

Plasma neurofilament light for early differentiation of multiple system atrophy from Parkinson disease



Zhenwei Yu,^{a,*,m} Yuanchu Zheng,^{b,m} Wenyi Kou,^{c,m} Huihui Cai,^b Siming Li,^b Jiayi Wu,^b Chen Yang,^b Bingxu Zhu,^b Genliang Liu,^b Yuke Wang,^{d,e} Hao Tian,^f Xinning Zhang,^g Ying Zhou,^h Xuedong Liu,ⁱ Ying Chang,^j Zhigang Chen,^{k,****} Wei Mao,^{d,e,***} and Tao Feng^{b,l,*}



^aDepartment of Pathophysiology, Beijing Neurosurgical Institute, Capital Medical University, Beijing, China

^bDepartment of Neurology, Centre for Movement Disorders, Beijing Tiantan Hospital, Capital Medical University, Beijing, China

^cDepartment of Neurology, Beijing Hospital, National centre of Gerontology; Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing, China

^dDepartment of Neurology, Xuanwu Hospital, Capital Medical University, Beijing, China

^eNational centre for Neurological Disorders, Beijing, China

^fBeijing University of Chinese Medicine, 100029, Beijing, China

^gXi'an Hospital of Traditional Chinese Medicine, 710021, Xi'an, China

^hDepartment of Traditional Chinese Medicine, The Second Affiliated Hospital of Army Medical University, Chongqing, 400030, China

ⁱDepartment of Neurology, Xijing Hospital of Air Force Military Medical University, Xi'an, China

^jDepartment of Neurology, China-Japan Union Hospital of Jilin University, Changchun, China

^kNeurology Department One, Dongfang Hospital, Beijing University of Chinese Medicine, No. 6, Fangxingyuan Community, Fangzhuang, Fengtai District, Beijing, 100078, China

^lChina National Clinical Research centre for Neurological Diseases, Beijing, China

Summary

Background Early differentiation of multiple system atrophy (MSA) from Parkinson's disease (PD) remains difficult, particularly within two years of symptom onset, when diagnostic uncertainty has major implications for prognosis, referral, and trial enrolment. Because MSA is rare, previous biomarker studies have often been limited by relatively small samples and restricted multicentre validation. We aimed to determine the diagnostic performance of plasma neurofilament light chain (NfL) for differentiating MSA from PD and to assess the incremental value of glial fibrillary acidic protein (GFAP), total tau (t-tau), and phosphorylated tau at threonine 217 (p-tau217).

Methods In this multicentre cross-sectional diagnostic study, participants were enrolled from five movement-disorder referral centres in China between Jan 1, 2018, and June 30, 2024, and were divided by enrolment period into discovery and temporally separated validation datasets. Plasma NfL, GFAP, t-tau, and p-tau217 were measured using light-initiated chemiluminescence assays. Group comparisons used age- and sex-adjusted models, and discrimination was assessed using receiver-operating-characteristic analysis with sensitivity, specificity, predictive values, and robustness analyses.

Findings The analysis included 2408 participants: 782 (32.5%) with PD, 796 (33.1%) with MSA, and 830 (34.5%) healthy controls. NfL was the best single biomarker for differentiating MSA from PD in the discovery dataset (AUC 0.920, 95% CI 0.903–0.936). The discovery-derived cutoff of 41.3 pg/mL yielded sensitivity of 90.5% and specificity of 83.4% in the discovery dataset, and sensitivity of 86.7% and specificity of 85.5% in the validation dataset (AUC 0.924, 95% CI 0.898–0.948). In the early-stage subgroup, NfL retained strong performance (AUC 0.943, 95% CI 0.915–0.966). Integrated multimarker models provided limited incremental discrimination over NfL alone.

Interpretation Plasma NfL may support the differentiation between clinically diagnosed MSA and PD, including early in the disease course. The limited added value of multimarker panels supports a simpler and more immediately translatable NfL-first strategy for diagnostically uncertain parkinsonism.

eBioMedicine
2026;131: 106461
Published Online xxx
<https://doi.org/10.1016/j.ebiom.2026.106461>

*Corresponding author. Department of Neurology, Centre for Movement Disorders, Beijing Tiantan Hospital, Capital Medical University, Beijing, China.

**Corresponding author.

***Corresponding author. Department of Neurology, Xuanwu Hospital, Capital Medical University, No. 45 Changchun Street, Xicheng District, Beijing, 100053, China.

****Corresponding author.

E-mail addresses: bxkyjs@sina.com (T. Feng), jasonyzw@163.com (Z. Yu), wei.mao@ccmu.edu.cn (W. Mao), chenzhigang64@126.com (Z. Chen).

^mThese authors contributed equally.

Funding National Natural Science Foundation of China; the Capital's Fund for Health Improvement and Research; Beijing Natural Science Foundation; Beijing Municipal Science and Technology Commission; Beijing Neurosurgical Institute; Beijing Traditional Chinese Medicine Science and Technology Development Fund.

Copyright © 2026 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Parkinsonism; Multiple system atrophy; Biomarkers

Research in context

Evidence before this study

Blood neurofilament light chain (NfL) has consistently emerged as one of the most promising fluid biomarkers for differentiating atypical parkinsonian syndromes from Parkinson's disease, with multiple system atrophy (MSA) typically showing the highest concentrations. However, previous studies have often been limited by small single-centre samples, relatively few early-stage MSA cases, restricted external validation, and uncertainty about how blood NfL might be used in practice across different prevalence settings.

Added value of this study

This study provides a large, focused, 5-centre clinical evaluation of plasma NfL for MSA-PD differentiation, including 796 patients with MSA, a temporally separated validation dataset, and an early-stage subgroup. We

additionally assessed the robustness of the plasma NfL findings using disease-duration-adjusted and disease-duration-matched analyses, centre stratification, cognitive stratification, orthostatic-hypotension summaries, MSA-P/MSA-C analyses, and exploratory PSP/CBS comparators.

Implications of all the available evidence

The available evidence supports plasma NfL as the leading blood-based biomarker for differentiating MSA from Parkinson's disease. Our findings suggest NfL as a triage biomarker in diagnostically uncertain parkinsonism, particularly early in the disease course, and that a simple 'NfL-first' strategy may be more scalable and clinically translatable than more complex multimarker panels. Longitudinal and pathology-confirmed studies are needed to define prognostic utility and broader applicability across atypical parkinsonian syndromes.

Introduction

Distinguishing multiple system atrophy (MSA) from Parkinson's disease (PD) remains a major unmet need in clinical neurology, particularly early in the disease course, when overlapping motor and non-motor features often lead to diagnostic uncertainty.^{1–3} Both conditions can manifest with parkinsonism, leading to frequent initial misdiagnosis.^{4,5} This distinction matters because an incorrect or delayed diagnosis affects prognostic counselling, subsequent investigations, and eligibility for trials of disease-modifying therapies. Although fluid and imaging biomarkers have improved the biological characterisation of parkinsonian disorders, clinically scalable tools that can support early PD and MSA differentiation in routine practice remain limited.⁶

Current diagnostic approaches rely on clinical criteria, which have limited sensitivity in early disease.⁷ Dopamine transporter imaging can help differentiate parkinsonian syndromes from non-degenerative conditions, but it often fails to reliably distinguish MSA from PD.^{8,9} Biomarkers in cerebrospinal fluid (CSF), such as seed-amplification assays for α -synuclein, have shown considerable promise for improving diagnostic accuracy,^{10–13} but lumbar puncture limits scalability. Therefore, blood-based biomarkers represent a less

invasive alternative, yet their utility for differentiating MSA from PD remains to be definitively established.¹⁴

Recent advances in ultrasensitive immunoassays have enabled reliable quantification of neurofilament light chain (NfL) in blood and CSF, providing an accessible marker of neuroaxonal injury. Multiple cohorts have demonstrated that circulating NfL is consistently higher in atypical parkinsonian syndromes, including MSA, progressive supranuclear palsy (PSP), and corticobasal syndrome (CBS), than in PD, including in relatively early disease stages.^{14–17} Recent meta-analyses further support this concept, reporting strong overall diagnostic accuracy for blood NfL in differentiating PD from atypical parkinsonism and consistently ranking MSA as the condition with the highest NfL elevations.^{18,19} Nevertheless, between-study heterogeneity (assay platforms, preanalytical handling) has limited translation into harmonised reference intervals and clinically actionable cutoffs across laboratories and populations.²⁰

Glial fibrillary acidic protein (GFAP), a marker of astroglial activation, has also been associated with synucleinopathies and may complement NfL for differential diagnosis and prognosis; however, its incremental value for separating PD from MSA remains uncertain and potentially cohort-dependent.^{21,22}

Total tau (t-tau) and phosphorylated tau at threonine 217 (p-tau217) are plasma biomarkers of neuronal injury and Alzheimer-type tau pathology, respectively. Plasma NfL can also be modestly elevated in mild cognitive impairment and Alzheimer's disease (AD), where it reflects neurodegeneration rather than AD-specific amyloid or tau pathology.²³ Because concomitant amyloid/tau pathology can occur in older individuals and in PD with cognitive impairment, incorporating p-tau217 may help identify comorbid Alzheimer's disease (AD) pathology that could confound the interpretation of neurodegeneration markers (including NfL) and influence clinical phenotyping.^{24,25} However, in the absence of systematic CSF or PET AD-biomarker stratification, p-tau217 was not considered a primary discriminator of MSA from PD.

Despite evidence that blood NfL differentiates PD from atypical parkinsonism, several barriers still limit routine clinical adoption. The rarity of MSA has constrained many studies to small samples, few clinically early MSA cases, and limited validation across centres. In addition, assay and preanalytical heterogeneity has also hindered transferable cutoffs.^{17–20}

Using a large multicentre cross-sectional dataset with a discovery-validation design and an early-stage subgroup, we aimed to evaluate the diagnostic performance of plasma NfL for PD-MSA differentiation and the incremental value of GFAP, t-tau, and p-tau217.

Methods

Study design and participants

This multicentre cross-sectional study enrolled participants from 5 movement-disorder centres in China between January 2018 and June 2024: Beijing Tiantan Hospital, Xuanwu Hospital, Dongfang Hospital, Xijing Hospital of Air Force Military Medical University, and China–Japan Union Hospital of Jilin University. The discovery dataset comprised enrolments from Beijing Tiantan and Xuanwu hospitals between January 2018 and December 2021; the temporally separated validation dataset comprised later enrolments from all five centres between January 2022 and June 2024. Centre-specific contributions and demographic and clinical summaries are reported in [Supplementary Tables S1 and S2](#).

Eligible patients met established Movement Disorder Society criteria for clinically probable or clinically established PD or MSA and were diagnosed by at least 2 movement-disorder specialists.^{2,3} The study reference standard was expert clinical diagnosis rather than neuropathological confirmation. A prespecified early-stage subgroup was defined as disease duration of two years or less from motor symptom onset. These participants underwent follow-up assessment one to two years after enrolment; those whose expert clinical diagnosis changed were excluded from the early-stage

analysis to avoid evaluating a fixed baseline label that was subsequently judged incorrect. This strategy improves label stability but may reduce generalisability to patients with evolving or indeterminate diagnoses.

Exclusion criteria were other major neurological disorders (eg, stroke, dementia, or epilepsy), severe systemic disease, or clinician-documented renal impairment or major vascular/metabolic conditions judged at enrolment likely to materially affect circulating biomarker concentrations. Uniform eGFR or disease-severity thresholds were not available across centres; these eligibility exclusions therefore cannot be interpreted as laboratory-threshold-based adjustment. Healthy controls had no neurological signs or cognitive impairment and no self-reported family history of movement disorders. The study was approved by the institutional review board of Beijing Tiantan Hospital (KY2023-285-03) and was conducted in accordance with the approved protocol, the principles of the Declaration of Helsinki, and applicable institutional and national ethical guidelines. Written informed consent from all participants was obtained.

Clinical assessments

Motor severity was assessed using the MDS-Unified Parkinson's Disease Rating Scale Part III (MDS-UPDRS III). For patients with MSA, disease-specific impairment and functional disability were evaluated using the MDS-UPDRS III and Unified MSA Rating Scale (UMSARS). Cognitive function was assessed using the Mini-Mental State Examination (MMSE) and the Montreal Cognitive Assessment (MoCA). Demographic information and disease duration were collected through structured interviews and a review of medical records. Sex was self-reported by study participants. MSA subtype and orthostatic hypotension were analysed when explicitly recorded. Harmonised variables for urinary dysfunction, erectile dysfunction, MRI supportive findings, cardiac MIBG imaging, formal autonomic testing, PD-MCI, Parkinson's disease dementia, and CSF/PET AD biomarkers were not available across centres.

Plasma collection and biomarker measurement

Venous plasma was collected in EDTA tubes, processed within 4 h, and stored at -80°C . Plasma NfL, GFAP, t-tau, and p-tau217 were measured on the LiCA 800 platform (Chemclin, Beijing, China) using light-initiated chemiluminescence assays (LiCA[®]). All four biomarkers were measured using corresponding reagent kits (Lot No.: 202408). Laboratory operators were blinded to clinical diagnosis. Assay analytical performance was validated according to CLSI guidance; detailed analytical metrics are provided in [Supplementary Table S3](#).

Statistical analysis

Continuous variables were inspected for distributional form and are summarised as median (IQR) because the

principal biomarker and clinical variables were non-normally distributed; categorical variables are reported as counts and percentages. Biomarker concentrations were natural log-transformed for modelling. Analyses used available-case data for descriptive summaries and complete cases for each model; biomarker missingness and non-positive values are reported in [Supplementary Table S4](#). Group differences were assessed using linear models adjusted for age and sex, with Holm correction for pairwise comparisons within each biomarker.

Diagnostic performance was evaluated using ROC analysis, AUCs, sensitivity, and specificity. The primary NfL cutoff was derived in the discovery dataset using the Youden index and was applied unchanged to the temporally separated validation dataset and fixed-cutoff sensitivity analyses. Confidence intervals for AUC, sensitivity, and specificity were estimated using 2000 bootstrap resamples. An integrated logistic model included log NfL, log GFAP, log t-tau, log p-tau217, age, and sex. AUC comparisons used DeLong tests with Holm correction. Positive and negative predictive values were calculated at assumed MSA prevalences of 10%, 20%, and 30%. Decision curve analysis (DCA) was prevalence-recalibrated and treated as exploratory because the dataset was case-enriched.

To assess disease-duration confounding, we fitted a complete-case multivariable linear model with log NfL as the dependent variable and diagnosis, age, sex, disease duration, and recruitment centre as predictors. We also performed 1:1 nearest-neighbour matching of MSA to PD on disease duration without replacement using a 0.5-year calliper; the discovery-derived cutoff was applied unchanged. Additional analyses evaluated MSA-P and MSA-C, centre-stratified performance, cognitive strata defined by MoCA <26, orthostatic-hypotension summaries, and exploratory PSP/CBS comparators.

Spearman correlations were retained as descriptive exploratory analyses and were corrected using the Benjamini-Hochberg false-discovery rate (FDR). Inter-biomarker and biomarker-clinical associations were also examined in separate multivariable linear regression models within PD and MSA. Continuous outcomes and biomarker predictors were standardised, models adjusted for age, sex, and disease duration, and HC3 heteroskedasticity-robust 95% CIs were used. FDR correction was applied separately to the inter-biomarker and biomarker-clinical model families. Complete adjusted results are provided in the [Supplementary Materials](#). Further analytical details are provided in the [Supplementary Methods](#).

Role of funders

The Funders had no role in study design, data collection, data analyses, interpretation, or writing of report.

Results

Participants

We enrolled 2408 participants: 782 (32.5%) with PD, 796 (33.1%) with MSA, and 830 (34.5%) healthy controls (HCs). The discovery dataset included 1685 participants (PD = 548 [32.5%], MSA = 556 [33.0%], HC = 581 [34.5%]) and the validation dataset included 723 participants (PD = 234 [32.4%], MSA = 240 [33.2%], HC = 249 [34.4%]). Participant characteristics are summarised in [Table 1](#). Centre contributions are reported in [Supplementary Table S1](#); centre-level age, sex, disease duration, MoCA, and NfL summaries are reported in [Supplementary Table S2](#).

The early-stage subgroup included 484 patients with disease duration of two years or less (PD, 120 [24.8%]; MSA, 364 [75.2%]) after exclusion of participants whose expert diagnosis changed during one to two years of follow-up. Available exploratory comparators comprised 50 PSP and 10 CBS participants and were analysed only in the [Supplementary Materials](#). Explicit MSA subtype was available for 672 participants (351 MSA-P and 321 MSA-C); 124 MSA participants without an explicit P/C label were retained in overall analyses but excluded from subtype analyses.

MoCA was available for 1540 of 1578 participants with PD or MSA. Using MoCA <26, 1192 (77.4%) were classified as cognitively impaired and 348 (22.6%) as cognitively unimpaired; PD-MCI and Parkinson's disease dementia were not documented as harmonised diagnostic categories. Orthostatic-hypotension status was available for 772 of 796 MSA participants and 775 of 782 PD participants; orthostatic hypotension was recorded in 505/772 (65.4%) MSA and 218/775 (28.1%) PD participants.

Discovery analysis

In the discovery dataset, with adjustment for age and sex, plasma NfL, GFAP, t-tau, and p-tau217 were significantly higher in MSA than in PD. The adjusted beta for NfL was 0.862, compared with 0.112 for GFAP, 0.061 for t-tau, and 0.067 for p-tau217 ([Supplementary Table S5](#); [Fig. 1A–D](#)). NfL was the best single discriminator of MSA from PD (AUC, 0.920; 95% CI 0.903–0.936), with an optimal cutoff of 41.3 pg/mL yielding sensitivity of 90.5% (95% CI 88.1–92.8) and specificity of 83.4% (80.3–86.4). GFAP, t-tau, and p-tau217 had AUCs of 0.576, 0.565, and 0.562, respectively. The integrated model AUC was 0.931; its improvement over NfL alone was small (DeLong Holm-adjusted $P < 0.001$; [Supplementary Tables S6 and S7](#); [Fig. 1E and F](#)).

Validation analysis

Findings were similar in the temporally separated validation cohort. After adjustment for age and sex, MSA showed higher plasma NfL, GFAP, and t-tau levels than PD and HCs, whereas p-tau217 did not differ significantly between MSA and PD ([Fig. 2A–D](#)). NfL AUC for MSA versus PD was 0.924 (95% CI

Characteristic	MSA	PD	Healthy controls	P value
Discovery dataset	n = 556	n = 548	n = 581	
Age, median (IQR), years	61.00 (54.25–66.00)	62.00 (56.00–66.00)	60.00 (56.00–66.00)	0.445
Sex ^a , Female, n (%)	268 (48.4%) ^b	257 (46.9%)	285 (49.1%)	0.761
Disease duration, median (IQR), years	2.06 (1.50–3.13)	6.00 (3.00–8.00)	NA	<0.001
MMSE, median (IQR)	27.00 (24.00–28.00)	27.00 (24.00–29.00)	29.00 (28.00–29.00)	<0.001
MoCA, median (IQR)	22.00 (18.00–24.00)	22.00 (19.00–25.00)	NA	0.012
MDS-UPDRS III, median (IQR)	38.00 (26.00–51.00)	38.00 (26.00–48.00)	NA	0.828
UMSARS, median (IQR)	37.00 (28.00–48.00)	NA	NA	NA
Temporally separated validation dataset	n = 240	n = 234	n = 249	
Age, median (IQR), years	60.00 (55.00–66.00)	62.00 (55.00–66.00)	59.00 (56.00–65.00)	0.442
Sex ^a , Female, n (%)	115 (47.9%)	102 (43.6%)	118 (47.4%)	0.588
Disease duration, median (IQR), years	2.00 (1.50–4.00)	6.00 (3.00–8.00)	NA	<0.001
MMSE, median (IQR)	26.00 (23.00–29.00)	27.00 (25.00–29.00)	29.00 (27.00–29.00)	<0.001
MoCA, median (IQR)	22.00 (19.00–26.00)	23.00 (19.00–26.00)	NA	0.508
MDS-UPDRS III, median (IQR)	35.00 (24.00–49.00)	39.00 (27.00–48.00)	NA	0.184
UMSARS, median (IQR)	40.00 (29.00–49.00)	NA	NA	NA
Early-stage analysis (disease duration ≤2 years)	n = 364	n = 120	NA	
Age, median (IQR), years	59.50 (54.00–65.00)	61.00 (54.00–67.00)	NA	0.578
Sex ^a , Female, n (%)	165 (45.3%)	64 (53.3%)	NA	0.156
Disease duration, median (IQR), years	1.42 (1.00–2.00)	2.00 (1.00–2.00)	NA	0.016
MMSE, median (IQR)	27.00 (24.00–29.00)	28.00 (26.00–29.00)	NA	0.003
MoCA, median (IQR)	22.00 (18.00–25.00)	23.00 (20.00–26.00)	NA	0.049
MDS-UPDRS III, median (IQR)	31.00 (22.00–44.00)	27.00 (21.00–36.00)	NA	0.023
UMSARS, median (IQR)	33.00 (26.00–41.50)	NA	NA	NA

Values are based on non-missing observations. Continuous variables are median (IQR); categorical variables are n (% of available observations). P values use Kruskal–Wallis tests for continuous variables and chi-squared tests for categorical variables. MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; MDS-UPDRS III, Movement Disorder Society–Unified Parkinson's Disease Rating Scale Part III; UMSARS, Unified Multiple System Atrophy Rating Scale; NA, not applicable. ^aSex was self-reported by study participants. ^bSex was missing for two discovery-dataset MSA participants.

Table 1: Demographic and clinical characteristics of participants in the discovery, validation, and early-stage analyses.

0.898–0.948). Applying the discovery-derived cutoff of 41.3 pg/mL yielded sensitivity of 86.7% (82.3–90.9), and specificity of 85.5% (80.9–89.6) for MSA vs PD (Fig. 2E). The integrated model AUC was 0.930 and did not differ significantly from NfL alone on the same complete-case sample (Holm-adjusted $P = 0.0849$; [Supplementary Tables S6 and S7](#); Fig. 2E). For MSA versus HCs, NfL and the integrated model yielded AUCs of 0.977 (95% CI 0.964–0.989) and 0.984 (95% CI 0.972–0.994), respectively (Fig. 2F).

Because the referral-enriched study prevalence was not considered representative of routine neurology practice, predictive values were calculated at assumed MSA prevalence of 10%, 20%, and 30%. In the validation dataset, PPV was 39.9%, 59.9%, and 71.9%, respectively, and NPV was 98.3%, 96.2%, and 93.7% ([Supplementary Table S8](#)). Prevalence-recalibrated DCA further suggested higher net benefit for NfL than default strategies across portions of the clinically relevant threshold range, while the NfL and integrated-model curves largely overlapped; these exploratory results are shown in [Supplementary Figure S1](#).

Early-stage disease (≤2 years)

In the early-stage subgroup, MSA exhibited higher plasma NfL than PD (Holm-adjusted $P < 0.001$), though

GFAP ($P = 0.192$), t-tau ($P = 0.192$), and p-tau217 ($P = 0.574$) did not differ after Holm correction (Fig. 3A–D). NfL AUC was 0.943 (95% CI 0.915–0.966). The early-stage Youden threshold was 41.34 pg/mL, nearly identical to the discovery-derived cutoff; using the fixed 41.3 pg/mL cutoff yielded sensitivity of 90.4% and specificity of 88.3% ([Supplementary Table S9](#); Fig. 3E).

The early-stage subgroup was enriched for MSA because all participating centres are tertiary referral centres for atypical parkinsonism and suspected MSA frequently prompts early referral. Excluding patients whose diagnosis changed during follow-up also selected a diagnostically more stable subset. Early-stage performance estimates are therefore interpreted in light of this referral and selection pattern.

Robustness and phenotypic analyses

Before disease-duration matching, median disease duration was 2.05 years (IQR 1.50–3.58) in MSA and 6.00 years (3.00–8.00) in PD. In the complete-case multivariable model, disease duration was independently associated with log NfL (beta = 0.008 per year, 95% CI 0.002–0.014; $P = 0.0129$). After adjustment for age, sex, disease duration, and centre, MSA remained strongly associated with higher log NfL than PD

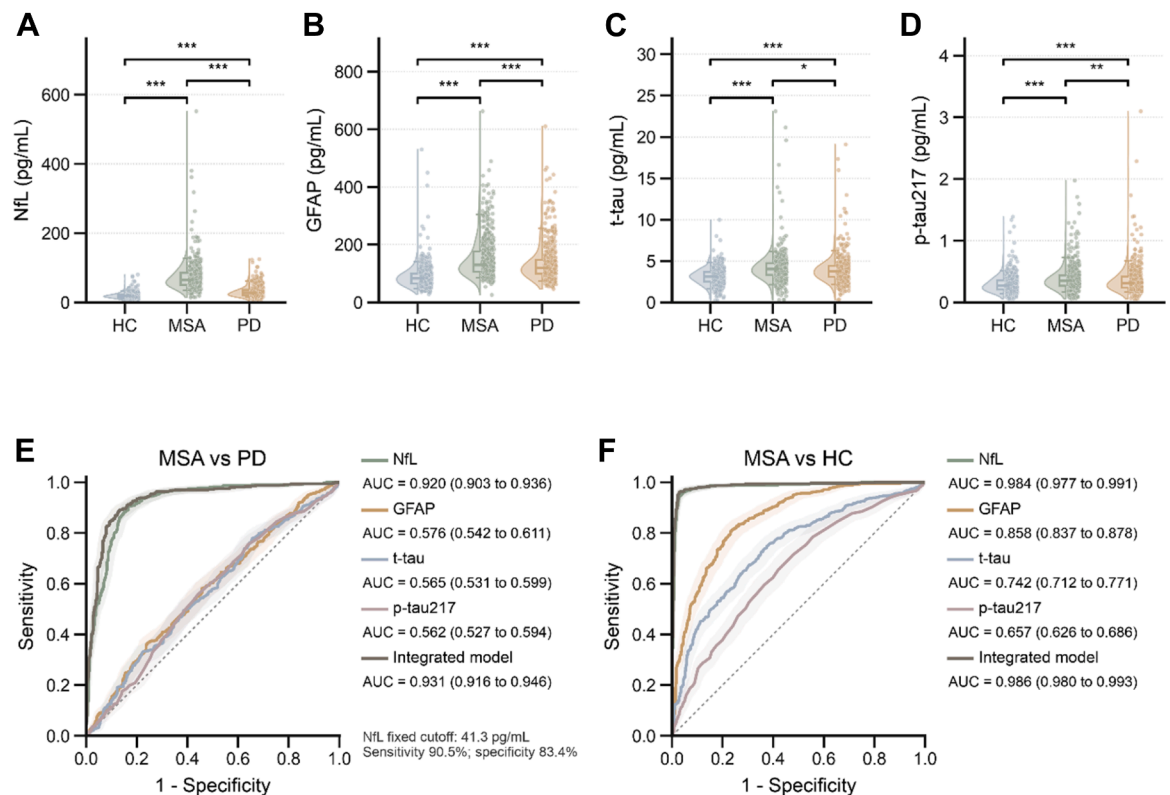


Fig. 1: Plasma biomarkers in the discovery dataset and diagnostic performance. Plasma NfL (A), GFAP (B), t-tau (C), and p-tau217 (D) in MSA, PD, and healthy controls. ROC curves for MSA versus PD (E) and MSA versus healthy controls (F) compare the four biomarkers and an integrated model comprising log-transformed biomarkers, age, and sex. AUCs and bootstrap 95% CIs are shown; sensitivity and specificity in E use the discovery-derived NfL cutoff of 41.3 pg/mL. Asterisks denote age- and sex-adjusted pairwise comparisons after Holm correction: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; NS, not significant.

(adjusted beta = 0.891, 95% CI 0.843–0.939; $P < 0.001$; $n = 1531$; [Supplementary Table S10](#)).

The 1:1 disease-duration-matched analysis included 396 MSA and 396 PD participants. Median disease duration was 3.00 years in both groups (IQR 1.74–5.00 for MSA and 2.00–5.00 for PD). In the matched sample, plasma NfL AUC was 0.925 (95% CI 0.905–0.944). At the unchanged discovery-derived cutoff, sensitivity was 89.6% (86.7–92.6) and specificity was 85.4% (82.0–89.0; [Supplementary Table S11](#) and [Fig. 3F](#)).

NfL discriminated both MSA-P and MSA-C from PD. For MSA-P versus PD, AUC was 0.926 and the fixed cutoff yielded sensitivity of 91.5% and specificity of 84.0%; for MSA-C versus PD, AUC was 0.915, sensitivity was 87.9%, and specificity was 84.0%. In an age-, sex-, and disease-duration-adjusted model, log NfL did not differ between MSA-C and MSA-P (beta = -0.038, 95% CI -0.107 to 0.031; $P = 0.275$; $n = 635$; [Supplementary Table S12](#) and [Supplementary Figure S2](#)).

Centre-stratified NfL AUCs using the same fixed cutoff ranged from 0.806 to 0.959. Estimates were broadly consistent in the four centres with at least 63

PD/MSA participants, whereas the Jilin estimate ($n = 12$; AUC 0.806, 95% CI 0.438–1.000) was imprecise due to limited patient number. In the multivariable NfL model, individual centre terms were not statistically significant after adjustment for diagnosis, age, sex, and disease duration ([Supplementary Tables S10 and S13](#)).

Cognitive-stratified performance remained strong among participants classified as cognitively impaired (AUC 0.910) and cognitively unimpaired (AUC 0.960; [Supplementary Table S14](#)). Orthostatic-hypotension availability, prevalence, and descriptive NfL distributions are reported in [Supplementary Table S15](#). In exploratory comparisons, plasma NfL was higher in MSA, PSP, and CBS than in PD after adjustment for age and sex (all Holm-adjusted $P < 0.001$; [Supplementary Table S16](#) and [Supplementary Figure S3](#)); PSP and CBS were not used to derive or validate the MSA cutoff.

Biomarker associations

Spearman analyses showed multiple statistically significant but generally weak biomarker-clinical correlations ([Supplementary Table S17](#); [Fig. 4A and B](#)).

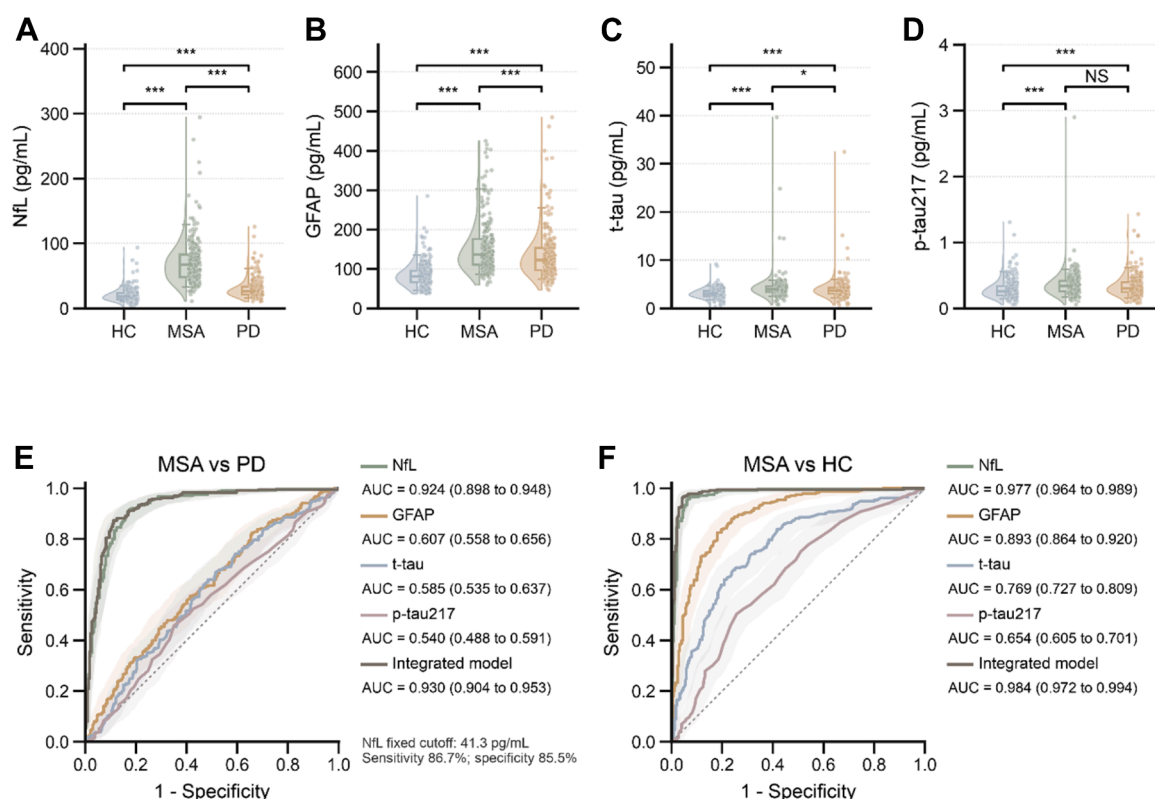


Fig. 2: Plasma biomarkers and diagnostic performance in the temporally separated validation dataset. Plasma NfL (A), GFAP (B), t-tau (C), and p-tau217 (D) in MSA, PD, and healthy controls. ROC curves for MSA versus PD (E) and MSA versus healthy controls (F) compare the four biomarkers and the integrated model. AUCs and bootstrap 95% CIs are shown; sensitivity and specificity in E use the unchanged discovery-derived NfL cutoff of 41.3 pg/mL. Asterisks denote age- and sex-adjusted pairwise comparisons after Holm correction: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; NS, not significant.

In PD, NfL concentrations were weakly associated with MMSE (Fig. 4C; $\rho = -0.261$; exact $P = 1.56 \times 10^{-13}$) and MoCA ($\rho = -0.269$; exact $P = 3.08 \times 10^{-14}$). In MSA, higher GFAP correlated with greater motor and functional impairment on MDS-UPDRS III and UMSARS (Fig. 4D; MDS-UPDRS III: $\rho = 0.153$; exact $P = 6 \times 10^{-5}$; UMSARS: $\rho = 0.294$; exact $P = 2.78 \times 10^{-8}$). However, these effect sizes do not support strong severity-marker claims.

In age-, sex-, and disease-duration-adjusted models, associations remained small. In PD, NfL was associated with lower MMSE (standardised beta = -0.182 , 95% CI -0.265 to -0.099 ; $P = 1.73 \times 10^{-5}$; $q = 9.68 \times 10^{-5}$) and MoCA (beta = -0.185 , -0.264 to -0.107 ; $P = 3.97 \times 10^{-6}$; $q = 2.78 \times 10^{-5}$). In MSA, NfL was associated with MDS-UPDRS III (beta = 0.171 , 0.105 – 0.237 ; $P = 3.69 \times 10^{-7}$; $q = 1.03 \times 10^{-5}$), and GFAP was associated with UMSARS (beta = 0.235 , 0.135 – 0.334 ; $P = 3.9 \times 10^{-6}$; $q = 2.78 \times 10^{-5}$). Adjusted inter-biomarker beta estimates ranged from 0.127 to 0.610 . Complete adjusted results are reported in [Supplementary Table S17](#).

Discussion

In this large multicentre study of 2408 participants, plasma NfL showed robust, reproducible performance for differentiating clinically diagnosed MSA from PD in discovery and temporally separated validation datasets. The finding remained robust in disease-duration adjustment and matching, centre stratification, cognitive stratification, and MSA-P/MSA-C analyses. GFAP, t-tau, and p-tau217 added little additional discrimination beyond NfL alone.

The NfL finding is consistent with multicentre studies and meta-analyses showing higher plasma NfL in atypical parkinsonian disorders than in PD.^{14–20} NfL reflects neuroaxonal injury rather than a disease-specific proteinopathy.²⁶ The widespread striatonigral, olivopontocerebellar, and autonomic degeneration of MSA provides a biological rationale for greater NfL release than the more topographically restricted neurodegeneration typical of early PD.¹ However, elevated NfL in PSP, CBS, and other disorders limits disease specificity.^{14,17,19}

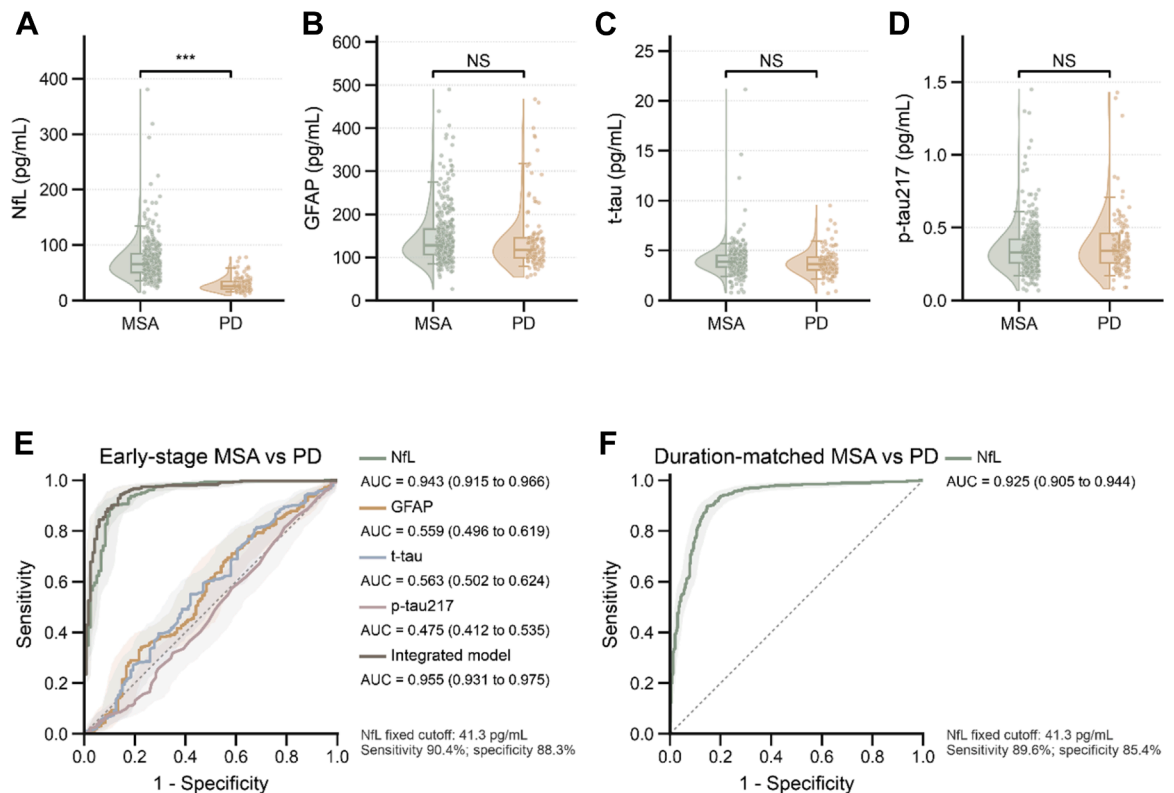


Fig. 3: Early-stage and disease-duration-matched analyses of MSA versus PD. Plasma NfL (A), GFAP (B), t-tau (C), and p-tau217 (D) in the early-stage subgroup, defined by disease duration of 2 years or less. (E) ROC curves for the four biomarkers and the integrated model in the early-stage subgroup. (F) NfL ROC curve in the disease-duration-matched sample of 396 MSA and 396 PD participants. AUCs and bootstrap 95% CIs are shown; sensitivity and specificity use the fixed discovery-derived NfL cutoff of 41.3 pg/mL. ***, Holm-adjusted $P < 0.001$; NS, not significant.

Although GFAP, t-tau, and p-tau217 were elevated in MSA, they offered limited value for PD-MSA discrimination. Plasma NfL can be modestly elevated in AD as a marker of neurodegeneration,²³ whereas p-tau217 more directly reflects Alzheimer-type amyloid/tau pathology.^{24,25} Without systematic CSF or PET AD biomarkers, p-tau217 can only be interpreted as an exploratory contextual marker in this study.

Within specialist referral settings, plasma NfL may have an adjunctive triage-support role when the distinction between PD and MSA remains uncertain. Its clinical utility depends on the pre-test probability and clinical spectrum of the population being tested. Accordingly, we acknowledge that the 41.3 pg/mL threshold should not be used as a standalone rule-in or rule-out test, and its performance in unselected routine-neurology populations remains to be established.

The disease-duration-adjusted and duration-matched analyses addressed an important potential confounder. The persistence of the MSA-PD NfL difference after multivariable adjustment, together with strong discrimination in the matched sample, indicates

that the primary finding was not explained solely by measured disease duration. However, disease duration is an imperfect proxy for neurodegeneration severity and progression tempo, and residual confounding remains possible in this cross-sectional referral-centre dataset.

Correlation analyses suggested potentially distinct associations in PD and MSA. In PD, higher NfL correlated with worse cognitive scores (MMSE/MoCA), suggesting that, though warranting further investigation, elevated NfL may have a role in tracking the evolution of non-motor symptoms in this population. In contrast, these correlations with cognitive scores were substantially weaker in patients with MSA, potentially because plasma NfL levels are already significantly elevated at an early stage of the disease, which may obscure a more direct relationship with cognitive status.

Another observation was that GFAP correlated with greater motor and functional impairment (MDS-UPDRS III and UMSARS) in MSA, but not in PD, suggesting that while GFAP is not a suitable differential diagnostic marker, it may have value as a tool for

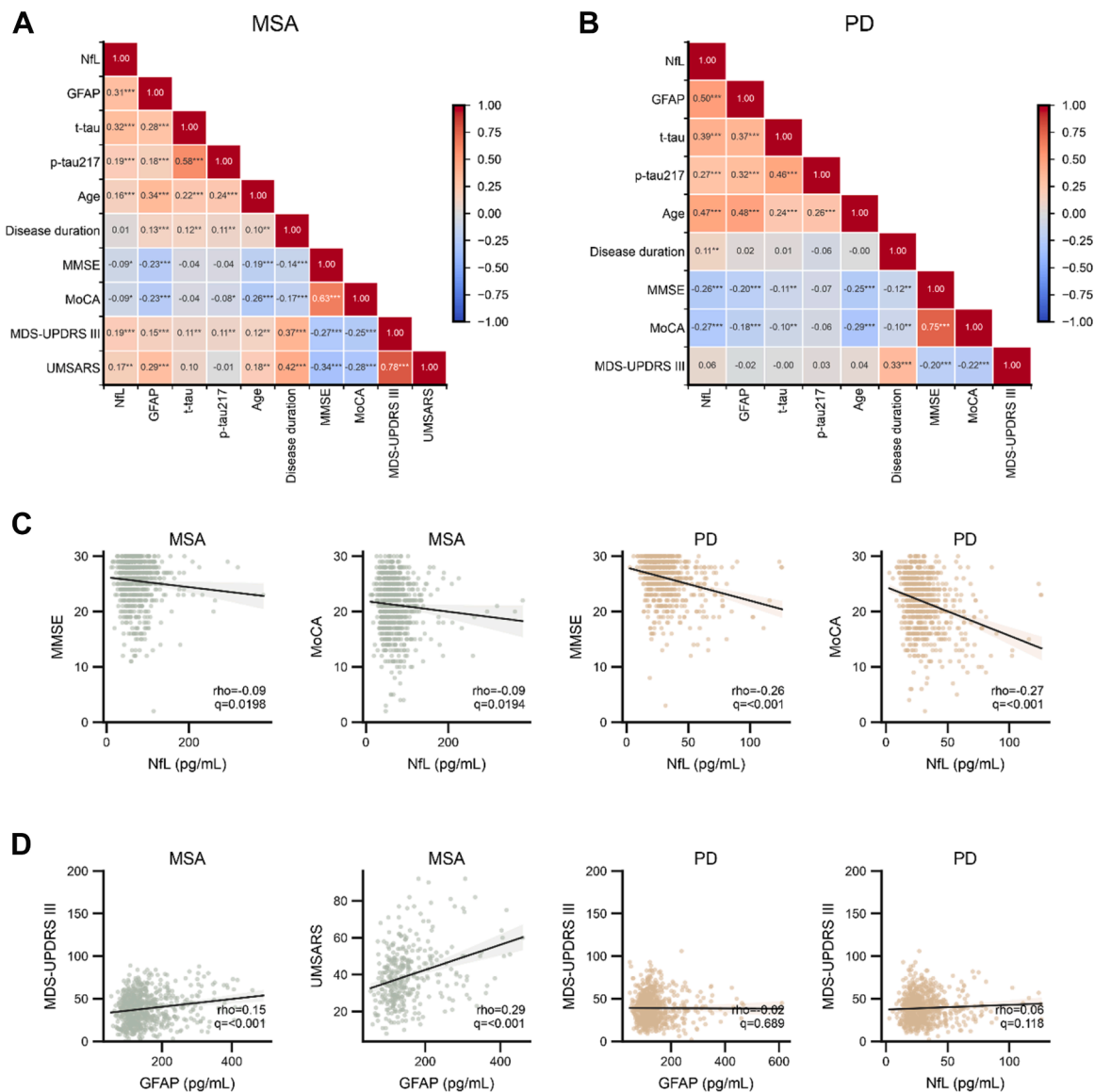


Fig. 4: Exploratory associations between plasma biomarkers and clinical measures. Spearman correlation heatmaps in MSA (A) and PD (B); cells show Spearman ρ , with significance based on Benjamini-Hochberg FDR-adjusted q values. (C) Associations of NfL with MMSE and MoCA in MSA and PD. (D) Associations of GFAP with motor and functional measures in MSA, and of GFAP or NfL with MDS-UPDRS III in PD. Lines show linear fits with 95% CIs; annotations report Spearman ρ and FDR-adjusted q values. *, $q < 0.05$; **, $q < 0.01$; ***, $q < 0.001$.

monitoring disease severity or progression specifically in MSA. These differential correlations suggest that NfL and GFAP not only distinguish the two disorders but also potentially reflect independent aspects of disease severity and progression within each unique synucleinopathy. However, the association analyses identified statistically significant but limited effects, and NfL and GFAP should therefore not be described as established progression or pharmacodynamic biomarkers on the basis of these data.

Notably, recent evidence showed that LiCA and Simoa present good cross-platform comparability in

measuring plasma biomarkers including NfL.²⁷ Accordingly, the cutoff value of 41.3 pg/mL may serve as a useful reference for cross-platform translation, although its applicability across different laboratories, assay conditions, and populations still requires further validation.

Several limitations should be acknowledged. First, diagnoses were based on established clinical criteria rather than post-mortem neuropathological confirmation, which remains the gold standard. Although follow-up assessment was used to improve diagnostic certainty in the early-stage subgroup, diagnostic instability and

residual misclassification cannot be excluded, particularly early in the disease course. Given the known rate of diagnostic revision in parkinsonian disorders, some degree of misclassification is unavoidable.

Second, the cross-sectional design precludes conclusions regarding the longitudinal prognostic utility of these biomarkers. In addition, the temporally separated validation dataset was not an external validation dataset because it originated from the same national referral infrastructure and assay environment. Available PSP and CBS samples were small exploratory comparators, and systematically collected ataxia comparators were unavailable. Accordingly, the specificity of plasma NfL for MSA relative to other atypical syndromes remains incompletely defined.

Third, recruitment exclusively from tertiary movement-disorder referral centres introduces spectrum bias. The study population may overrepresent clinically well-characterised MSA and PD while underrepresenting patients with subtle, overlapping, or still-uncertain phenotypes encountered in routine neurology practice. The present performance estimates should therefore be interpreted as referral-centre estimates, and prospective validation in consecutively recruited routine-practice populations is required.

Fourth, low MoCA values in both MSA and PD probably reflect referral-centre case mix and cognitive heterogeneity. MoCA stratification showed that NfL discrimination persisted in impaired and unimpaired strata, but harmonised PD-MCI, Parkinson's disease dementia, and education-adjusted cognitive diagnoses were unavailable.

Fifth, orthostatic hypotension was the only autonomic feature available in harmonised form; urinary and erectile dysfunction, MRI, cardiac MIBG, and formal autonomic testing could not be analysed. Furthermore, although clinician-documented renal impairment and major vascular/metabolic comorbidities expected to affect circulating biomarkers were excluded at enrolment, uniform eGFR and disease-severity thresholds were not available and these factors could not be adjusted quantitatively. Residual effects of subclinical renal dysfunction or vascular/metabolic disease therefore cannot be excluded, limiting generalisability of the cutoff.^{26,28–30} Therefore, the proposed 41.3 pg/mL cutoff should be interpreted as an implementation-oriented reference rather than a universally transferable threshold and may require physiological recalibration before broader clinical adoption.³¹ Finally, systematic CSF or PET AD biomarkers were unavailable, limiting interpretation of p-tau217 and possible AD co-pathology.

Future work should prospectively enrol consecutive patients with diagnostically ambiguous parkinsonism in routine neurology settings; include pathology-confirmed or deeply phenotyped longitudinal follow-

up; broaden comparators to PSP, CBS, dementia with Lewy bodies, vascular parkinsonism, and degenerative or immune-mediated ataxias; evaluate renal, vascular, and metabolic covariates using harmonised thresholds; and test assay calibration across platforms and laboratories. Longitudinal studies are also required to determine whether within-person NfL change has value for prognosis, progression monitoring, or pharmacodynamic assessment.

In conclusion, this study provides evidence suggesting plasma NfL as an accurate and accessible biomarker for the differential diagnosis of MSA and PD, even in the earliest clinical stages. Clinically, a plasma NfL measurement above the threshold of 41.3 pg/mL in a patient presenting with parkinsonism should prompt intensified evaluation for MSA, including focused autonomic assessment and characteristic neuroimaging, whereas a value below this threshold supports PD as the more likely diagnosis in the appropriate clinical context. In research settings, plasma NfL could be evaluated as an enrichment biomarker for trials enrolling participants with clinically diagnosed MSA; longitudinal studies are required before it can be used for progression monitoring or pharmacodynamic assessment.

Contributors

Conceptualisation: Z.Y., Y.C.Z., W.K., and T.F. Methodology: Z.Y., Y.C.Z., and W.K. Resources: S.L., J.W., H.C., W.K., G.L., C.Y., B.Z., Y.C.Z., Y.W., H.T., X.L., Y.C., X.Z., Y.Z., Z.C., W.M., and T.F. Visualisation: Z.Y., W.K. Writing-original draft: Z.Y., and W.K. Writing-review and editing: Z.Y., W.K., Y.C.Z., and T.F. Z.Y., Y.C.Z., and W.K. accessed and verified the underlying data. All authors read and approved the final version of the manuscript.

Data sharing statement

De-identified individual participant data underlying the reported results will be available to qualified researchers from publication, with no prespecified end date, upon request to Tao Feng at bxkyjs@sina.com. Requests must include a scientifically sound proposal and are subject to investigator review, applicable ethical and institutional approvals, and a signed data-use agreement; approved data will be shared by secure electronic transfer.

Declaration of interests

We declare no competing interests.

Acknowledgements

This research was funded by the Capital's Fund for Health Improvement and Research (grant 2026-1Q-1112), the Beijing Natural Science Foundation (grant QG26024), the National Natural Science Foundation of China (grant 82501504, 82271459, U24A20694, 82071422, 82401480, and 82174328), the Beijing Science and Technology Plan Project (grant Z251100004625065), the Beijing Neurosurgical Institute (grant 11000025T000003319495-3), and the 2025 Beijing Traditional Chinese Medicine Science and Technology Development Fund (grant BJZYD-2025-01). The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

During the preparation of this work, we used ChatGPT for language editing. The authors reviewed and edited the output as needed and take full responsibility for the content of the published article.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.ebiom.2026.106461>.

References

- Krismer F, Fanciulli A, Meissner WG, Coon EA, Wenning GK. Multiple system atrophy: advances in pathophysiology, diagnosis, and treatment. *Lancet Neurol*. 2024;23(12):1252–1266.
- Postuma RB, Berg D, Stern M, et al. MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord*. 2015;30(12):1591–1601.
- Wenning GK, Stankovic I, Vignatelli L, et al. The movement disorder society criteria for the diagnosis of multiple system atrophy. *Mov Disord*. 2022;37(6):1131–1148.
- Miki Y, Tsushima E, Foti SC, et al. Identification of multiple system atrophy mimicking Parkinson's disease or progressive supranuclear palsy. *Brain*. 2021;144(4):1138–1151.
- Raty V, Kuusimäki T, Majuri J, et al. Stability and accuracy of a diagnosis of parkinson disease over 10 years. *Neurology*. 2025;104(9):e213499.
- Zarkali A, Thomas GEC, Zetterberg H, Weil RS. Neuroimaging and fluid biomarkers in Parkinson's disease in an era of targeted interventions. *Nat Commun*. 2024;15(1):5661.
- Sekiya H, Tipton PW, Kawazoe M, et al. Current landscape of clinical diagnosis in multiple system atrophy: a 15-Year analysis from 2008 to 2022. *Neurology*. 2024;103(11):e210021.
- Di Luca DG, Perlmutter JS. Time for clinical dopamine transporter scans in parkinsonism?: not DAT yet. *Neurology*. 2024;102(11):e209558.
- Ogawa T, Fujii S, Kuya K, et al. Role of neuroimaging on differentiation of Parkinson's disease and its related diseases. *Yonago Acta Med*. 2018;61(3):145–155.
- Okuzumi A, Hatano T, Matsumoto G, et al. Propagative alpha-synuclein seeds as serum biomarkers for synucleinopathies. *Nat Med*. 2023;29(6):1448–1455.
- Wiseman JA, Turner CP, Faull RLM, Halliday GM, Dieriks BV. Refining alpha-synuclein seed amplification assays to distinguish Parkinson's disease from multiple system atrophy. *Transl Neurodegener*. 2025;14(1):7.
- Poggiolini I, Gupta V, Lawton M, et al. Diagnostic value of cerebrospinal fluid alpha-synuclein seed quantification in synucleinopathies. *Brain*. 2022;145(2):584–595.
- Yu Z, Su D, Li S, Lin S, Ren K, Feng T. From phenotype to biology: a multi-modal roadmap of biofluid, tissue, imaging, and digital biomarkers in Parkinson's disease. *Transl Neurodegener*. 2026;15(1):25.
- Ashton NJ, Janelidze S, Al Khleifat A, et al. A multicentre validation study of the diagnostic value of plasma neurofilament light. *Nat Commun*. 2021;12(1):3400.
- Hansson O, Janelidze S, Hall S, et al. Blood-based NfL: a biomarker for differential diagnosis of parkinsonian disorder. *Neurology*. 2017;88(10):930–937.
- Lin CH, Li CH, Yang KC, et al. Blood NfL: a biomarker for disease severity and progression in Parkinson disease. *Neurology*. 2019;93(11):e1104–e1111.
- Quadalti C, Calandra-Buonaura G, Baiardi S, et al. Neurofilament light chain and alpha-synuclein RT-QuIC as differential diagnostic biomarkers in parkinsonisms and related syndromes. *NPJ Parkinsons Dis*. 2021;7(1):93.
- Demiri S, Veltsista D, Siokas V, et al. Neurofilament light chain in cerebrospinal fluid and blood in multiple system atrophy: a systematic review and meta-analysis. *Brain Sci*. 2025;15(3):241.
- Kou W, Li S, Yan R, Zhang J, Wan Z, Feng T. Cerebrospinal fluid and blood neurofilament light chain in Parkinson's disease and atypical parkinsonian syndromes: a systematic review and Bayesian network meta-analysis. *J Neurol*. 2025;272(4):311.
- Baiardi S, Rossi M, Giannini G, et al. Head-to-head comparison of four cerebrospinal fluid and three plasma neurofilament light chain assays in Parkinsonism. *NPJ Parkinsons Dis*. 2025;11(1):98.
- Guo Y, Shen XN, Huang SY, et al. Head-to-head comparison of 6 plasma biomarkers in early multiple system atrophy. *NPJ Parkinsons Dis*. 2023;9(1):40.
- Katzdöbler S, Nübling G, Kliez M, et al. GFAP and NfL as fluid biomarkers for clinical disease severity and disease progression in multiple system atrophy (MSA). *J Neurol*. 2024;271(10):6991–6999.
- Mattsson N, Cullen NC, Andreasson U, Zetterberg H, Blennow K. Association between longitudinal plasma neurofilament light and neurodegeneration in patients with alzheimer disease. *JAMA Neurol*. 2019;76(7):791–799.
- Musso G, Fiorenzato E, Misenti V, et al. Detecting amyloid and tau pathology in Parkinson's disease, 4R-tauopathies and control subjects with plasma pTau217. *Front Neurol*. 2025;16:1638852.
- Salvado G, Janelidze S, Bali D, et al. Plasma phosphorylated tau 217 to identify preclinical alzheimer disease. *JAMA Neurol*. 2025;82:1122–1134.
- 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):870–881.
- Wang Y, Cui L, Pan FF, et al. Validation of the robustness of blood-based biomarkers for predicting amyloid-beta positivity in Chinese populations. *Front Aging Neurosci*. 2025;17:1660755.
- Akamine S, Marutani N, Kanayama D, et al. Renal function is associated with blood neurofilament light chain level in older adults. *Sci Rep*. 2020;10(1):20350.
- Uyar M, Lezius S, Buhmann C, et al. Diabetes, glycated hemoglobin (HbA1c), and neuroaxonal damage in Parkinson's disease (MARK-PD study). *Mov Disord*. 2022;37(6):1299–1304.
- Baskaran G, Krisai P, Kuhne M, et al. Serum neurofilament light chain and cardiovascular outcomes in patients with atrial fibrillation. *JAMA Cardiol*. 2026;11:400–407.
- Abdelhak A, Petermeier F, Benkert P, et al. Serum neurofilament light chain reference database for individual application in paediatric care: a retrospective modelling and validation study. *Lancet Neurol*. 2023;22(9):826–833.