

RESEARCH

Open Access



Targeted CSF metabolomics and conformal prediction improve diagnostic accuracy of normal pressure hydrocephalus

Ulrika Hofling¹, Jenny Jakobsson^{2,3}, Ida Erngren², Oskar Ekman¹, Eva Freyhult⁴, Akshai Parakkal Sreenivasan², Jakob Siljebo², Sylwia Libard^{5,6}, Lena Kilander⁷, Malin Löwenmark⁷, Martin Ingelsson^{8,9,10}, Kim Kultima^{2†} and Johan Virhammar^{1*†}

Abstract

Background and objectives Idiopathic normal pressure hydrocephalus (iNPH) is a progressive but treatable neurological disorder. Yet, diagnosis is often confounded by overlapping symptoms and biomarker profiles with Alzheimer's disease (AD), mild cognitive impairment (MCI), and frontotemporal dementia (FTD). We aimed to determine whether cerebrospinal fluid (CSF) metabolomic profiling, combined with uncertainty-aware machine learning using conformal prediction (CP), could improve diagnostic differentiation of iNPH.

Methods CSF samples were collected from 120 patients with iNPH, 44 healthy controls, and 152 individuals with AD, MCI, or FTD. Targeted metabolomics of 59 metabolites was performed using liquid chromatography–high-resolution mass spectrometry. Group differences were assessed using age- and sex-adjusted regression models. Multivariate classification with partial least squares discriminant analysis (PLS-DA) incorporated metabolites, demographics, and conventional biomarkers (amyloid- β 42, tau, phosphorylated tau). CP was applied to address individual-level diagnostic uncertainty.

Results Eight metabolites (proline, threonine, histidine, tyrosine, tryptophan, isobutyrylcarnitine, citric acid, and dehydroascorbic acid) were consistently reduced in iNPH ($q < 0.05$), independent of ventricular volume and cortical tau or amyloid- β pathology. An integrated PLS-DA model combining metabolomic, demographic, and AD-biomarker data achieved excellent discrimination (AUC = 0.97). CP provided calibrated case-level confidence, identifying clear-cut and uncertain cases while maintaining high accuracy (94% for iNPH, 97% for not-iNPH).

Discussion iNPH exhibits a distinct CSF metabolomic signature reflecting altered amino acid metabolism, mitochondrial function, and oxidative stress. Integrating metabolomic data with established biomarkers enhances diagnostic accuracy, while CP adds individualized uncertainty estimates to improve diagnostic confidence and guide treatment decisions.

Keywords iNPH, CSF, Metabolomics, Biomarkers, LC-MS, Neurodegeneration, Oxidative stress, Glymphatic system

[†]Kim Kultima and Johan Virhammar contributed equally.

*Correspondence:
Johan Virhammar
johan.virhammar@uu.se

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



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Introduction

Idiopathic normal pressure hydrocephalus (iNPH) is a progressive neurological disorder affecting older adults, clinically defined by the classic triad of gait disturbance, cognitive impairment, and urinary incontinence [1, 2]. Among these, gait abnormalities are typically the earliest and most prominent symptom [3]. Disruption of cerebrospinal fluid (CSF) dynamics is a central feature of the condition, and dysfunction of the glymphatic system is discussed as a contributing factor [2, 4].

Despite its distinct clinical presentation, the diagnosis and management of iNPH remains challenging, and often includes different invasive methods to manipulate or measure the dynamics of the CSF-system. Diagnosing iNPH can also be a clinical challenge due to symptom overlap with neurodegenerative conditions such as Alzheimer's disease (AD), and the frequent coexistence of these pathologies in aging populations [4]. Some centers use CSF biomarkers, such as amyloid beta (amyloid- β 42) and tau proteins, to improve diagnostic accuracy in the diagnostic work-up [5–7]. While these biomarkers help to identify AD pathology, they lack specificity for iNPH and are insufficient to differentiate between conditions.

Recent advances in CSF proteomic and metabolomic profiling have opened new avenues for exploring disease mechanisms and enhancing diagnostic precision. Preliminary studies comparing CSF metabolite signatures in iNPH and AD suggest metabolomics may offer valuable insights into differential diagnosis [8–11]. However, despite advances in biomarker discovery, significant limitations remain in clinical applications, including the inability to communicate uncertainty in individual predictions, unlike the aggregate measures of model performance typically reported (e.g., area under the curve, AUC), and the lack of explainability or interpretability at the level of a specific prediction [12].

To capture individual uncertainty, advanced computational approaches such as predictive models augmented with conformal prediction (CP) offer a robust framework for clinical decision-making [13–15]. CP enhances machine learning classifiers by assigning valid confidence measures to each prediction, thereby explicitly quantifying diagnostic uncertainty. In idiopathic normal pressure hydrocephalus (iNPH), where symptom overlap with AD and other dementias complicates diagnosis, CP can mitigate misclassification risk by flagging ambiguous cases and enabling more reliable differentiation. This uncertainty-aware approach is particularly valuable for identifying patients who may benefit from additional confirmatory testing.

The aim of this study was to identify novel biomarker patterns and metabolic signatures that extend beyond previously examined non-specific CSF markers, such as amyloid- β 42 and tau, in order to distinguish iNPH from

its major diagnostic mimics AD, AD-related mild cognitive impairment (ADMCI), stable MCI, and frontotemporal dementia (FTD). Furthermore, by incorporating CP into the analysis, we sought to provide individualized measures of diagnostic uncertainty, thereby enhancing the clinical interpretability of the results.

Methods

Participants

A total of 316 individuals were included in the study. Patients diagnosed with iNPH were recruited prospectively at Uppsala University Hospital between 2014 and 2020. All eligible participants with available CSF samples during the study period were included, no formal sample size calculation was performed. All participants underwent a comprehensive evaluation by a specialized NPH multidisciplinary team, with diagnosis determined according to international guidelines for iNPH [16]. Radiological data confirming normal ventricular size were not available for controls or other diagnostic groups.

Patients with AD, ADMCI, and MCI were diagnosed according to the National Institute on Aging and Alzheimer's Association (NIA-AA) criteria. FTD diagnoses were determined using established clinical criteria combined with neuroimaging findings [17–19]. Diagnoses and lumbar punctures for participants in the neurodegenerative disease groups were performed between 2005 and 2018. Patients classified in the MCI group were longitudinally followed for 4–9 years following CSF sampling to ensure diagnostic stability. Importantly, none of these participants converted to AD dementia during the follow-up period, confirming their stable MCI classification.

The control group (C) consisted of neurologically healthy individuals without neurocognitive disorders, recruited at Uppsala University Hospital through local newspaper advertisements and from a cohort of older healthy individuals, explaining the higher median age. All analyses were adjusted for age to minimize bias. The study protocol received approval from the Swedish Ethical Review Authority. Written informed consent was obtained from all study participants, including patients and healthy controls, before enrollment and sample collection [20]. The sex and age distribution across all study groups is presented in Table 1.

Sample collection

CSF samples from all diagnostic groups and controls were obtained via lumbar puncture. All samples were stored in polypropylene tubes and frozen at -70°C . Before analysis, samples were thawed, aliquoted into microtubes, and refrozen at -70°C to maintain sample

Table 1 Demographic and biomarker characteristics across groups. Group differences were assessed using age- and sex-adjusted regression models. Values are presented as medians (Q1, Q3). In-group correlations between age and CSF biomarkers are provided in Supplementary table 1

Characteristics	iNPH (n = 120)	AD (n = 72)	ADMCI (n = 24)	FTD (n = 8)	MCI (n = 48)	Control (n = 44)
Age, median (Q1, Q3)	74 (70, 78)	72 (66, 76)	71 (67, 76)	71 (64, 76)	70 (61, 74)	88 (82, 88)
Female, n (%)	46 (38%)	43 (60%)	13 (54%)	3 (38%)	17 (35%)	10 (23%)
Male, n (%)	74 (62%)	29 (40%)	11 (46%)	5 (63%)	31 (65%)	34 (77%)
Amyloid- β 42, median (Q1, Q3), ng/L	586 (418, 680)	400 (340, 465)	387 (350, 483)	710 (595, 819)	754 (645, 995)	663 (494, 927)
Tau, median (Q1, Q3), ng/L	180 (139, 240)	680 (471, 920)	528 (445, 650)	265 (222, 460)	244 (186, 330)	425 (353, 681)
Ptau, median (Q1, Q3), ng/L	26 (21, 32)	84 (60, 115)	77 (55, 93)	46 (35, 55)	45 (35, 54)	64 (45, 78)

Abbreviations: AD: amyloid- β 42; Amyloid-beta 42, tau: total tau, ptau: phosphorylated tau

Normal reference ranges for biomarkers: amyloid- β 42: > 530 ng/L; tau: < 350 ng/L

integrity. No adverse events related to CSF sampling or cortical biopsies were reported.

Sample preparation

Thawed CSF samples were processed for targeted metabolomics using high-performance liquid chromatography coupled to high-resolution mass spectrometry (HPLC-HRMS), following a previously published protocol [21]. Analysts performing metabolomic profiling were blinded to clinical diagnoses and biopsy results. Briefly, 100 μ L samples underwent protein precipitation with 300 μ L cold methanol solution. Samples were vortexed for 15 seconds and incubated at -20°C for 1 hour. Following incubation, samples were centrifuged at 21,100 RCF at 4°C for 15 minutes. A 300 μ L aliquot of the supernatant was transferred to new tubes and concentrated by evaporation under a gentle nitrogen stream. The dried samples were reconstituted with 40 μ L of 5% methanol in 95% H_2O , transferred into HPLC vials, and stored at -80°C until analysis.

Mass spectrometry analysis

The HPLC-HRMS analysis was performed using an Ultimate 3000 HPLC system (Thermo Scientific, Waltham, MA, USA) coupled to a high-resolution hybrid quadrupole Q Exactive Orbitrap mass spectrometer (Thermo Scientific) operating in positive ionization mode with a target m/z range of 55–820. Quality control (QC) measures included pooled QC samples created by combining 10 μ L aliquots from prepared samples to generate a global QC sample, which was repeatedly injected alongside blank samples throughout the analytical sequence. Sample analysis involved injecting five μ L volumes using a method optimized for non-polar metabolites that employed reversed-phase liquid chromatography. Chromatographic separation was achieved using a C18 column (Accucore aQ RP C18 column, 100×2.1 mm, $2.6 \mu\text{m}$, Thermo Scientific). A 22-minute gradient elution was employed with mobile phase A consisting of H_2O with 0.1% formic acid and mobile phase B comprising 90% acetonitrile and 10% 2-propanol with 0.1% formic

acid. The mass spectrometer settings utilized have been previously described in detail [21].

Ventricular volume

Ventricular volume was measured using two methods. Evans' index was manually determined from axial sections of preoperative computed tomography scans [22]. In addition, in cases with an available preoperative 3DT1 MRI scan, ventricular volume was assessed with *Smart-Print*, a freely available, semi-automatic volumetric segmentation software [23].

Brain biopsy procedures

Patients with iNPH were treated with insertion of a shunt system with the burr hole placed over the frontal lobe (most often the right side). When no surgical contraindications were present, one needle biopsy from the cortex measuring approximately 2×3 mm was obtained immediately before shunt insertion, as previously described [24]. The biopsy specimens underwent standard histopathological processing, including fixation in formalin solution and paraffin embedding, followed by hematoxylin-eosin staining and additional immunohistochemical methods [24]. Qualified neuropathologists performed diagnostic evaluation of the samples and assessed the presence or absence of hyperphosphorylated tau (Thermo Fisher Scientific; Clone Tau8) and amyloid- β (Dako-Agilent; Clone 6F/3D) protein aggregations within the biopsy specimens. Here we define biopsy findings of both tau and amyloid- β 42 as AD+ and all other combinations as AD-.

Statistical analysis

Raw metabolomics data were processed in TraceFinder (version 4.1, Thermo Scientific). Relative metabolite concentrations were normalized for weekly signal drift using a locally estimated scatterplot smoothing (LOESS) model, and inter-week effects were corrected by dividing each metabolite intensity by the weekly median global QC intensity. In total, 219 metabolites spanning diverse classes and pathways were analyzed, of which 59 (26.9%)

met the quality control threshold ($\geq 90\%$ non-missing values) and were included in subsequent analyses.

Linear regression models were employed to find alterations in metabolite concentrations for the five patient groups and control subjects. All regression models were adjusted for age and sex to account for demographic differences between groups. In addition, a sensitivity analysis was performed using a subset of healthy controls closer in age to the iNPH group, which confirmed the main findings, although with slightly higher p-values due to smaller sample size. The linear regression models were evaluated using analysis of variance (ANOVA) F-tests, while pairwise group comparisons were conducted using post hoc t-statistics. The ANOVA p-values were adjusted for multiple testing using the Benjamini-Hochberg method for false discovery rate (FDR), with adjusted p-values (q values) below 5% considered statistically significant. Correlations between metabolite measurements were assessed using Pearson's correlation, and $p < 0.01$ was considered significant. We also computed within-group correlations between age and the clinically used CSF biomarkers amyloid β 42, tau, ptau for each diagnostic group (Supplementary Table S1).

A multivariate partial least-squares discriminant analysis (PLS-DA) model was constructed to distinguish iNPH from the other patient groups (AD, ADMCI, MCI, FTD) (controls excluded). Three different PLS-DA models were built. The first included age, sex, and the clinically used CSF biomarkers. The second included the 59 measured metabolites, and the third included age and sex, the established CSF biomarkers, and all metabolite data. Model performance was evaluated using ten five-fold cross-validations.

Conformal prediction

Conformal prediction (CP) is a model-agnostic framework that can be applied to any machine learning algorithm to quantify prediction uncertainty [13]. CP was implemented as a post-hoc calibration layer on top of a supervised PLS-DA model. As the calibration step relies on labeled data to compute nonconformity scores and guarantee coverage, the resulting conformal predictions are supervised. In classification problems, CP generates a prediction set that may include one, several, or no class labels. For each potential label, CP computes a p-value indicating how unusual a new data point appears compared with the calibration data when assigned that label. A low p-value suggests that the data point differs substantially from previously observed examples of that label. In the transductive CP (TCP) framework, each new data point is evaluated by temporarily adding it to the dataset under each possible class label. Conformity scores are then recomputed to assess how well the data point fits relative to the calibration data. This process is

repeated for every prediction, which makes TCP more computationally demanding but also particularly reliable in smaller datasets. For each label, a p-value is calculated as the proportion of calibration scores greater than or equal to the test data point score. Labels with p-values above the predefined significance level were retained in the final prediction set. Depending on the results, the prediction set could contain a single label (high confidence), multiple labels (ambiguous classification), or be empty (low confidence).

TCP was employed in this study due to its ability to maintain calibration of prediction regions under small-sample conditions. This calibration was assessed by comparing expected and observed error rates, with good calibration indicated by a diagonal relationship in the calibration plot.

Results

Study participants

The study cohort comprised participants with a median age of 74 years (44–89), with an even sex distribution of 42% women and 58% men. Established CSF biomarkers (amyloid β 42, tau, and ptau) analyzed as part of routine clinical work-up are presented in Table 1. Within-group correlations between age and biomarkers were weak or nonsignificant in most groups (Supplementary Table S1), indicating that between-group biomarker differences are unlikely to be explained by age.

The iNPH group demonstrated significantly lower levels of tau and ptau compared to the other diagnostic groups and reduced levels of amyloid- β 42 compared to healthy controls (Supplementary Figure 1). Principal component analysis (PCA) of the metabolite profiles demonstrated partial separation between patient groups, as shown in Supplementary Figure 2.

Metabolomic differences between iNPH and neurodegenerative disorders

Fifty-nine metabolites were semi-quantified in the CSF of all the included groups (Supplementary Table 2). Of the 59 metabolites, eight demonstrated overall statistically significant differences between groups (ANOVA F-test with FDR correction at 5%) and showed consistent pairwise separation of iNPH from all other diagnostic groups and controls ($p < 0.05$; Table 2, Fig. 1). These discriminatory metabolites included five amino acids: proline, threonine, histidine, tyrosine, and tryptophan, and in addition, isobutyrylcarnitine, citric acid, and the vitamin derivative dehydroascorbic acid. Notably, all eight metabolites were detected at lower concentrations in the iNPH group than in healthy controls and the neurodegenerative disorder groups.

Table 2 Linear regression, adjusted for age and sex, was used to identify metabolites differing between NPH and other diagnostic groups. ANOVA F-tests with FDR correction determined overall significance, and pairwise comparisons are shown as log₂ Fold changes (iNPH vs other group) with confidence intervals and p-values

Metabolite	AD	ADMCI	MCI	FTD	C
Isobutyrylcarnitine	-0.48 [-0.69, -0.27], 1.2e-5	-0.55 [-0.86, -0.23], 6.8e-4	-0.37 [-0.62, -0.12], 0.003	-1.12 [-1.63, -0.61], 2.2e-5	-0.43 [-0.71, -0.15], 0.003
Proline	-0.40 [-0.55, -0.25], 2.9e-7	-0.29 [-0.51, -0.07], 0.010	-0.35 [-0.52, -0.17], 1.4e-4	-0.59 [-0.95, -0.23], 0.001	-0.28 [-0.48, -0.08], 0.007
Threonine	-0.28 [-0.38, -0.17], 3.8e-7	-0.32 [-0.48, -0.17], 5.4e-5	-0.17 [-0.29, -0.04], 0.009	-0.27 [-0.52, -0.01], 0.040	-0.21 [-0.35, -0.07], 0.003
Histidine	-0.24 [-0.36, -0.13], 5.3e-5	-0.30 [-0.47, -0.13], 5.6e-4	-0.34 [-0.48, -0.20], 1.5e-6	-0.28 [-0.56, -0.00], 0.050	-0.22 [-0.37, -0.06], 0.006
Tyrosine	-0.29 [-0.43, -0.15], 9.2e-5	-0.31 [-0.52, -0.10], 0.004	-0.42 [-0.59, -0.25], 1.6e-6	-0.39 [-0.73, -0.05], 0.026	-0.27 [-0.46, -0.08], 0.010
Tryptophan	-0.31 [-0.45, -0.16], 4.7e-5	-0.47 [-0.68, -0.25], 2.3e-4	-0.32 [-0.49, -0.15], 2.6e-4	-0.46 [-0.81, -0.11], 0.010	-0.23 [-0.42, -0.03], 0.022
Citric acid	-0.21 [-0.30, -0.12], 2.5e-6	-0.18 [-0.31, -0.06], 0.005	-0.16 [-0.26, -0.06], 0.002	-0.33 [-0.54, -0.13], 0.002	-0.16 [-0.27, -0.05], 0.006
Dehydroascorbic acid	-0.19 [-0.28, -0.10], 4.6e-5	-0.18 [-0.31, -0.04], 0.010	-0.15 [-0.26, -0.05], 0.005	-0.29 [-0.51, -0.07], 0.009	-0.16 [-0.28, -0.04], 0.008

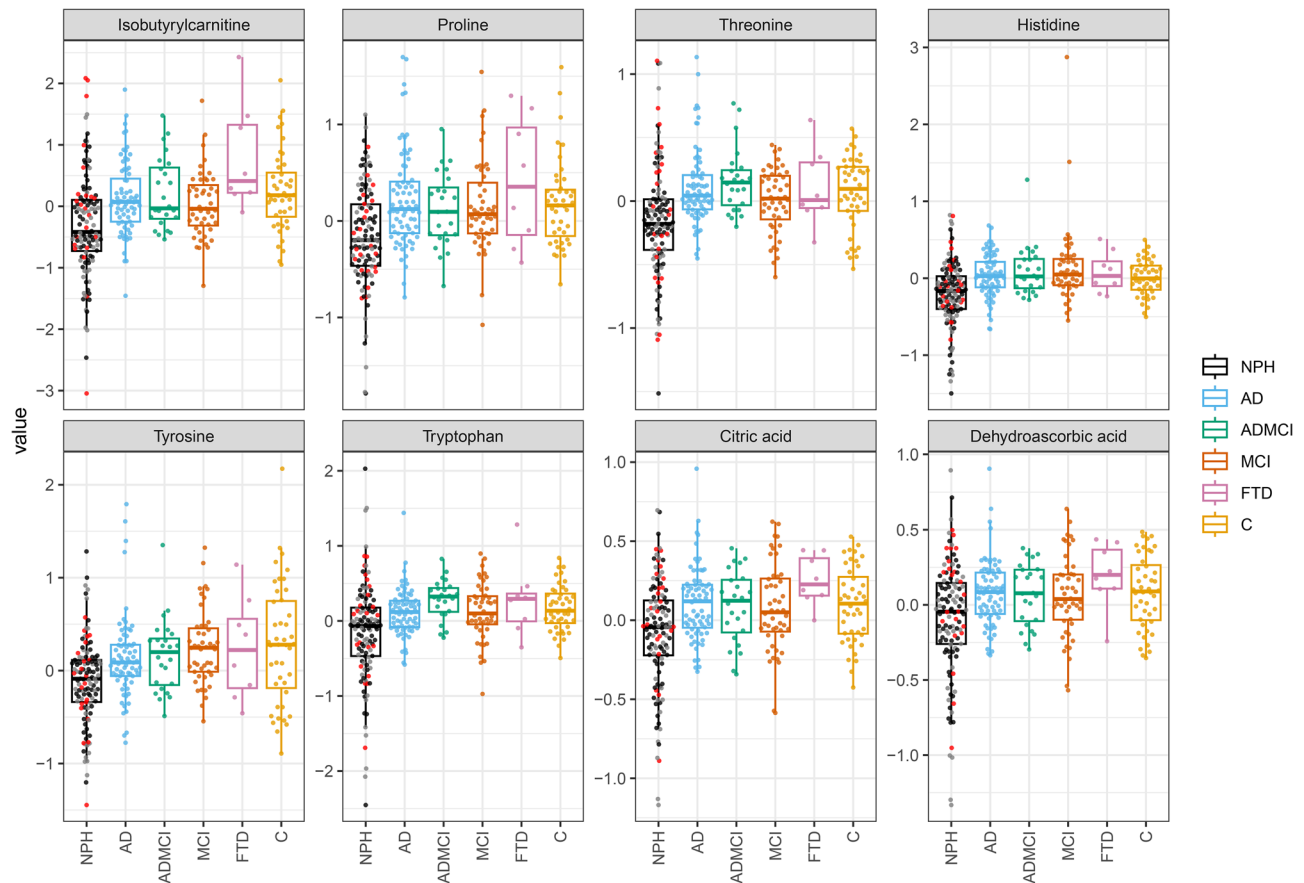


Fig. 1 Box-plots of the eight metabolites that were lower in iNPH compared with controls and all other groups (significant global p-value (ANOVA; FDR 5%) and also $p < 0.05$ in the pairwise iNPH comparisons). For iNPH, the individuals with AD+ biopsies are red, AD- are black, and patients with no biopsy are grey

Correlation with ventricular volume

No significant correlation was observed between ventricular volume and the eight metabolites found at lower concentrations in iNPH, whether ventricular volume was measured using the Evans index or MR volumetry (Supplementary Figure 3).

Association with Alzheimer's disease pathology in brain biopsies

Cortical brain biopsies were available in 80/120 (66,7%) of patients with iNPH. Among these, 58/80 (73%) were classified as AD+ (presence of both tau and amyloid- β 42). No significant differences in the levels of the eight discriminatory metabolites were observed between patients with brain biopsy findings consistent with AD pathology (AD+) and those without (AD-) (Fig. 2A). In contrast, amyloid- β 42 (A β 42) concentrations were significantly lower ($p=0.0098$) and total tau levels significantly higher ($p=0.039$) in the AD+ group (Fig. 2B).

Distinguishing iNPH using PLS-DA

Three partial least squares discriminant analysis (PLS-DA) models were constructed to differentiate iNPH from all other groups. The first included demographic variables (age and sex) together with AD CSF biomarkers (A β 42, total tau, and phosphorylated tau); the second was based solely on the panel of 59 CSF metabolites; and the third integrated all variables. Among these, the integrated model demonstrated the highest discriminative performance, achieving an AUC of 0.97 in distinguishing iNPH from the other diagnostic groups (Table 3, Fig. 3, Supplementary Figures 4–5).

Uncertainty-aware prediction for iNPH diagnosis

Although AUC-based evaluation reflects the overall discriminative ability of a model, it does not capture diagnostic uncertainty at the individual level. To address this limitation, we applied conformal prediction (CP), which provides case-level confidence estimates by quantifying uncertainty for each prediction.

The diagnostic performance was assessed by combining the PLS-DA analysis with transductive conformal prediction (TCP). The dataset comprised 118 iNPH cases (82 training, 36 test) and 149 non-iNPH cases (104 training, 45 test), using a 70/30 train–test split.

The calibration curve demonstrated that TCP yielded a well-calibrated model (Fig. 4a). At a significance level of 0.09 (corresponding to 91% confidence), the model achieved the highest proportion of single-label predictions (Fig. 4b). The corresponding confusion matrix (Fig. 4c) shows that the model confidently and correctly classified 27 of the 35 iNPH cases as single-label predictions. Among the six individuals predicted as having

concurrent iNPH, four underwent perioperative cortical biopsy; three were confirmed as AD-positive, supporting the presence of coexisting iNPH and AD spectrum pathology. Figure 4d illustrates the predicted p-values for iNPH (y-axis) and non-iNPH (x-axis), where higher values reflect greater confidence in the corresponding classification, and predictions closer to the origin indicate lower-confidence assignments.

Discussion

In this study, we applied targeted metabolomic profiling to CSF samples from patients with iNPH, neurodegenerative disorders, and healthy controls. Eight metabolites, five amino acids, one carnitine, one citric acid cycle intermediate, and one vitamin-related compound, were consistently reduced in iNPH compared with all other groups. These alterations were independent of ventricular size and of AD pathology observed in cortical biopsies. A multivariate model integrating metabolite data, demographic variables, and established CSF biomarkers demonstrated excellent diagnostic accuracy in distinguishing iNPH from other conditions. To further enhance clinical applicability, we incorporated transductive conformal prediction, which provided individualized measures of diagnostic uncertainty and improved interpretability at the patient level.

Previous metabolomics studies in iNPH are limited, but differences in the metabolic profiles of AD and iNPH have been reported. These findings particularly included metabolites related to glycerolipid and amino acid metabolism [8]. In the present study, eight metabolites were significantly reduced in the CSF of iNPH patients compared to controls and individuals with MCI, AD, or FTD. One possible interpretation is that these changes reflect a general disturbance in CNS metabolism in iNPH. All analyses were adjusted for age and sex, and a sensitivity analysis using a subset of controls closer in age to the iNPH group confirmed the main findings, although p values were somewhat higher due to the smaller sample size. While this suggests that the differences are not solely explained by age, residual confounding cannot be excluded, and we therefore limit mechanistic interpretation to what is directly supported by the data.

Several altered metabolites, including amino acids and intermediates related to energy metabolism and oxidative stress (e.g., citric acid and dehydroascorbic acid), may indicate impaired cellular metabolic activity in the brain parenchyma [25]. However, it remains unclear whether the lower CSF levels are due to reduced production of these metabolites within the brain, possibly due to diminished neuronal or glial function, or altered transport, turnover, or clearance mechanisms, i.e., dysfunction of the glymphatic system. The glymphatic pathway facilitates exchange between interstitial fluid and CSF

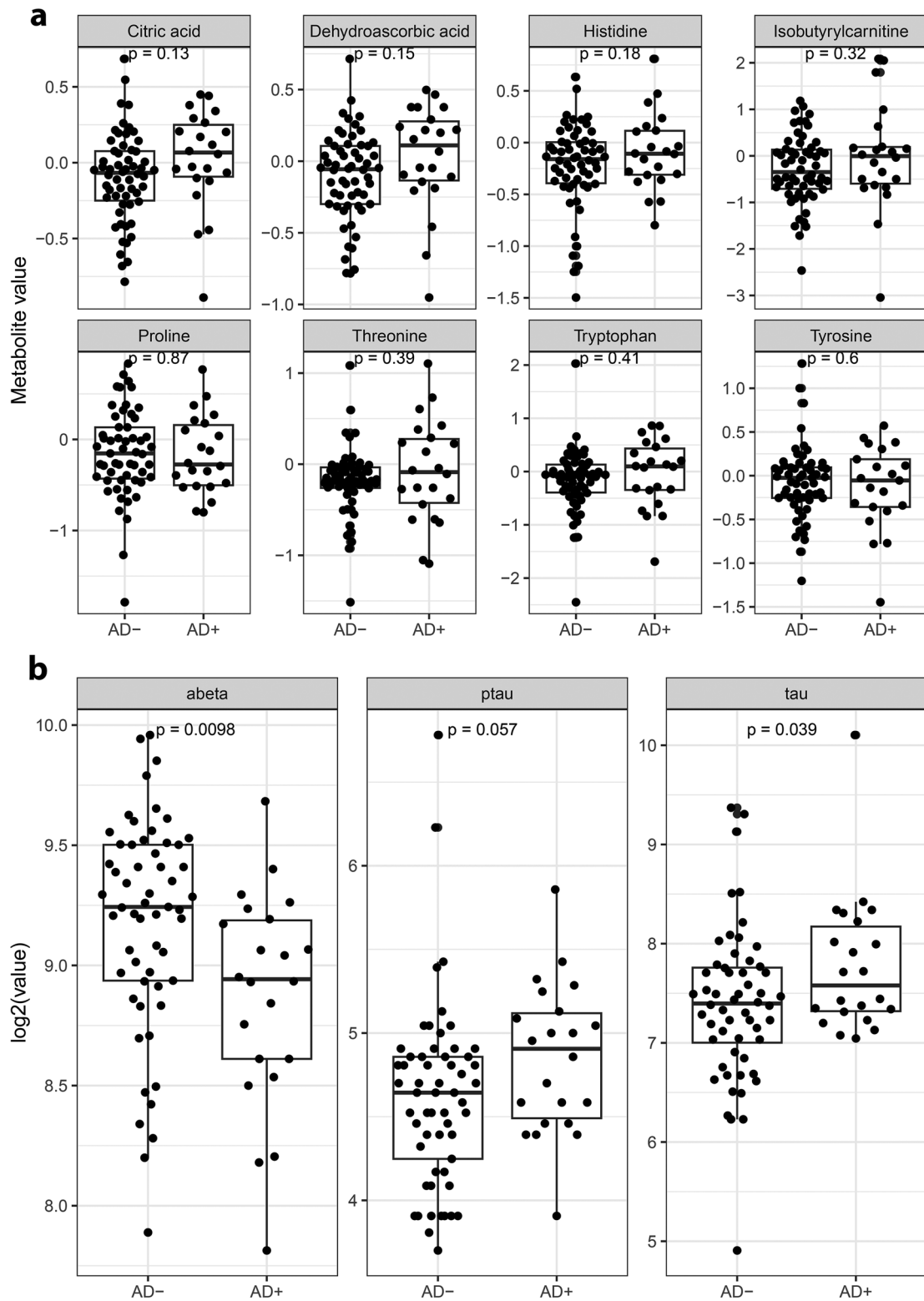


Fig. 2 AD biopsies with manually evaluated presence or absence of hyperphosphorylated tau and total amyloid- β aggregations within the biopsies. The association between measured values and AD+/AD- was analyzed with a t-test

Table 3 Performance of PLS-DA models. Mean (standard deviation) of AUC (area under the ROC curve) and error rate (ER) are computed over the 50 test sets. mean (SD)

Variables	AUC	ER
age, sex, amyloid-β42, tau, ptau	0.89 (0.04)	0.19 (0.04)
All 59 metabolites	0.91 (0.04)	0.13 (0.04)
age, sex, amyloid-β42, tau, ptau, all 59 metabolites	0.97 (0.03)	0.08 (0.04)

and is believed to be involved in the clearance of metabolic waste products from the brain [2]. Impaired glymphatic flow could theoretically lead to the accumulation of certain substances in brain tissue or to the reduced appearance of these substances in CSF, depending on the dynamics of transport and clearance. Although glymphatic function is known to decline with age [26–28], the observed metabolite differences are unlikely to be explained solely by aging, as our comparisons included

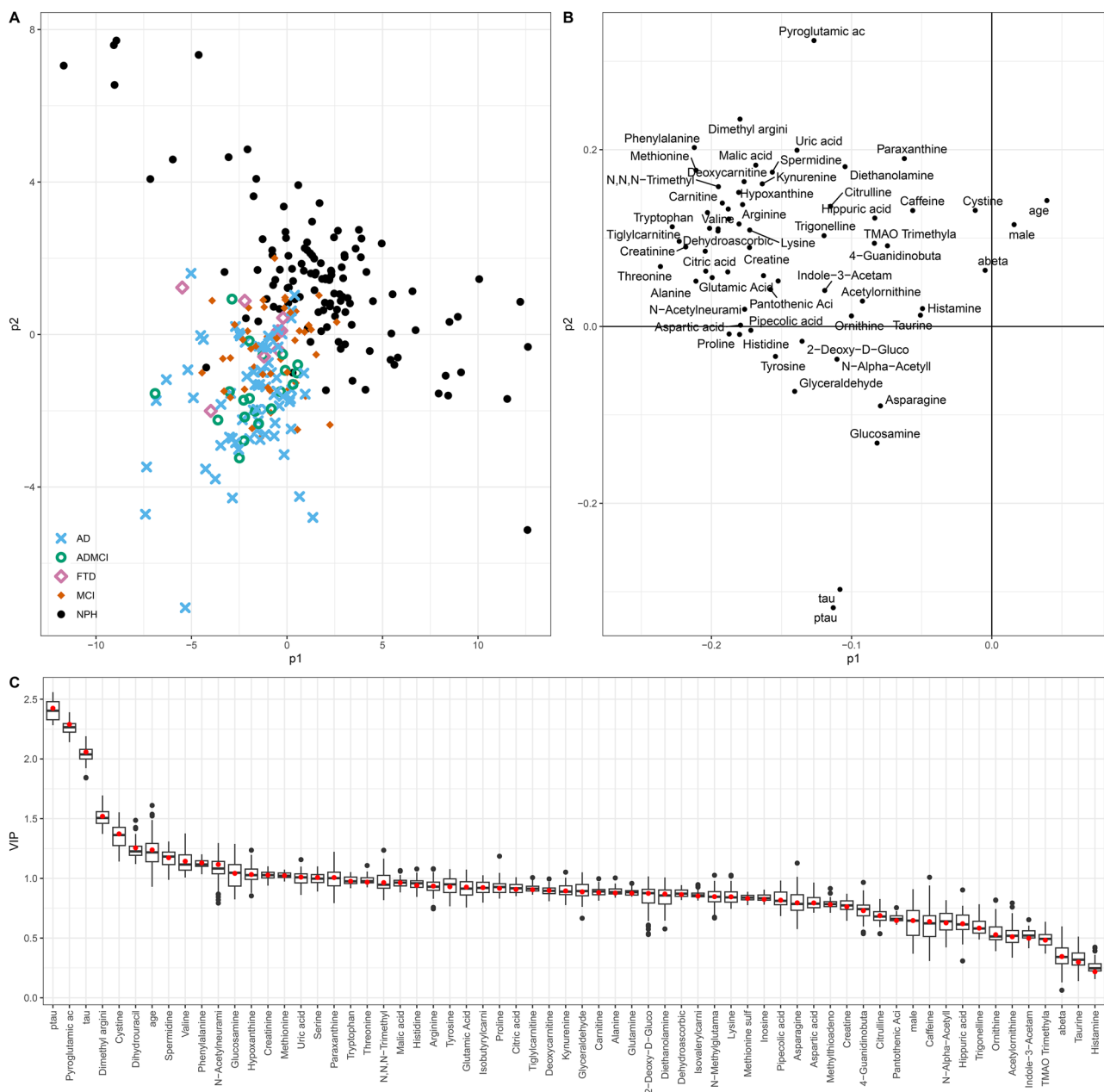


Fig. 3 PLS-DA scores and loadings for the model based on age, sex, amyloid-β42, ptau, tau and all metabolites comparing NPH to AD, AD/MCI, MCI, and FTD

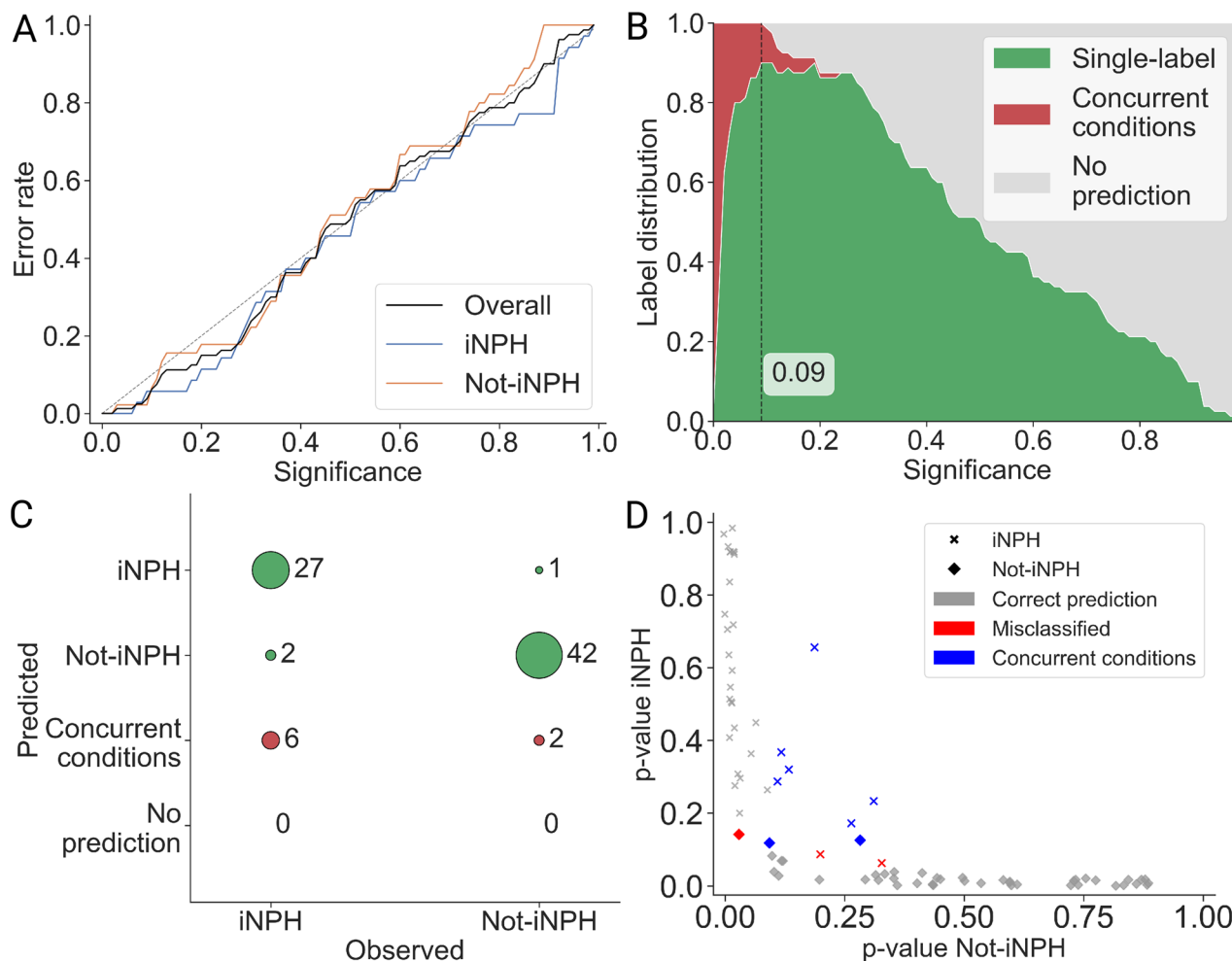


Fig. 4 Evaluation of conformal prediction. **A)** calibration curve showing significance levels on the x-axis and observed error rates on the y-axis. The diagonal line represents perfect calibration. **B)** label distribution across significance levels, illustrating the proportion of single-label, multiple-label, and empty-label predictions. The significance level yielding the highest number of single-label predictions was selected for model evaluation (significance level of 0.09). **C)** confusion matrix summarizing predictive performance at the chosen significance level (0.09). The model achieved an accuracy of 94% for iNPH patients and 97% for the not-iNPH group. **D)** predicted p-values for each test individual, with p-values for iNPH on the y-axis and p-values for not-iNPH on the x-axis

multiple neurodegenerative groups of similar age to the iNPH cohort.

Decreased CSF levels of dehydroascorbic acid, an antioxidant and neurotransmitter precursor, may contribute to oxidative stress and neuroinflammation in iNPH [29]. Citric acid dysregulation could impair neuronal energy metabolism and glymphatic clearance, while altered isobutyrylcarnitine points to mitochondrial β -oxidation deficits [30, 31]. The affected amino acids reflect key neurophysiological roles: proline supports blood–brain barrier integrity and redox balance. Through the kynurenine pathway, tryptophan metabolism may link to inflammation and neurochemical shifts [32, 33]. Tyrosine, a catecholamine precursor, has been associated with ventricular changes, and reduced histidine and threonine may affect neurotransmission and mTOR signaling,

potentially influencing cognitive and motor symptoms in iNPH [25, 34, 35].

Previous studies have shown that specific CSF biomarkers in iNPH, such as amyloid precursors, may be lower than in controls and other neurodegenerative disorders [5–7]. It has been suggested that this reduction could be due to a dilution effect related to the enlarged ventricles commonly seen in iNPH [36–38]. Because the eight metabolites distinguishing iNPH patients were consistently lower than in all other groups, we examined whether their concentrations correlated with ventricular volume. No such associations were observed using either the Evans index or MRI-based volumetric measurements. This suggests that the metabolite alterations are not simply a result of ventricular enlargement or CSF dilution,

but may instead reflect underlying metabolic changes in the CNS.

Comorbidity with AD is common in iNPH, and previous studies have demonstrated a high prevalence of histopathological AD findings in cortical brain biopsies from iNPH patients [39]. Given that our comparison groups included cohorts representing different stages of AD, we also examined whether AD-related histological findings influenced metabolite levels within the iNPH group. However, no significant differences were observed in any of the eight metabolites between iNPH patients with and without histopathological evidence of AD. Our findings for amyloid β 42 are consistent with the established pattern of lower levels in AD than in other conditions and somewhat lower levels in iNPH than in healthy controls. Within-group correlations between amyloid β 42, tau, and ptau and age were weak or nonsignificant in most groups (Supplementary Table 1), indicating that age does not account for the observed between-group differences. In addition, lower concentrations of tau, phosphorylated tau, and amyloid β 42 in iNPH compared to healthy controls have been consistently reported in previous studies [6, 7, 40], and our findings are in line with this established pattern.

Glymphatic impairment has also been reported in Alzheimer's disease and other neurodegenerative conditions [41–43], which may contribute to overlapping biomarker profiles and underscores the value of multimodal assessment.

To evaluate the diagnostic utility of the metabolomic data, we also developed PLS-DA models incorporating various combinations of clinical and biochemical variables. A model that integrated all available data, demographics, clinically established CSF biomarkers (abeta, tau, and p-tau), and all 59 identified metabolites showed the highest diagnostic performance. This model achieved substantially better discrimination of iNPH cases than models using only traditional CSF biomarkers or metabolite data alone. These findings highlight the added value of CSF metabolomics in improving diagnostic accuracy and suggest that incorporating metabolic profiling could have a potential role in future diagnostic workflows of iNPH.

In addition, integrating TCP with PLS-DA enhanced the diagnostic utility of our approach by introducing an uncertainty-aware framework. At a prediction confidence of 91%, the model confidently identified 27 of the 35 iNPH patients as single-label predictions, underscoring its reliability in classifying the majority of true iNPH cases. Six iNPH patients in the test set were flagged as having concurrent conditions, reflecting overlap with the not-iNPH group and suggesting the presence of AD spectrum pathology. The model misclassified two iNPH and one not-iNPH patient; however, increasing the

confidence threshold highlighted these cases as potential concurrent conditions rather than simple misclassifications. This underscores the value of incorporating uncertainty, as higher confidence thresholds identify additional patients who may benefit from closer monitoring [44]. Such patients should be considered for further diagnostic evaluation before treatment decisions are made.

Strengths

This study benefits from using a well-characterized, single-center cohort, ensuring clinical assessment and sample handling consistency. The inclusion of multiple diagnostic groups allows for relevant differential comparisons and increases the clinical relevance of the findings. A notable strength is the use of HPLC-HRMS, which enables broad, sensitive, and reproducible detection of metabolites across a wide dynamic range. Targeted assays, pooled quality control samples, and statistical adjustments for potential confounders enhance the reliability and comparability of the metabolomic data.

Limitations

One limitation of this study is the difference in age distribution between the control group and other diagnostic groups. Although all analyses were adjusted for age and sex, and a sensitivity analysis using age-matched controls confirmed the main findings, residual confounding cannot be excluded. This should be considered when interpreting the results. Also, the sample size in some diagnostic subgroups, such as FTD, was small.

The cross-sectional design limits conclusions about longitudinal changes or treatment effects. Furthermore, the chromatographic method used was optimized for semi-polar to non-polar metabolites. Analysis of highly polar metabolites would require a different LC-HRMS approach, such as HILIC, which we plan for future studies. Confirmational studies on separate cohorts are also needed to be more certain about how generalizable the results are.

Finally, we did not assess potential differences between ventricular and lumbar CSF metabolomic profiles. To our knowledge, no studies have directly compared these compartments, and future research should investigate whether such differences could influence biomarker interpretation.

Conclusion

This study identifies a distinct CSF metabolite profile that differentiates iNPH from other neurodegenerative disorders, including AD, MCI, and FTD. Eight metabolites related to neurotransmission, oxidative stress, and energy metabolism were consistently reduced in iNPH, independent of ventricular size or AD pathology. The strong performance of a multivariate model

combining metabolomic, demographic, and established biomarker data underscores the diagnostic potential of this approach, and CP extends this by providing patient-specific confidence estimates. These findings support the integration of CSF metabolomics into diagnostic workflows to improve accuracy and guide clinical decision-making in iNPH.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12987-026-00771-z>.

Supplementary Material 1

Acknowledgements

Region Uppsala (ALF-grant and R&D funds), Swedish Research Council (2024-03161).

Author contributions

JV and KK conceived and designed the study. JV, OE, SL, LK, ML, and MI collected the data. JJ, IE, EF, and KK performed the statistical analyses and interpretations. UH drafted the manuscript, and all authors reviewed, revised, and approved the final version of the manuscript.

Funding

Open access funding provided by Uppsala University. Johan Virhammar is funded by the Swedish Society for Medical Research (SG-22-0192-H-01).

Data availability

De-identified data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

The study was conducted in accordance with the principles of the Declaration of Helsinki and approved by the Swedish Ethical Review Authority (approval number: 2018/168 and 2021-05439-02). Written informed consent was obtained from all participants, including patients and healthy controls, before enrollment and sample collection.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Medical Sciences, Neurology, Uppsala University, Akademiska Sjukhuset, Ing 85, Uppsala 751 85, Sweden

²Department of Medical Sciences, Clinical Chemistry, Uppsala University, Uppsala, Sweden

³Department of Surgical Sciences, Uppsala University, Uppsala, Sweden

⁴Department of Cell and Molecular Biology, National Bioinformatics Infrastructure Sweden, Science for Life Laboratory, Uppsala University, Uppsala, Sweden

⁵Department of Pathology, Uppsala University Hospital, Uppsala, Sweden

⁶Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden

⁷Department of Public Health and Caring Sciences, Clinical Geriatrics, Rudbeck Laboratory, Uppsala University, Uppsala, Sweden

⁸Krembil Brain Institute, University Health Network, Toronto, Ontario, Canada

⁹Tanz Centre for Research in Neurodegenerative Diseases, Departments of Medicine and Laboratory Medicine & Pathobiology, University of Toronto, Toronto, Ontario, Canada

¹⁰Department of Public Health and Caring Sciences, Molecular Geriatrics, Rudbeck Laboratory, Uppsala University, Uppsala, Sweden

Published online: 07 February 2026

References

1. Adams RD, Fisher CM, Hakim S, Ojemann RG, Sweet WH. Symptomatic occult hydrocephalus with normal cerebrospinal-fluid pressure. *N Engl J Med*. 1965 July 15;273(3):117–26.
2. Wang Z, Zhang Y, Hu F, Ding J, Wang X. Pathogenesis and pathophysiology of idiopathic normal pressure hydrocephalus. *CNS Neurosci Ther*. 2020 Dec;26(12):1230–40.
3. Panciani PP, Migliorati K, Muratori A, Gelmini M, Padovani A, Fontanella M. Computerized gait analysis with inertial sensor in the management of idiopathic normal pressure hydrocephalus. *Eur J Phys Rehabil Med*. 2018 Oct;54(5):724–29.
4. Reeves BC, Karimy JK, Kundishora AJ, Mestre H, Cerci HM, Matouk C, et al. Glymphatic system impairment in Alzheimer's disease and idiopathic normal pressure hydrocephalus. *Trends Mol Med*. 2020 Mar;26(3):285–95.
5. Miyajima M, Nakajima M, Ogino I, Miyata H, Motoi Y, Arai H. Soluble amyloid precursor protein α in the cerebrospinal fluid as a diagnostic and prognostic biomarker for idiopathic normal pressure hydrocephalus. *Eur J Neurol*. 2013 Feb;20(2):236–42.
6. Craven CL, Baudracco I, Zetterberg H, Lunn MPT, Chapman MD, Lakdawala N, et al. The predictive value of T-tau and AB1-42 levels in idiopathic normal pressure hydrocephalus. *Acta Neurochir (Wien)*. 2017 Dec;159(12):2293–300.
7. Agren-Wilsson A, Lekman A, Sjöberg W, Rosengren L, Blennow K, Bergenheim AT, et al. CSF biomarkers in the evaluation of idiopathic normal pressure hydrocephalus. *Acta Neurol Scand*. 2007 Nov;116(5):333–39.
8. Nagata Y, Hirayama A, Ikeda S, Shirahata A, Shoji F, Maruyama M, et al. Comparative analysis of cerebrospinal fluid metabolites in Alzheimer's disease and idiopathic normal pressure hydrocephalus in a Japanese cohort. *Biomark Res*. 2018 Jan 22;6(1):5.
9. Torretta E, Arosio B, Barbacini P, Capitanio D, Rossi PD, Moriggi M, et al. Novel insight in idiopathic normal pressure hydrocephalus (iNPH) biomarker discovery in CSF. *Int J Mol Sci*. 2021 July 27;22(15):8034.
10. Braun M, Boström G, Ingelsson M, Kilander L, Löwenmark M, Nyholm D, et al. Levels of inflammatory cytokines MCP-1, CCL4, and PD-L1 in CSF differentiate idiopathic normal pressure hydrocephalus from neurodegenerative diseases. *Fluids Barriers CNS*. 2023 Oct 13;20(1):72.
11. Kamalian A, Shirzadeh Barough S, Ho SG, Albert M, Luciano MG, Yasar S, et al. Molecular signatures of normal pressure hydrocephalus: a large-scale proteomic analysis of cerebrospinal fluid. *Fluids Barriers CNS*. 2024 Aug 8;21(1):64.
12. Pierce RL, Van Biesen W, Van Cauwenberge D, Decruyenaere J, Sterckx S. Explainability in medicine in an era of AI-based clinical decision support systems. *Front Genet*. 2022 Sept 19;13:903600.
13. Vovk V, Gammerman A, Shafer G. Algorithmic learning in a random world [Internet]. Cham: Springer International Publishing; 2022 [cited 2025 Sept 10]. Available from: <https://link.springer.com/10.1007/978-3-031-06649-8>.
14. Sreenivasan AP, Vaivade A, Noui Y, Khoonsari PE, Burman J, Spjuth O, et al. Conformal prediction enables disease course prediction and allows individualized diagnostic uncertainty in multiple sclerosis. *Npj Digit Med*. 2025 Apr 24;8(1):224.
15. Olsson H, Kartasalo K, Mulliqi N, Capuccini M, Ruusuvoori P, Samartunga H, et al. Estimating diagnostic uncertainty in artificial intelligence assisted pathology using conformal prediction. *Nat Commun*. 2022 Dec 15;13(1):7761.
16. Relkin N, Marmarou A, Klinge P, Bergsneider M, Black PM. Diagnosing idiopathic normal-pressure hydrocephalus. *Neurosurgery*. 2005 Sept;57(3 Suppl):S4–16; discussion ii-v.
17. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement J Alzheimers Assoc*. 2011 May;7(3):263–69.
18. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement J Alzheimers Assoc*. 2011 May;7(3):270–79.

19. Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain J Neurol*. 2011 Sept;134(Pt 9):2456–77.
20. Boström G, Freyhult E, Virhammar J, Alcolea D, Tuman H, Otto M, et al. Different inflammatory signatures in Alzheimer's disease and frontotemporal dementia cerebrospinal fluid. *J Alzheimers Dis JAD*. 2021;81(2):629–40.
21. Herman S, Khoonsari PE, Tolf A, Steinmetz J, Zetterberg H, Åkerfeldt T, et al. Integration of magnetic resonance imaging and protein and metabolite CSF measurements to enable early diagnosis of secondary progressive multiple sclerosis. *Theranostics*. 2018 Aug 7;8(16):4477–90.
22. Evans WA Jr. An encephalographic ratio for estimating ventricular enlargement and cerebral atrophy. *Arch Neurol Psychiatry*. 1942 June 1;47(6):931–37.
23. Snöbom C, Malmberg F, Freyhult E, Kultima K, Fällmar D, Virhammar J. White matter changes should not exclude patients with idiopathic normal pressure hydrocephalus from shunt surgery. *Fluids Barriers CNS*. 2022 May 23;19(1):35.
24. Libard S, Alafuzoff I. Alzheimer's disease neuropathological change and loss of matrix/neuropil in patients with idiopathic normal pressure hydrocephalus, a model of Alzheimer's disease. *Acta Neuropathol Commun*. 2019 May 29;7(1):98.
25. He W, Wu G. Metabolism of amino acids in the brain and their roles in regulating food intake. In: Wu G, editor. *Amino acids in nutrition and health: amino acids in systems function and health* [Internet]. Cham: Springer International Publishing; 2020 [cited 2025 July 21]. p. 167–85. Available from: https://doi.org/10.1007/978-3-030-45328-2_10.
26. Wang J, Zhou Y, Zhang K, Ran W, Zhu X, Zhong W, et al. Glymphatic function plays a protective role in ageing-related cognitive decline. *Age Ageing*. 2023 July 1;52(7):afad107.
27. Chong S, Wang S, Gao T, Yuan K, Han Y, Shi L, et al. Glymphatic function decline as a mediator of core memory-related brain structures atrophy in aging. *J Transl Intern Med*. 2025 Feb;13(1):65–77.
28. Han G, Zhou Y, Zhang K, Jiao B, Hu J, Zhang Y, et al. Age- and time-of-day dependence of glymphatic function in the human brain measured via two diffusion MRI methods. *Front Aging Neurosci* [Internet]. 2023 May, 22 [cited 2026 Jan 2];15. Available from: <https://www.frontiersin.org/journals/aging-neuroscience/articles/10.3389/fnagi.2023.1173221/full>.
29. García-Krausa A, Ferrada L, Astuya A, Salazar K, Cisternas P, Martínez F, et al. Dehydroascorbic acid promotes cell death in neurons under oxidative stress: a protective role for astrocytes. *Mol Neurobiol*. 2016 Nov 1;53(9):5847–63.
30. Wang R, Lou L. The central role of the citric acid cycle in energy metabolism: from metabolic intermediates to regulatory mechanisms. *Biol Evid* [Internet]. 2024 May 28 [cited 2025 July 21];14 Available from: <https://bioscipublisher.com/index.php/be/article/view/3895>.
31. Jensen O, Matthaai J, Klemp HG, Meyer MJ, Brockmöller J, Tzvetkov MV. Isobutyrylcarnitine as a biomarker of OCT1 activity and interspecies differences in its membrane transport. *Front Pharmacol* [Internet]. 2021 May 10 [cited 2025 July 21];12. Available from: <https://www.frontiersin.org/journals/pharmacology/articles/10.3389/fphar.2021.674559/full>.
32. Vettore LA, Westbrook RL, Tennant DA. Proline metabolism and redox; maintaining a balance in health and disease. *Amino Acids*. 2021 Dec 1;53(12):1779–88.
33. Sarawi WS. Neurochemical insights into the role of tryptophan metabolites and kynurenine pathway in insomnia and its psychological and neurological comorbidities. *Mol Neurobiol*. 2025 July 19.
34. Ichinose *, Aida R, Nago-Iwashita Y, Moriya Y, Ide S, Ikeda K. Modulation of the dopaminergic neurotransmission by controlling the activity of Tyrosine Hydroxylase. *Int J Neuropsychopharmacol*. 2025 Feb 1;28(Supplement_1):i9.
35. van Berlekom A B, Kübler R, Hoogeboom JW, Vonk D, Sluijs JA, Pasterkamp RJ, et al. Exposure to the amino acids histidine, lysine, and threonine reduces mTOR activity and affects neurodevelopment in a human cerebral organoid model. *Nutrients*. 2022 Jan;14(10):2175.
36. Jeppsson A, Wikkelsö C, Blennow K, Zetterberg H, Constantinescu R, Remes AM, et al. CSF biomarkers distinguish idiopathic normal pressure hydrocephalus from its mimics. *J Neurol Neurosurg Psychiatry*. 2019 Oct;90(10):1117–23.
37. Lidén S, Farahmand D, Laurell K. Ventricular volume in relation to lumbar CSF levels of amyloid- β 1–42, tau and phosphorylated tau in iNPH, is there a dilution effect? *Fluids Barriers CNS*. 2022 July 17;19(1):59.
38. Pyrgelis ES, Boufidou F, Constantinides VC, Papaioannou M, Papageorgiou SG, Stefanis L, et al. Cerebrospinal fluid biomarkers in iNPH: a narrative Review. *Diagnostics*. 2022 Dec;12(12):2976.
39. Libard S, Laurell K, Cesarini KG, Alafuzoff I. Progression of Alzheimer's disease-related pathology and cell counts in a patient with idiopathic normal pressure hydrocephalus. *J Alzheimers Dis*. 2018 Jan 23;61(4):1451–62.
40. Jeppsson A1, Wikkelsö C1, Blennow 2 K, Zetterberg H3, Constantinescu R4, Remes AM 5, et al. CSF biomarkers distinguish idiopathic normal pressure hydrocephalus from its mimics. 2019 Oct;1117–23.
41. Keil SA, Jansson D, Braun M, Iliff JJ. Glymphatic dysfunction in Alzheimer's disease: a critical appraisal. *Science*. 2025 July 10;389(6756):eadv8269.
42. Beschorner N, Nedergaard M. Glymphatic system dysfunction in neurodegenerative diseases. *Curr Opin Neurol*. 2024;37(2):182–88.
43. Jia L, Chen Y, Li H, Zhao K, Ge S, Wang C, et al. The glymphatic system in neurodegenerative diseases and brain tumors: mechanistic insights, biomarker advances, and therapeutic opportunities. *Acta Neuropathol Commun* [Internet]. 2025 Dec 13 [cited 2025 Dec 29]; Available from: <https://doi.org/10.1186/s40478-025-02203-9>.
44. Banerji CRS, Chakraborti T, Harbron C, MacArthur BD. Clinical AI tools must convey predictive uncertainty for each individual patient. *Nat Med*. 2023 Dec;29(12):2996–98.

Publisher's Note

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