medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Title

Blood biomarker profiles in young-onset neurocognitive disorders: a cohort study

Authors

Oneil G. Bhalala[1,2,*], Jessica Beamish[3], Dhamidhu Eratne[3,4], Patrick Summerell[3], Tenielle Porter[5], Simon M. Laws[5], Matthew JY Kang[3,4], Aamira J. Huq[1], Wei-Hsuan Chiu[4], Claire Cadwallader[6], Mark Walterfang[3,4,7,8], Sarah Farrand[3,4], Andrew H. Evans[1], Wendy Kelso[3], Leonid Churilov[1], Rosie Watson[1,2], Nawaf Yassi[1,2], Dennis Velakoulis[3,4], Samantha M. Loi[3,4,*]

[1]Department of Medicine, Royal Melbourne Hospital, University of Melbourne, Parkville, VIC, Australia
[2]Population Healthy and Immunity Division, Walter and Eliza Hall Institute of Medical Research, Parkville, VIC, Australia
[3]Neuropsychiatry Centre, Royal Melbourne Hospital, Parkville, VIC, Australia
[4]Department of Psychiatry, University of Melbourne, Parkville, VIC, Australia
[5]Centre for Precision Health, Edith Cowan University, Joondalup, Western Australia, Australia
[6] Memory and Aging Center, UCSF Weill Institute for Neurosciences, University of California, San Francisco, CA, USA
[7]Florey Institute of Neuroscience and Mental Health, Parkville, VIC, Australia
[8]School of Medical and Health Sciences, Edith Cowan University, Joondalup, WA, Australia

***Corresponding Authors**

Oneil G. Bhalala
Walter and Eliza Hall Institute
1G Royal Parade
Parkville, VIC 3052
bhalala.o@wehi.edu.au

Samantha M. Loi
The RMH – Elizabeth St
Level 7, 635 Elizabeth St
Melbourne, VIC 3000
samantha.loi@unimelb.edu.au

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Abstract

INTRODUCTION: This study investigates the blood biomarkers associated with young-onset neurocognitive disorders.

METHODS: Sixty-five participants less than 65 years old with neurocognitive symptoms (median age at assessment of 58 years, 42% female) were categorised as either early-onset Alzheimer's disease (EOAD, n=18), non-AD neurodegeneration (nAD-ND, n=23) or primary psychiatric disorders (PPD, n=24). Levels of neurofilament light chain, glial fibrillary acidic protein and phosphorylated-tau 181, apolipoprotein E genotype and late-onset AD polygenic risk scores were determined. Information-theoretic model selection identified discriminatory factors.

RESULTS: Glial fibrillary acidic protein was up to 3.5-fold higher in individuals with EOAD compared to other diagnostic categories. A combination of cognitive and blood biomarkers, but not the polygenic risk score, discriminated between diagnostic categories. Phosphorylated-tau 181 alone significantly discriminated between EOAD and nAD-ND causes.

DISCUSSION: Discriminating between EOAD, nAD-ND and PPD causes of young-onset neurocognitive symptoms is possible by combining cognitive profiles with blood biomarkers.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

KEYWORDS

young-onset dementia, apolipoprotein E, polygenic risk scores, neurofilament light, glial fibrillary acidic protein, phosphorylated tau 181, primary psychiatric disorder, neurodegeneration, Alzheimer's disease, multi-omic analyses

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

1. INTRODUCTION

Young-onset dementia, often defined as dementia diagnosed before the age of 65 years[1], is comprised of a heterogenous group of disorders that accounts for approximately 5% of all cases of dementia and have a prevalence of 119 per 100,000 individuals worldwide.[2,3] Diagnosing young-onset dementia is challenging due to the atypical symptoms and cognitive profiles that can overlap with other causes of neurocognitive disorders in this age group, such as primary psychiatric disorders (PPD). In fact, nearly one-third of individuals are misdiagnosed with PPD prior to the identification of a neurodegenerative cause.[4] Furthermore, nearly 40% of individuals with young-onset neurocognitive symptoms have their initial diagnosis revised during follow-up, reflecting low diagnostic certainty with standard clinical work-up.[5] Establishing a timely and accurate diagnosis is important to facilitate appropriate management and understand prognosis.

Blood biomarkers are emerging as powerful predictors of late-onset neurodegenerative conditions. Improvements in immunoassay technology, such as single molecule array (Simoa), have allowed for ultra-sensitive detection of protein biomarkers in blood, where concentrations are often in the pico- and femtomolar ranges.[6] Proteins such as neurofilament light chain (NfL) and glial fibrillary acidic protein (GFAP) are detected in blood samples from individuals with a variety of neurodegenerative conditions including late-onset Alzheimer's disease (LOAD), Lewy body disease and frontotemporal dementia[7-9] Furthermore, elevated levels of hyperphosphorylated species of tau, such as those phosphorylated at threonine 181 (p-tau181), have demonstrated high specificity in distinguishing LOAD from other neurodegenerative conditions.[10]

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

While different protein blood biomarkers reflect dynamic changes occurring in the setting of neurodegeneration, especially when symptoms and neuroimaging changes are well established, they vary in their ability to predict onset of disease prior to symptom onset. Contrastingly, genetics can be considered as capturing an individual's static risk of developing neurodegeneration. Genetic variants have been identified using genome-wide association studies (GWAS) for conditions such as LOAD, Lewy body disease and Parkinson's disease.[11-14] Unique polygenic risk scores (PRS) have been derived from GWAS findings to calculate an individual's risk of developing a particular neurodegenerative condition based on their combination of single nucleotide polymorphisms. For example, those in the highest 10% of PRS values have nearly a 1.9-fold increase in the risk of AD compared to those in the lowest 10%, and this effect is additive to age and the apolipoprotein E (*APOE*) locus, with the latter being the gene most strongly associated with LOAD.[13]

Leveraging multi-omic data by combining the static risk captured by genetic analyses such as the PRS, the dynamic risk identified by protein biomarkers, and standard clinical assessments can improve diagnostic accuracy for late-onset neurodegenerative conditions such as LOAD.[15-17] However, there is a paucity of data on how such an approach applies to diagnosing individuals with young-onset neurocognitive symptoms. Identifying clinical variables and blood biomarkers that are useful in discriminating young-onset dementia due to neurodegenerative causes from PPD may improve diagnostic accuracy and timeliness, thereby potentially reducing morbidity.

In this study, we sought to address these outstanding questions by analysing the Biomarkers in Younger-Onset Neurocognitive Disorders (BeYOND) cohort[18], a clinical cohort of individuals experiencing young-onset neurocognitive symptoms. In brief, BeYOND is a

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

longitudinal prospective observational cohort study evaluating clinical presentation, cognition, blood biomarkers and genetics. The inclusion criteria included individuals referred to the Neuropsychiatry Centre at the Royal Melbourne Hospital for a possible diagnosis of young-onset dementia who had psychiatric, behavioural, neurological and/or cognitive symptom onset prior to the age of 65 years. Participants were prospectively recruited during their index presentation. The study aimed to determine which combinations of clinical, cognitive, blood and genetic biomarker variables accurately differentiated causes of neurocognitive symptoms into the categories of early-onset AD (EOAD), non-AD neurodegeneration (nAD-ND) and PPD within this group (Supplemental File 1).

2. METHODS

2.1 Participant recruitment

The BeYOND study protocol has been described previously, with participant recruitment from June 2019 to December 2020 and follow-up over one year.[19] For this study, we evaluated cognition, genetics and blood biomarkers with clinical characteristics cross-sectionally. This project was approved by the Royal Melbourne Hospital/Melbourne Health Human Research Ethics Committee (MH 2018.371).

2.2 Diagnosis

Dementia diagnosis was made according to consensus criteria including the McKhann criteria for AD[20] and the Rascovsky criteria for behavioural-variant FTD.[21] Cerebrospinal fluid protein levels of amyloid-β and phosphorylated-tau at threonine 181 (p-tau181) were utilized for diagnosis of AD. Psychiatric diagnoses were made according to the Diagnostic Statistical Manual, Fifth Edition.[22]

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

2.3 Cognitive testing and classification

Participants underwent formal neuropsychological assessments. The overarching domains of global cognition, memory (delayed recall) and executive function were used for this study. Global cognitive function was determined using overall performance on the Neuropsychiatry Cognitive Assessment tool[23], which measures the five domains of cognition (attention, memory, visuospatial, executive and language functions). A range of domain-specific neuropsychological tests were used to further assess memory and executive function, as shown in Supplementary Table 1.

Raw scores from the neuropsychological tests were converted into Z-scores based on normative data stratified by age and education as provided in the test manuals.[24,25] A Z-score was categorized into "impaired" at 1.5 standard deviations (SD) below the normative score and "normal" (not impaired") if the Z-score was within the normal level. If a participant completed more than one neuropsychological test in a particular domain, the classification for each test was determined. If these were incongruent, then two clinical neuropsychologists evaluated the data and formed a consensus opinion about the participant's overall performance in that domain.

2.4 Biomarker analyses

EDTA blood samples were collected from fasted participants and stored at -80˚C. Plasma was tested in duplicates for selected biomarkers using Quanterix Simoa HD-X Neurology 2 Plex B for NfL and GFAP, and single Plex for p-tau181, according to the manufacturer's specifications. Age at assessment is the participant's age at which blood samples were collected.

2.5 APOE genotyping and polygenic risk score calculation

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

As described previously[26,27], DNA was extracted from peripheral whole-blood samples and *APOE* genotype determined using TaqMan® genotyping assays (rs7412, assay ID: C____904973_10; rs429358, assay ID: C____3084793_20), (Life Technologies, Carlsbad, CA) on a QuantStudio 12K FlexTM Real-Time-PCR system (Applied Biosystems, Foster City, CA), using TaqMan® GTXpressTM Master Mix (Life Technologies) as per manufacturer instructions. Allele loads were quantified for ε2 and ε4 and values were assigned to each individual (0,1 or 2) per allele, representing their *APOE* ε2 and ε4 status. Genetic data used to construct polygenic risk scores (PRS) were derived from an Axiom™ Precision Medicine Diversity Array (Applied Biosystems™) and imputed using the Haplotype Reference Consortium panel, for greatest cross-over of SNPs.

Beta weights for PRS calculation and *p*-value thresholds for variant inclusion were sourced from the International Genomics of Alzheimer's Project (IGAP) Alzheimer's disease genome-wide association study (GWAS).[11] PRS were calculated by first multiplying the number of effect alleles for each individual variant by the beta weights from the GWAS summary statistics. These weighted variant scores were then summed. PRS were calculated by including individual genetic variants at eight different GWAS *p*-value thresholds, from $p < 5 \times 10^{-8}$, reflecting the standard stringent GWAS-threshold, to $p < 0.1$, reflecting a suggestive-association threshold.

2.6 Statistical Analyses

R-package (version 4.3.1)[28] was used for statistical analyses. Due to the potential bias that may be observed with covariate-based multiple imputations using relatively small sample sizes, missing values were not imputed. For non-parametric data, the Kruskal-Wallis test was used when comparing across three or more groups, with the Dunn test applied for post-hoc

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

analyses. For non-parametric data comparison across two groups, the Wilcoxon rank sum exact test was used. For categorical data, Fisher's exact test, with pairwise post-hoc analyses using the R package rcompanion[29], was used. For other post-hoc analyses, Benjamini-Hochberg correction, with false discovery rate of 0.05 was used. p_{nom} represent un-adjusted (nominal) *p*-values and p_{adj} represent adjusted *p*-values using post-hoc analyses. Significance level α was set to 0.05.

For multinomial logistic regression, the R package nnet[30] was used, with comparison between the diagnostic categories of EOAD, nAD-ND and PPD. Only participants with complete data (on variables analysed) were used for model formation. Multinomial logistic regression was performed using diagnostic categories as the dependent variable and age at assessment, sex and individually-tested protein biomarker levels as the independent variables. PPD was set as the reference in the first model to obtain statistics for comparisons of EOAD vs PPD and nAD-ND vs PPD; nAD-ND was set as the reference in the second model for the EOAD vs nAD-ND comparison. p_{nom} values were calculated using the Wald statistic.

The R package MuMIn[31] was used for information-theoretic model selection of general linear mixed models composed of all available participant variables, with only participants with complete data being analyzed. DeLong test was used to compare resulting models.

Plots were made the R packages ggplot2, ggbreak and ggstatsplot.[32]

3. RESULTS

3.1 BeYOND participant demographics

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Seventy-two participants were recruited, with 65 of them classified within the diagnostic categories of either EOAD (n = 18), nAD-ND (n = 23), or PPD (n = 24, Table 1). nAD-ND causes consisted of behavioural variant frontotemporal dementia with definite frontotemporal lobar degeneration pathology (n = 6), probable behavioural variant frontotemporal dementia (n = 4), Parkinson disease (n = 3), Niemann-Pick disease type C (n = 3), cerebellar degeneration (n = 2), cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (n=1), progressive supranuclear palsy (n = 1), dementia with Lewy bodies (n = 1), vascular dementia (n = 1), and that of unclear etiology (n = 1). PPD causes consisted of depression (n = 7), schizophrenia (n = 4), bipolar affective disorder (n = 3), subjective cognitive impairment (n = 2), schizoaffective disorder (n = 2), obsessive compulsive disorder (n = 1), psychosis (n = 1), delusional disorder (n = 1), functional neurological disorder (n = 1) and 'other specified mental disorder' (n = 2). The EOAD, nAD-ND and PPD groups were similar for age of symptom onset, age at assessment, sex and *APOE* genotypes (Table 1). Missing data ranged between 0% and 26% (Table 1, Supplementary Table 2).

Of the seven individuals that were not classified as having one of the three diagnostic categories listed above, five had a diagnosis of mild cognitive impairment of unclear aetiology, one had autoimmune encephalitis and one had primary angiitis of the central nervous system. These seven individuals were excluded from further analyses as the primary aim of the study was to differentiate the three aforementioned diagnostic categories based on cognitive profiles and multi-omic blood biomarkers.

3.2 Cognitive profiles amongst diagnostic categories

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Executive function was similar between the EOAD, nAD-ND and PPD diagnostic groups (p_{nom} = 0.14) while global cognitive function and memory function were impaired in proportionally more individuals from the EOAD group than those from nAD-ND and PPD groups (Table 2 and Supplementary Figure 1). In particular, 100% of those in the EOAD group had global cognitive impairment, compared to 55% of those in the PPD group (p_{adj} < 0.005). Similarly, memory was classified as impaired in 82% of those with EOAD, compared to 26% of those with nAD-ND and 46% of those with PPD (p_{adj} < 0.005 for EOAD vs nAD-ND and p_{adj} < 0.05 for AD vs PPD).

3.3 Protein blood biomarker levels

Blood biomarkers levels for NfL, GFAP and p-tau181 were significantly different between the three categories, with levels higher in individuals with EOAD compared to the other categories (Figure 1, Table 2).

3.3.1 NfL

Median NfL levels were 1.4-fold higher in participants with EOAD (21 pg/ml) compared to nAD-ND (15 pg/ml, p_{adj} = 0.02, Supplementary Table 3) and 2.1-fold higher compared to those with PPD (10 pg/ml, p_{adj} = 5.6 x 10^{-5}). Median NfL levels were also elevated in nAD-ND compared to PPD (1.5-fold, p_{adj} = 0.02).

3.3.2 GFAP

In those with EOAD, median GFAP levels were 3.5-fold higher (198 pg/ml) than in those with PPD (56 pg/ml, p_{adj} = 1.3 x 10^{-5}) and nearly 2.6-fold higher those with nAD-ND (77 pg/ml, 1.8-fold, p_{adj} = 4.0 x 10^{-4}). While median GFAP levels were over 1.3-fold higher in nAD-ND compared to PPD, statistical significance was not found (p_{adj} = 0.14).

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

3.3.3 P-tau181

Median p-tau181 levels were nearly 2-fold higher in EOAD (3.33 pg/ml) compared to nAD-ND (1.69 pg/ml, p_{adj} = 3.2 x 10^{-5}) and 2.9-fold higher compared to PPD (1.15 pg/ml, p_{adj} = 1.9 x 10^{-8}). Similarly, nAD-ND had 1.5-fold higher median p-tau181 levels than PPD (p_{adj} = 0.04). Pairwise comparisons between diagnoses with the adjusted (Dunn) *p*-values for the three protein blood biomarkers are provided in Supplementary Table 3.

3.3.4 Covariate analysis

We further tested the association between protein blood biomarkers and diagnostic categories, adjusted for age at assessment and sex. Participants without missing data for these variables were analysed using multinomial logistic regression models. Demographic variables between the diagnostic categories of participants analyzed in these models were similar (Supplementary Table 4).

Compared to PPD, NfL demonstrated an odds ratio (OR) = 1.16 [95% CI: 1.05-1.29] for EOAD (p_{adj} = 0.011) and OR = 1.16 [1.04-1.28] for nAD-ND (p_{adj} = 0.011, Supplementary Table 5). Comparing between EOAD and nAD-ND, OR = 1.00 ([0.99-1.02], p_{adj} = 0.589). While the OR for GFAP was statistically significant (p_{adj} < 0.05) for EOAD vs PPD and EOAD vs nAD-ND comparisons, the magnitude of the effect was quite small with OR values between 1.01 and 1.02. Contrastingly, p-tau181 analyses yielded a large OR of 49.90 ([6.03-410.53], p_{adj} = 0.003) and of 20.09 ([2.92-138.47], p_{adj} = 0.007) for EOAD vs PPD and EOAD vs nAD-ND, respectively.

3.4 Polygenic risk score associations

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

There were no statistically significant associations between PRS values, calculated using eight different LOAD GWAS[11] *p*-value thresholds, and diagnostic categories (Supplementary Figure 2, Supplementary Tables 6 and 7). There was a trend for association using the stringent GWAS *p*-value threshold of 5×10^{-8} to calculate the PRS, with a median PRS value of 5.50 (interquartile range [IQR]: 4.14-6.14) in those with EOAD compared to 3.91 (IQR: 2.68-6.89) in those with nAD-ND and 2.45 (IQR: 0.70-4.69) in those with PPD ($p_{nom} = 0.085$ for group comparison, Supplementary Figure 3).

3.5 Information-theoretic model selection identifies key variables for pairwise discrimination between diagnostic categories.

To determine which combination of clinical and biomarker variables best discriminated between the diagnostic categories in the BeYOND cohort, we utilized an information-theoretic approach for model selection. Models with the lowest Akaike information criterion with correction for finite sample sizes (AIC) were selected as they represented the optimal trade-off between model fit and model complexity.[15,31] Variables included in the model selection were age at symptom onset, age at assessment, sex, cognitive function (global, memory and executive), blood biomarker levels (NfL, GFAP, and p-tau181), *APOE* ε2 and ε4 allele status and PRS (using GWAS *p*-value threshold of 5×10^{-8}). Due to the model structure, only individuals with complete data could be analyzed, yielding 46 participants across the three diagnostic categories with similar demographic variables amongst them (Supplementary Tables 8-10).

For discriminating between EOAD and PPD within the BeYOND cohort, the best model (Model 1A, as denoted by producing the lowest AIC) contained five variables: *APOE* ε2 and ε4 status, global cognitive function, NfL and p-tau181 (Table 3). P-tau181 had the largest

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

odds ratio of the included variables in Model 1, with an OR = 1.48 [95% CI: 1.33-1.65]. Interestingly, *APOE* ε2 and ε4 status both had OR < 1 for association with EOAD, compared to PPD. Three parsimonious models (Models 2A – 4A) equivalent to Model 1A (as denoted by having a ΔAIC < 2 compared to Model 1A) were identified. The area under the receiver operating curves (AUCs) for Models 1A-4A were very high (Figure 2A), ranging from 0.975 to 1.000 and were not statistically different (DeLong p_{nom} > 0.05, Table 3). We next tested how models containing only one of the variables identified in Model 1A performed. Models 5A-7A and 9A were inferior to Models 1A-4A with respect to both ΔAIC and AUC values. Model 8A, containing only the p-tau181 variable, demonstrated a high AUC of 0.954 [95% CI: 0.887-1.000], but with a ΔAIC > 2, indicated an inferior model compared to Models 1A-4A. Assessing all models analysed for information-theoretic model selection revealed that p-tau181 had a sum of weights (SoW) = 1.00 (indicating this variable was present in all models), while only NfL, *APOE* ε2 and global cognitive function has SoW ≥ 0.5 (indicating the importance of these variables as they were present in most of the well supported models based on AIC, Supplementary Figure 4).

To discriminate EOAD from nAD-ND, the best model (Model 1B) required only the variable p-tau181 with an AUC of 0.950 [95% CI: 0.877-1.000] and an OR = 1.63 [95% CI: 1.38-1.93] (Supplementary Table 11, Figure 2B). P-tau181 had a SoW = 1 (Supplementary Figure 4). No other variable had a SoW ≥ 0.5, indicating that no other variables were considered significant for this comparison. In determining which variables would discriminate between all neurodegenerative cases (combining EOAD and nAD-ND) and PPD, the best model (Model 1C) contained seven variables: age at assessment, *APOE* ε2 and ε4 status, global cognitive function, PRS, p-tau181 and sex (Supplementary Table 12, Figure 2C), yielding an AUC = 0.975 [95% CI: 0.940-1.000]. Interestingly, global cognitive function had the largest

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

magnitude of effect size with OR of 0.58 (95% CI: 0.45-0.75), followed by sex (OR = 1.19), p-tau181 (OR = 1.27) and PRS (OR = 1.12). Two parsimonious models (Models 2C and 3C) each contained a different combination of five variables and demonstrated similar AUC values of 0.957 and 0.947, respectively. Models 4C-11C, each containing only one variable from Model 1C, were all inferior to Models 1C-3C with ΔAIC > 2. SoW for p-tau181 and global cognitive function was 0.99 and 0.95, respectively (Supplementary Figure 4). Other variables with SoW ≥ 0.5 included age at assessment (0.71), *APOE* ε2 (0.66), sex (0.61), PRS (0.60) and memory function (0.55), indicating that more variables may contribute to discriminating between neurodegeneration and PPD compared to between EOAD vs PPD and EOAD vs nAD-ND.

4. DISCUSSION

We sought to determine how clinical, cognitive, protein blood biomarker and genetic variables differentiated causes of neurocognitive symptoms into the categories of EOAD, nAD-ND and PPD within the BeYOND cohort. We found that levels of NfL, GFAP and p-tau181 were higher in individuals with EOAD compared to nAD-ND and PPD. The association of the PRS (based on a stringent GWAS *p*-value threshold of 5×10^{-8}) and diagnostic categories did not reach statistical significance, but was suggestive ($p_{nom} = 0.085$). A model containing global cognitive function and levels of p-tau181 and NfL significantly discriminated between EOAD and PPD, while a model containing only p-tau181 significantly discriminated between EOAD and nAD-ND causes.

Much attention has been given to how blood biomarkers differentiate and diagnose causes of late-onset (older than 65 years of age) cognitive impairment, especially LOAD.[8,33-35] Numerous proteins have been implicated. In particular, NfL has been implicated in a wide

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

range of neurodegenerative conditions[9] including LOAD[36], frontotemporal dementia[37] and amyotrophic lateral sclerosis.[38] This protein biomarker has shown utility in distinguishing neurodegenerative conditions from PPD.[7,39] In our cohort, we found that NfL could function similarly, with individuals with EOAD demonstrating higher blood levels than those with nAD-ND and PPD. NfL continued to differentiate between neurodegenerative and PPD cases when adjusted for age at assessment and sex, factors which can influence protein blood biomarker levels.[7,40-42] The predictive power of NfL, however, was not different (OR = 1.00) between EOAD and nAD-ND causes, supporting emerging data that NfL can be considered as a general marker of neurodegeneration, and not as a specific marker for any particular neurodegenerative cause.[43]

We found elevated GFAP levels in those with EOAD compared to nAD-ND and PPD. However, odds ratio for GFAP (OR = 1.01-1.02) was low for EOAD vs PPD and nAD-ND, when adjusting for age at assessment and sex. This differs from the conclusion reached in a prospective study of 110 individuals with EOAD and 50 controls, where blood GFAP levels demonstrated an AUC of 96% in discriminating between these two groups.[44] A potential confounder in comparing these two studies is that the comparator group is individuals with PPD in our study, while healthy controls were used in the other, as well as the difference in samples sizes between the two studies. Overall, our findings support the hypothesis that GFAP may be a marker in a number of neurodegenerative conditions including EO/LOAD, frontotemporal dementia and Lewy body dementia.[45-48] More work is needed to elucidate what blood GFAP levels add in discerning the cause of neurocognitive symptoms, especially in the younger populations.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

A strength of this study is the identification of specific variables that contribute to diagnosis, as the diagnosis of neurodegenerative diseases is moving towards a multi-omics approach.[48] To the best of our knowledge, this is the first use of information-theoretic model selection[31] of multi-omic variables in a clinical cohort of individuals with young-onset neurocognitive symptoms. A benefit of multi-model inference is the ability to determine important variables in a less biased manner as all studied variables can be included in these analyses, with the best model (as quantified by the lower AIC value), as well as equivalent parsimonious models (as those with ΔAIC < 2 from the best model), identified. Using this approach, we found that diagnostic models that included global cognitive function and the blood biomarker levels of NfL and p-tau181 were sufficient to discriminate between individuals with EOAD and PPD. Similarly, p-tau181 alone was sufficient to discriminate EOAD from nAD-ND cases. Discriminating between all causes of neurodegeneration from PPD required more variables including *APOE* ε2 status, sex, p-tau181 levels, as well as global cognitive and memory function. These findings are similar to model selections performed in LOAD, where inclusion of protein blood biomarkers (p-tau181 or p-tau217), *APOE* status, brain imaging and cognitive testing (memory and executive) were able to predict conversion to LOAD in individuals with subjective cognitive impairment or mild cognitive impairment.[15] While our study did not have access to standardized neuroimaging or p-tau217 data, it is encouraging that many of the variables overlap between LOAD and EOAD. Additional studies of individuals with young-onset neurocognitive symptoms may reveal divergent variables between early-onset and late-onset causes. These findings support further hypothesis testing that single blood biomarkers may be used in the appropriate context to risk stratify individuals for neurodegeneration, and specifically for EOAD, which may be both useful in clinical care and clinical trials.[8,49,50]

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Multiple studies have demonstrated that p-tau181 is a very specific biomarker for LOAD.[10,51,52] However, studies are lacking in describing the relationship between EOAD and p-tau181, especially in non-familial cases. One study did find elevated cerebrospinal fluid levels of p-tau181 in those with EOAD compared to healthy controls.[53] Our study supports this finding in blood, with plasma p-tau181 levels demonstrating a robust ability to discriminate EOAD from other causes of young-onset neurocognitive symptoms. Further research into this area may help establish p-tau181 as a marker for AD in both early- and late-onset cases.

PRS for AD have been constructed based on late-onset cases, which has shown strong ability to discriminate LOAD cases from healthy controls and even predict progression from mild cognitive impairment to LOAD.[13] In this study, we utilized a PRS derived from a GWAS including over 90,000 clinically diagnosed LOAD cases and controls, which identified 25 genetic loci associated with LOAD.[11] In our cohort, there was a trend towards a higher PRS in those with EOAD compared to nAD-ND and PPD when restricting the PRS to include only those SNPs with stringent genome-wide significance ($p < 5 \times 10^{-8}$). Overall, the transferability of a PRS derived from a LOAD population to an EOAD population is mixed, with some studies demonstrating poor correlation.[54,55] Other studies have shown a potential role of a LOAD-derived PRS in predicting EOAD cases, though differences in age of participants and *APOE* status may confound direct comparisons.[56-59] Our findings, along with those previously published, indicate the need for an EOAD-derived PRS, especially in individuals with non-familial cases.[60]

The relationship between *APOE* ε4 status and odds of being diagnosed with EOAD versus PPD was a surprising result, as the model selection suggested that a higher *APOE* ε4 allele

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

count was associated with a higher probability of being diagnosed with PPD compared to EOAD. While *APOE* ε4 status is associated with a younger age-of-onset in LOAD and familial cases of AD, its role in sporadic cases of AD is less clear.[61,62] This study was not powered to detail the relationship between *APOE* ε4 status, age of onset and diagnosis. Further work is needed to elucidate this relationship and to determine if our findings are due to differences in the genetic contributions to EOAD or limitations due to sample size.

There are additional limitations to our study that warrant careful consideration. A significant limitation is that the BeYOND cohort is relatively small with 72 total participants. Diagnostic heterogeneity, as to the cause of young-onset neurocognitive symptoms, was very high and necessitated formation of three diagnostic categories to perform analyses. Therefore, the effect magnitudes for included variables (cognitive testing and blood biomarkers) may change greatly depending on the relative proportions of the different diagnostic categories within the study cohort. We attempted to mitigate this issue by accounting for covariates and performing information-theoretic model selection, reducing bias and penalizing model over-fitting. However, these are also affected, to a certain degree, by the diagnostic proportions included in each model. External cohort studies of individuals with young-onset neurocognitive symptoms are needed to validate our findings. With respect to biomarkers, we did not have access to standardized neuroimaging or emerging protein biomarkers like p-tau231, with the latter showing a differing role in LOAD diagnosis.[50] We also performed PRS using a GWAS derived from LOAD cases;[11] there are recent LOAD PRS with more implicated genetic loci.[13,57] However, these recent studies use 'proxy' cases, where an individual who is asymptomatic is considered a GWAS case if they have at least one first degree relative diagnosed with LOAD, which may distort the relative contribution of GWAS loci to the risk of developing AD.[63] It is unclear how proxy cases affect LOAD GWAS

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

transferability in EOAD cases. Biomarker results presented in this study are static measurements from a single time point for each participant. Further studies with serial biomarker measurements may yield dynamic changes that may improve diagnostic accuracy and model simplicity.

The BeYOND cohort studied here is similar to other studies in its distribution of diagnostic causes for young-onset neurocognitive symptoms, thereby adding to this field of research and supporting the generalizability of our results.[64,65] Overall, this study adds to our understanding of the multi-omic variables aiding in the diagnosis of young-onset neurocognitive symptoms. Our findings support further research into the use of protein blood biomarkers and cognitive profiles in the diagnostic pathway as well as identify the need for the development and validation of an EOAD PRS.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Acknowledgements & Funding

We would like to thank all the participants, their families and caregivers involved in the BeYOND study.

OGB was supported by the Melbourne Genomics Health Alliance Genomics Immersion Fellowship. Funders did not have a role in the generation or analysis of the data or in the development of the manuscript.

Data Availability

Participant data access can be discussed with Samantha M. Loi. Standard R packages were used for data analyses and can be sourced directly from public repositories.

Conflicts of Interest

The authors declare that there are no conflicts of interests in this study.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

References

1. van de Veen D, Bakker C, Peetoom K, et al. An Integrative Literature Review on the Nomenclature and Definition of Dementia at a Young Age. *J Alzheimers Dis*. 2021;83(4):1891-1916. doi:10.3233/JAD-210458

2. Hendriks S, Peetoom K, Bakker C, et al. Global Prevalence of Young-Onset Dementia: A Systematic Review and Meta-analysis. *JAMA Neurol*. Sep 1 2021;78(9):1080-1090. doi:10.1001/jamaneurol.2021.2161

3. Loi SM, Cations M, Velakoulis D. Young-onset dementia diagnosis, management and care: a narrative review. *Med J Aust*. Mar 6 2023;218(4):182-189. doi:10.5694/mja2.51849

4. Woolley JD, Khan BK, Murthy NK, Miller BL, Rankin KP. The diagnostic challenge of psychiatric symptoms in neurodegenerative disease: rates of and risk factors for prior psychiatric diagnosis in patients with early neurodegenerative disease. *J Clin Psychiatry*. Feb 2011;72(2):126-33. doi:10.4088/JCP.10m06382oli

5. Tsoukra P, Velakoulis D, Wibawa P, et al. The Diagnostic Challenge of Young-Onset Dementia Syndromes and Primary Psychiatric Diseases: Results From a Retrospective 20-Year Cross-Sectional Study. *J Neuropsychiatry Clin Neurosci*. Winter 2022;34(1):44-52. doi:10.1176/appi.neuropsych.20100266

6. Mankhong S, Kim S, Lee S, et al. Development of Alzheimer's Disease Biomarkers: From CSF- to Blood-Based Biomarkers. *Biomedicines*. Apr 5 2022;10(4)doi:10.3390/biomedicines10040850

7. Eratne D, Kang M, Malpas C, et al. Plasma neurofilament light in behavioural variant frontotemporal dementia compared to mood and psychotic disorders. *Aust N Z J Psychiatry*. Jan 2024;58(1):70-81. doi:10.1177/00048674231187312

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

8. Hansson O, Blennow K, Zetterberg H, Dage J. Blood biomarkers for Alzheimer's disease in clinical practice and trials. *Nat Aging*. May 2023;3(5):506-519. doi:10.1038/s43587-023-00403-3

9. Bacioglu M, Maia LF, Preische O, et al. Neurofilament Light Chain in Blood and CSF as Marker of Disease Progression in Mouse Models and in Neurodegenerative Diseases. *Neuron*. Jul 20 2016;91(2):494-496. doi:10.1016/j.neuron.2016.07.007

10. Janelidze S, Mattsson N, Palmqvist S, et al. Plasma P-tau181 in Alzheimer's disease: relationship to other biomarkers, differential diagnosis, neuropathology and longitudinal progression to Alzheimer's dementia. *Nat Med*. Mar 2020;26(3):379-386. doi:10.1038/s41591-020-0755-1

11. Kunkle BW, Grenier-Boley B, Sims R, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat Genet*. Mar 2019;51(3):414-430. doi:10.1038/s41588-019-0358-2

12. Chia R, Sabir MS, Bandres-Ciga S, et al. Genome sequencing analysis identifies new loci associated with Lewy body dementia and provides insights into its genetic architecture. *Nat Genet*. Mar 2021;53(3):294-303. doi:10.1038/s41588-021-00785-3

13. Bellenguez C, Küçükali F, Jansen IE, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet*. Apr 2022;54(4):412-436. doi:10.1038/s41588-022-01024-z

14. Kim JJ, Vitale D, Otani DV, et al. Multi-ancestry genome-wide association meta-analysis of Parkinson's disease. *Nat Genet*. Jan 2024;56(1):27-36. doi:10.1038/s41588-023-01584-8

15. Palmqvist S, Tideman P, Cullen N, et al. Prediction of future Alzheimer's disease dementia using plasma phospho-tau combined with other accessible measures. *Nat Med*. 06 2021;27(6):1034-1042. doi:10.1038/s41591-021-01348-z

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

16. Ramanan VK, Gebre RK, Graff-Radford J, et al. Genetic risk scores enhance the diagnostic value of plasma biomarkers of brain amyloidosis. *Brain*. Jun 5 2023;doi:10.1093/brain/awad196

17. Bhalala OG, Watson R, Yassi N. Multi-Omic Blood Biomarkers as Dynamic Risk Predictors in Late-Onset Alzheimer's Disease. *Int J Mol Sci*. Jan 19 2024;25(2)doi:10.3390/ijms25021231

18. Loi SME, D.; Walterfang, M.; Kelso, W; Farrand, S.; Goh, A. M. Y.; Chadunow, C.; Ravenfar, P. S.; Watson, R.; Evans, A. H.; Vidaurre, L.; Valkoulis, D. BeYOND: Biomarkers in younger-onset neurocognitive disorders — The first 50 participants. presented at: Alzheimer's Association International Conference; December 7 2020 2020;

19. Loi SME, D; Cadwallader, C; Ravanfar, P; Chadunow, C; Vidaurre, L; Farrand, S; Kelso, W; Goh, A. M. Y.; Watson, R; Evans, A; Walterfang, M; Velakoulis, D. Investigating the clinical utility of biomarkers and other novel tests in younger onset neurocognitive disorders: the BeYOND study, protocol for a longitudinal clinical study in a real-world setting. *medRxiv*. 2021-01-01 00:00:00 2021;

20. McKhann GM, Knopman DS, Chertkow H, 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*. May 2011;7(3):263-9. doi:10.1016/j.jalz.2011.03.005

21. Rascovsky K, Hodges JR, Knopman D, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*. Sep 2011;134(Pt 9):2456-77. doi:10.1093/brain/awr179

22. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. 2013.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

23. Walterfang M, Siu R, Velakoulis D. The NUCOG: validity and reliability of a brief cognitive screening tool in neuropsychiatric patients. *Aust N Z J Psychiatry*. Nov-Dec 2006;40(11-12):995-1002. doi:10.1080/j.1440-1614.2006.01923.x

24. Griffith S, Wesselingh R, Broadley J, et al. Psychometric deficits in autoimmune encephalitis: A retrospective study from the Australian Autoimmune Encephalitis Consortium. *Eur J Neurol*. Aug 2022;29(8):2355-2366. doi:10.1111/ene.15367

25. Griffith SP, Wesselingh R, Seery N, et al. Characterizing cognitive function in patients with autoimmune encephalitis: an Australian prospective study. *J Neurol*. Jan 2024;271(1):310-324. doi:10.1007/s00415-023-11967-w

26. Porter T, Burnham SC, Milicic L, et al. Utility of an Alzheimer's Disease Risk-Weighted Polygenic Risk Score for Predicting Rates of Cognitive Decline in Preclinical Alzheimer's Disease: A Prospective Longitudinal Study. *J Alzheimers Dis*. 2018;66(3):1193-1211. doi:10.3233/JAD-180713

27. Porter T, Burnham SC, Savage G, et al. A Polygenic Risk Score Derived From Episodic Memory Weighted Genetic Variants Is Associated With Cognitive Decline in Preclinical Alzheimer's Disease. *Front Aging Neurosci*. 2018;10:423. doi:10.3389/fnagi.2018.00423

28. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2023. https://www.R-project.org/

29. *rcompanion: Functions to Support Extension Education Program Evaluation*. Rutgers Cooperative Extension; 2023. https://CRAN.R-project.org/package=rcompanion/

30. Venables WNR, B. D. *Modern Applied Statistics with S*. Fourth ed. Springer; 2002.

31. *MuMIn: Multi-Model Inference*. 2023. https://CRAN.R-project.org/package=MuMIn

32. Patil I. Visualizations with statistical details: The 'ggstatsplot' approach. *Journal of Open Source Software*. 2021;6(61)doi:10.21105/joss.03167

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

33. Qu Y, Ma YH, Huang YY, et al. Blood biomarkers for the diagnosis of amnestic mild cognitive impairment and Alzheimer's disease: A systematic review and meta-analysis. *Neurosci Biobehav Rev*. Sep 2021;128:479-486. doi:10.1016/j.neubiorev.2021.07.007

34. Chouliaras L, Thomas A, Malpetti M, et al. Differential levels of plasma biomarkers of neurodegeneration in Lewy body dementia, Alzheimer's disease, frontotemporal dementia and progressive supranuclear palsy. *J Neurol Neurosurg Psychiatry*. Jun 2022;93(6):651-658. doi:10.1136/jnnp-2021-327788

35. Saunders TS, Pozzolo FE, Heslegrave A, et al. Predictive blood biomarkers and brain changes associated with age-related cognitive decline. *Brain Commun*. 2023;5(3):fcad113. doi:10.1093/braincomms/fcad113

36. Mattsson N, Andreasson U, Zetterberg H, Blennow K, Alzheimer's Disease Neuroimaging I. Association of Plasma Neurofilament Light With Neurodegeneration in Patients With Alzheimer Disease. *JAMA Neurol*. May 1 2017;74(5):557-566. doi:10.1001/jamaneurol.2016.6117

37. Rohrer JD, Woollacott IO, Dick KM, et al. Serum neurofilament light chain protein is a measure of disease intensity in frontotemporal dementia. *Neurology*. Sep 27 2016;87(13):1329-36. doi:10.1212/WNL.0000000000003154

38. Gaiani A, Martinelli I, Bello L, et al. Diagnostic and Prognostic Biomarkers in Amyotrophic Lateral Sclerosis: Neurofilament Light Chain Levels in Definite Subtypes of Disease. *JAMA Neurol*. May 1 2017;74(5):525-532. doi:10.1001/jamaneurol.2016.5398

39. Eratne D, Loi SM, Li QX, et al. Cerebrospinal fluid neurofilament light chain differentiates primary psychiatric disorders from rapidly progressive, Alzheimer's disease and frontotemporal disorders in clinical settings. *Alzheimers Dement*. Nov 2022;18(11):2218-2233. doi:10.1002/alz.12549

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

40. Simren J, Andreasson U, Gobom J, et al. Establishment of reference values for plasma neurofilament light based on healthy individuals aged 5-90 years. *Brain Commun*. 2022;4(4):fcac174. doi:10.1093/braincomms/fcac174

41. Tsiknia AA, Edland SD, Sundermann EE, et al. Sex differences in plasma p-tau181 associations with Alzheimer's disease biomarkers, cognitive decline, and clinical progression. *Mol Psychiatry*. Oct 2022;27(10):4314-4322. doi:10.1038/s41380-022-01675-8

42. Sarto J, Esteller-Gauxax D, Tort-Merino A, et al. Impact of demographics and comorbid conditions on plasma biomarkers concentrations and their diagnostic accuracy in a memory clinic cohort. *J Neurol*. Dec 27 2023;doi:10.1007/s00415-023-12153-8

43. Ashton NJ, Janelidze S, Al Khleifat A, et al. A multicentre validation study of the diagnostic value of plasma neurofilament light. *Nat Commun*. Jun 7 2021;12(1):3400. doi:10.1038/s41467-021-23620-z

44. Lv X, Cheng Z, Wang Q, et al. High burdens of phosphorylated tau protein and distinct precuneus atrophy in sporadic early-onset Alzheimer's disease. *Sci Bull (Beijing)*. Nov 30 2023;68(22):2817-2826. doi:10.1016/j.scib.2023.10.019

45. Oeckl P, Halbgebauer S, Anderl-Straub S, et al. Glial Fibrillary Acidic Protein in Serum is Increased in Alzheimer's Disease and Correlates with Cognitive Impairment. *J Alzheimers Dis*. 2019;67(2):481-488. doi:10.3233/JAD-180325

46. Benussi A, Ashton NJ, Karikari TK, et al. Serum Glial Fibrillary Acidic Protein (GFAP) Is a Marker of Disease Severity in Frontotemporal Lobar Degeneration. *J Alzheimers Dis*. 2020;77(3):1129-1141. doi:10.3233/JAD-200608

47. Cicognola C, Janelidze S, Hertze J, et al. Plasma glial fibrillary acidic protein detects Alzheimer pathology and predicts future conversion to Alzheimer dementia in patients with mild cognitive impairment. *Alzheimers Res Ther*. Mar 27 2021;13(1):68. doi:10.1186/s13195-021-00804-9

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

48. Abdelhak A, Foschi M, Abu-Rumeileh S, et al. Blood GFAP as an emerging biomarker in brain and spinal cord disorders. *Nat Rev Neurol*. Mar 2022;18(3):158-172. doi:10.1038/s41582-021-00616-3

49. Hansson O, Edelmayer RM, Boxer AL, et al. The Alzheimer's Association appropriate use recommendations for blood biomarkers in Alzheimer's disease. *Alzheimers Dement*. Dec 2022;18(12):2669-2686. doi:10.1002/alz.12756

50. Self WK, Holtzman DM. Emerging diagnostics and therapeutics for Alzheimer disease. *Nat Med*. Sep 2023;29(9):2187-2199. doi:10.1038/s41591-023-02505-2

51. Karikari TK, Pascoal TA, Ashton NJ, et al. Blood phosphorylated tau 181 as a biomarker for Alzheimer's disease: a diagnostic performance and prediction modelling study using data from four prospective cohorts. *Lancet Neurol*. May 2020;19(5):422-433. doi:10.1016/S1474-4422(20)30071-5

52. Thijssen EH, La Joie R, Wolf A, et al. Diagnostic value of plasma phosphorylated tau181 in Alzheimer's disease and frontotemporal lobar degeneration. *Nat Med*. Mar 2020;26(3):387-397. doi:10.1038/s41591-020-0762-2

53. Kaur G, Poljak A, Braidy N, Crawford JD, Lo J, Sachdev PS. Fluid Biomarkers and APOE Status of Early Onset Alzheimer's Disease Variants: A Systematic Review and Meta-Analysis. *J Alzheimers Dis*. 2020;75(3):827-843. doi:10.3233/JAD-200052

54. Mantyh WG, Cochran JN, Taylor JW, et al. Early-onset Alzheimer's disease explained by polygenic risk of late-onset disease? *Alzheimers Dement (Amst)*. Oct-Dec 2023;15(4):e12482. doi:10.1002/dad2.12482

55. Schott JM, Crutch SJ, Carrasquillo MM, et al. Genetic risk factors for the posterior cortical atrophy variant of Alzheimer's disease. *Alzheimers Dement*. Aug 2016;12(8):862-71. doi:10.1016/j.jalz.2016.01.010

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

56. Cruchaga C, Del-Aguila JL, Saef B, et al. Polygenic risk score of sporadic late-onset Alzheimer's disease reveals a shared architecture with the familial and early-onset forms. *Alzheimers Dement*. Feb 2018;14(2):205-214. doi:10.1016/j.jalz.2017.08.013

57. de Rojas I, Moreno-Grau S, Tesi N, et al. Common variants in Alzheimer's disease and risk stratification by polygenic risk scores. *Nat Commun*. Jun 07 2021;12(1):3417. doi:10.1038/s41467-021-22491-8

58. Fulton-Howard B, Goate AM, Adelson RP, et al. Greater effect of polygenic risk score for Alzheimer's disease among younger cases who are apolipoprotein E-epsilon4 carriers. *Neurobiol Aging*. Mar 2021;99:101 e1-101 e9. doi:10.1016/j.neurobiolaging.2020.09.014

59. Huq AJ, Fulton-Howard B, Riaz M, et al. Polygenic score modifies risk for Alzheimer's disease in APOE epsilon4 homozygotes at phenotypic extremes. *Alzheimers Dement (Amst)*. 2021;13(1):e12226. doi:10.1002/dad2.12226

60. Mol MO, van der Lee SJ, Hulsman M, et al. Mapping the genetic landscape of early-onset Alzheimer's disease in a cohort of 36 families. *Alzheimers Res Ther*. Jun 1 2022;14(1):77. doi:10.1186/s13195-022-01018-3

61. Olarte L, Schupf N, Lee JH, et al. Apolipoprotein E epsilon4 and age at onset of sporadic and familial Alzheimer disease in Caribbean Hispanics. *Arch Neurol*. Nov 2006;63(11):1586-90. doi:10.1001/archneur.63.11.1586

62. Liu CC, Liu CC, Kanekiyo T, Xu H, Bu G. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. *Nat Rev Neurol*. Feb 2013;9(2):106-18. doi:10.1038/nrneurol.2012.263

63. Escott-Price V, Hardy J. Genome-wide association studies for Alzheimer's disease: bigger is not always better. *Brain Commun*. 2022;4(3):fcac125. doi:10.1093/braincomms/fcac125

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

64. Rossor MN, Fox NC, Mummery CJ, Schott JM, Warren JD. The diagnosis of young-onset dementia. *Lancet Neurol*. Aug 2010;9(8):793-806. doi:10.1016/S1474-4422(10)70159-9

65. Loi SM, Eratne D, Goh AMY, et al. A 10 year retrospective cohort study of inpatients with younger-onset dementia. *Int J Geriatr Psychiatry*. Feb 2021;36(2):294-301. doi:10.1002/gps.5424

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

BeYOND Manuscript
Figures

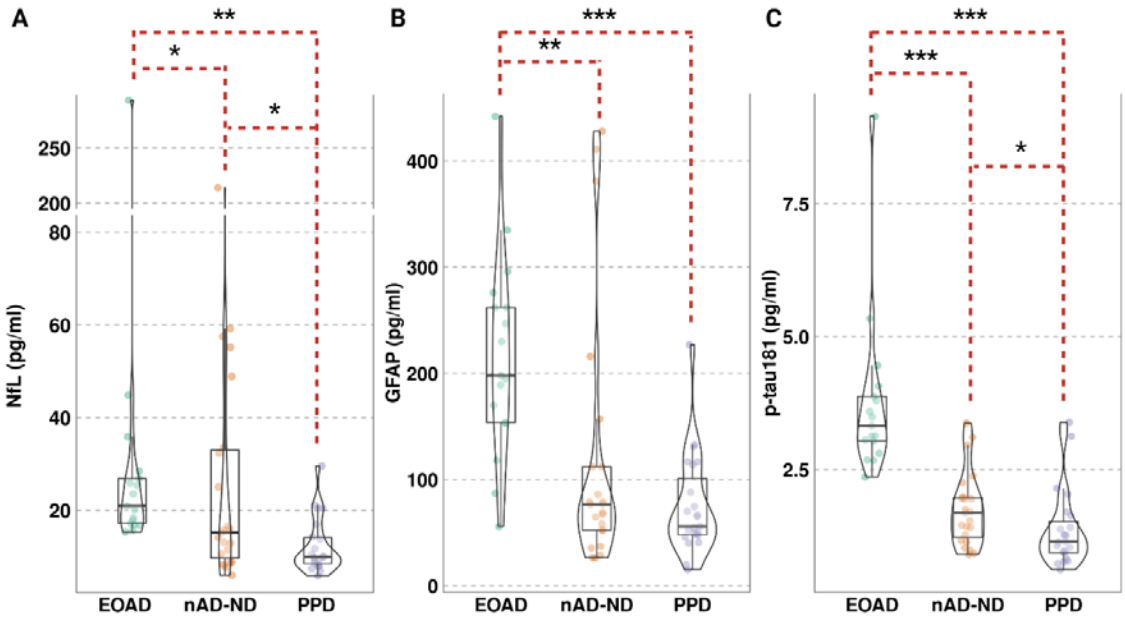

Figure 1: Blood biomarkers levels for NfL, GFAP and p-tau181 by diagnostic category. Violin plots of blood biomarker levels (pg/ml) for (A) neurofilament light chain (NfL), (B) glial fibrillary acidic protein (GFAP) and (C) phosphorylated-tau 181 (p-tau181) per diagnostic categories of early-onset Alzheimer's disease (EOAD), non-AD neurodegeneration (nAD-ND) and primary psychiatric disorders (PPD). Boxplots represent the median and interquartile range. Adjusted (Dunn) *P*-value notation* $p_{adj} < 0.05$, ** $p_{adj} < 5 \times 10^{-3}$, *** $p_{adj} < 5 \times 10^{-5}$.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

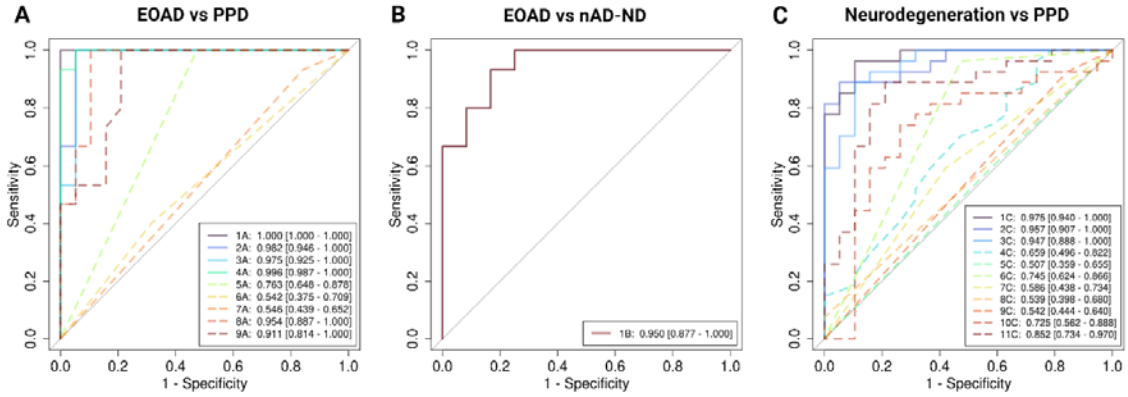

Figure 2: Area under the receiver operating curve plots for selected models discriminating between diagnostic categories. Best model (Model 1A/B/C), parsimonious models (solid lines beginning with Model 2A/B/C) and single variable models (dashed lines) in discriminating between A) early-onset Alzheimer's disease (EOAD) and primary psychiatric disorders (PPD), B) EOAD and non-AD neurodegenerative cases (nAD-ND), and C) neurodegenerative cases (of all types) and PPD. Values for each model indicate area under the receiver operating curve (AUC) and 95% confidence interval. Variables included in each model for the different diagnostic category comparisons are listed in Table 3 and Supplementary Tables 11 and 12.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Table 1. BeYOND Study Demographics and *APOE* Genotypes.

	Overall N = 65	EOAD N = 18	nAD-ND N = 23	PPD N = 24
Age of Onset[1]	56 (50, 60)	55 (52, 57)	59 (48, 61)	56 (50, 57)
*(Missing)**	*7 (11%)*	*0*	*6 (26%)*	*1 (4.2%)*
Age at Assessment[1]	58 (55, 62)	59 (55, 62)	61 (54, 64)	57 (54, 60)
Sex[2]				
Female	27 (42%)	9 (50%)	8 (35%)	10 (42%)
Male	38 (58%)	9 (50%)	15 (65%)	14 (58%)
***APOE* Genotype**[2]				
ε2/ε3	5 (8.8%)	1 (6.3%)	1 (5.3%)	3 (14%)
ε3/ε3	32 (56%)	8 (50%)	12 (63%)	12 (55%)
ε3/ε4	18 (32%)	7 (44%)	4 (21%)	7 (32%)
ε4/ε4	2 (3.5%)	0 (0%)	2 (11%)	0 (0%)
*(Missing)**	*8 (12%)*	*2 (11%)*	*4 (17%)*	*2 (8.3%)*

[1]Median (interquartile range).

[2]n (%).

*Number (percentage) of participants with missing/unknown data for particular variable.

If not indicated, there were no missing data for the variable.

APOE , apolipoprotein E; EOAD, early-onset Alzheimer's disease; nAD-ND, non-AD neurodegeneration; PPD, primary psychiatric disorder.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Table 2: Cognitive outcomes and protein blood biomarker levels by diagnostic category

	EOAD	nAD-ND	PPD	*p* -value[1]
Global Cognition				
Impaired	18 (100%)	15 (83%)	12 (55%)	0.001
Intact	0 (0%)	3 (17%)	10 (45%)	
Memory Function				
Impaired	14 (82%)	5 (26%)	11 (46%)	0.002
Intact	3 (18%)	14 (74%)	13 (54%)	
Executive Function				
Impaired	13 (76%)	15 (75%)	12 (50%)	0.14
Intact	4 (24%)	5 (25%)	12 (50%)	
Biomarker*				***p* -value[2]**
NfL	21 (17, 27)	15 (10, 33)	10 (9, 14)	<0.001
GFAP	198 (154, 262)	77 (53, 112)	56 (48, 101)	<0.001
p-tau181	3.33 (3.04, 3.87)	1.69 (1.22, 1.97)	1.15 (0.94, 1.52)	<0.001

Values indicate number (%) or median (interquartile range).

[1]Fisher's Exact test between diagnostic groups.

[2]Kruskal-Wallis rank sum test.

EOAD, early-onset Alzheimer's disease; nAD-ND, non-AD neurodegeneration; PPD, primary psychiatric disorder.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.30.24309746; this version posted July 1, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Table 3. Selected models in discriminating early-onset Alzheimer's disease from primary psychiatric disorders.

	Variable	OR [95% CI]	AUC	ΔAIC*	*p* -value[1]
1A	*APOE* ε2 Status	0.71 [0.54 - 0.94]	1.000 [1.000 - 1.000]	--	--
	APOE ε4 Status	0.84 [0.70 - 1.02]			
	Global Cognitive Function	0.78 [0.63 - 0.97]			
	NfL	1.10 [1.02 - 1.19]			
	p-tau181	1.48 [1.33 - 1.65]			
2A	*APOE* ε2 Status	0.77 [0.58 - 1.02]	0.982 [0.946 - 1.000]	1.41	3.5E-01
	NfL	1.10 [1.01 - 1.19]			
	p-tau181	1.55 [1.40 - 1.72]			
3A	Global Cognitive Function	0.81 [0.64 - 1.02]	0.975 [0.925 - 1.000]	1.61	3.4E-01
	NfL	1.09 [1.00 - 1.19]			
	p-tau181	1.47 [1.31 - 1.65]			
4A	*APOE* ε2 Status	0.76 [0.57 - 1.01]	0.996 [0.987 - 1.000]	1.78	4.8E-01
	Global Cognitive Function	0.78 [0.62 - 0.99]			
	p-tau181	1.48 [1.32 - 1.66]			
5A	Global Cognitive Function	0.54 [0.39 - 0.73]	0.763 [0.648 - 0.878]	28.34	5.7E-05
6A	*APOE* ε4 Status	1.10 [0.77 - 1.57]	0.542 [0.375 - 0.709]	41.64	8.1E-08
7A	*APOE* ε2 Status	0.81 [0.47 - 1.37]	0.546 [0.439 - 0.652]	41.22	6.6E-17
8A	p-tau181	1.57 [1.40 - 1.75]	0.954 [0.887 - 1.000]	4.26	1.9E-01
9A	NfL	1.15 [0.98 - 1.34]	0.911 [0.814 - 1.000]	38.77	6.8E-02

OR, odds ratio; CI, confidence interval; AUC, area under the curve; ΔAIC, difference in Akaike information criterion with correction for finite sample sizes.

* ΔAIC is difference from Model 1A.

[1]DeLong *p* -value compared to Model 1A.