Machine Learning Driven Profiling of Cerebrospinal Fluid Core Biomarkers in Alzheimer's Disease and Other Neurological Disorders

Giovanni Bellomo[1†], Antonio Indaco[2†], Davide Chiasserini[3†], Emanuela Maderna[2], Federico Paolini Paoletti[4], Lorenzo Gaetani[4], Silvia Paciotti[1], Maya Petricciuolo[1], Fabrizio Tagliavini[2], Giorgio Giaccone[2], Lucilla Parnetti[1,4*‡] and Giuseppe Di Fede[2*‡]

[1] Laboratory of Clinical Neurochemistry, Section of Neurology, Department of Medicine and Surgery, University of Perugia, Perugia, Italy, [2] Neurology 5/Neuropathology Unit, Fondazione IRCCS Istituto Neurologico C. Besta, Milan, Italy, [3] Section of Biochemistry, Department of Medicine and Surgery, University of Perugia, Perugia, Italy, [4] Section of Neurology, Department of Medicine and Surgery, University of Perugia, Perugia, Italy

Amyloid-beta (Aβ) 42/40 ratio, tau phosphorylated at threonine-181 (p-tau), and total-tau (t-tau) are considered core biomarkers for the diagnosis of Alzheimer's disease (AD). The use of fully automated biomarker assays has been shown to reduce the intra- and inter-laboratory variability, which is a critical factor when defining cut-off values. The calculation of cut-off values is often influenced by the composition of AD and control groups. Indeed, the clinically defined AD group may include patients affected by other forms of dementia, while the control group is often very heterogeneous due to the inclusion of subjects diagnosed with other neurological diseases (OND). In this context, unsupervised machine learning approaches may overcome these issues providing unbiased cut-off values and data-driven patient stratification according to the sole distribution of biomarkers. In this work, we took advantage of the reproducibility of automated determination of the CSF core AD biomarkers to compare two large cohorts of patients diagnosed with different neurological disorders and enrolled in two centers with established expertise in AD biomarkers. We applied an unsupervised Gaussian mixture model clustering algorithm and found that our large series of patients could be classified in six clusters according to their CSF biomarker profile, some presenting a typical AD-like profile and some a non-AD profile. By considering the frequencies of clinically defined OND and AD subjects in clusters, we subsequently computed cluster-based cut-off values for Aβ42/Aβ40, p-tau, and t-tau. This approach promises to be useful for large-scale biomarker studies aimed at providing efficient biochemical phenotyping of neurological diseases.

Keywords: Alzheimer's disease, biomarkers, dementia, cerebrospinal fluid, amyloid-beta, tau, machine learning, clustering analysis

INTRODUCTION

Alzheimer's disease (AD) is the most common neurodegenerative disorder evolving to dementia (Goedert and Spillantini, 2006). Increasing knowledge of the molecular mechanisms underlying the pathogenesis of AD has progressively improved the protocols employed for its diagnosis in clinical practice (Jack et al., 2018). However, AD is also recognized as a heterogeneous disorder that may occur under several distinct phenotypes which can mimic other forms of dementias and other neurodegenerative conditions (Di Fede et al., 2018). Such phenotypic heterogeneity sometimes makes the differential diagnosis between AD and other similar neurological diseases problematic (Sawyer et al., 2017; de Souza et al., 2019; Villain and Dubois, 2019). Cerebrospinal fluid (CSF) core biomarkers for AD – i.e., amyloid-beta (Aβ) 42/40 ratio, tau phosphorylated at threonine-181 (p-tau), and total-tau (t-tau) – are largely used in clinical settings, research, and drug trials (Paterson et al., 2018; Gaetani et al., 2020). However, their clinical utility to differentiate AD from non-AD neurodegenerative dementias, such as dementia with Lewy bodies (DLB) or frontotemporal dementia (FTD), is less established (Bartlett et al., 2012; Molinuevo et al., 2014). For a long time, manual enzyme-linked immunosorbent assay (ELISA) has been widely employed as the reference method for the analysis of the CSF AD biomarkers. However, its broad-scale use is critically hampered by the assay variability, which influences the measurement of the analytes and the interpretation of the outcome data, especially in the routine clinical context (Mattsson et al., 2012; Le Bastard et al., 2015). Due to these concerns, ELISA was recently replaced in worldwide laboratories by fully automated assays – such as chemiluminescence enzyme immunoassay (CLEIA) – which offer grounds to cut sample manipulation steps and to reduce the intra- and inter-laboratory variability for CSF biomarker measurement (Kollhoff et al., 2018). Nevertheless, there is still a need for harmonization of CSF biomarker assays across centers involved in AD diagnostics (Mattsson-Carlgren et al., 2020). For instance, the lack of established universal biomarker cut-offs makes the calculation of internal reference values mandatory for each laboratory both for clinical and research purposes. This calculation is often critically influenced by the choice and composition of AD and control groups; the clinically defined AD group may include patients affected by other forms of dementia (e.g., FTD and DLB) due to misdiagnosis, while the control group is often very heterogeneous due to the inclusion of subjects diagnosed with other neurological diseases (OND) who underwent lumbar puncture (LP) for diagnostic purposes. On the one hand, OND may better represent the real cases afferent to neurology clinics compared to healthy subjects. However, the heterogeneity of the inclusion criteria adopted in each center for the definition of OND controls represents a source of variability for the calculation of biomarkers cut-off values. In addition, the absence of standardized methodological and statistical approaches represents one of the most critical issues to study the distribution of CSF core AD biomarkers in different subgroups of patients and to validate in larger cohorts the cut-off values able to discriminate between AD and other AD-mimicking disorders (Simrén et al., 2020). In this context, unsupervised machine learning approaches may overcome both misdiagnosis and the lack of standardization of inclusion criteria providing unbiased cut-off values and data-driven patient stratification according to the sole distribution of biomarkers. In this work, we took advantage of the reproducibility of automated determination of the CSF core AD biomarkers to compare two large cohorts of patients diagnosed with different neurological disorders and enrolled in two centers with established expertise in AD biomarkers. We applied an unsupervised Gaussian mixture model (GMM) clustering algorithm and found that our large series of patients could be classified in six clusters according to their CSF biomarker profile, some presenting a typical AD-like profile and some a typical non-AD profile. By considering the frequencies of clinically defined OND and AD subjects in clusters, we subsequently computed cluster-based cut-off values for Aβ42/Aβ40, p-tau, and t-tau.

MATERIALS AND METHODS

Patients

A total of 616 prospectively collected CSF samples from patients referring to the Neurology Clinic, University of Perugia (cohort 1), and from the Carlo Besta Neurological Institute, Milan (cohort 2), were used in this study. All patients underwent a standardized assessment including medical history, physical and neurological examination, laboratory tests, neuropsychological evaluation, and brain imaging (computed tomography or magnetic resonance imaging, MRI). 18Fluoro-2-deoxyglucose positron emission tomography (FDG-PET), dopamine transporter single photon emission computed tomography (DaT-Scan), and electroencephalogram were also performed in selected cases, according to clinical suspicion. According to the purposes of our investigation, clinical diagnoses were made by consensus in a multidisciplinary meeting of neurologists with a deep expertise in the field of neurodegenerative diseases, without knowledge of CSF results. Therefore, we did not consider the most updated criteria for AD diagnosis, based on A/T/(N) classification (Jack et al., 2018), but rather we defined patients as affected by AD or other neurological disorders only according to the available clinical criteria, as follows. Patients with neurodegenerative disorders included 257 patients with probable AD (Dubois et al., 2007) both at dementia and prodromal (MCI) stages, 50 frontotemporal dementia (FTD) patients (Faber, 1999), 56 patients with Parkinson's disease (PD) (Postuma et al., 2015), 7 PD with dementia (PDD) patients (Emre et al., 2007), 21 patients with dementia with Lewy bodies (DLB) (McKeith et al., 2017), 58 patients with atypical parkinsonism or parkinsonism of different etiology (Gilman et al., 2008; Armstrong et al., 2013; Höglinger et al., 2017; Rektor et al., 2018), 1 patients with amyotrophic lateral sclerosis (ALS) (Traynor et al., 2000), 8 patients with Creutzfeldt-Jakob disease (CJD) (Manix et al., 2015), 27 patients with normal pressure hydrocephalus (NPH) (Relkin et al., 2005), and 2 patients with genetically confirmed degenerative spinocerebellar ataxia (SCA). Patients were classified as having subjective cognitive decline (SCD) if they complained cognitive

deficits but neuropsychological evaluation was normal or showed subtle deficits not fulfilling criteria for mild cognitive impairment (MCI) (8 patients). Patients with stable MCI (sMCI) showed unchanged neuropsychological results after 1-year follow-up (20 patients). Other diagnostic groups included vascular dementia (5 patients) (Román et al., 1993), cerebral amyloid angiopathy (CAA) (Smith and Greenberg, 2003) (3 patients), autoimmune encephalitis (8 patients) (Graus et al., 2016), encephalopathies of different etiology (2 patients), relapsing-remitting multiple sclerosis (MS) (Thompson et al., 2018) (6 patients), and cognitively impaired late-onset epilepsy (1 patient) (Scheffer et al., 2017). Patients categorized as having cerebrovascular diseases (CVD) showed significant brain small vessel disease at MRI (i.e., white matter changes, microbleeds, and lacunar infarcts) without fulfilling diagnostic criteria for VaD and CAA (3 patients). Dementia of unknown origin (uDEM) was defined for those subjects in which brain imaging including both MRI and nuclear imaging excluded vascular and neurodegenerative origins (3 patients). Cognitively unimpaired patients referring to our centers for psychiatric disorders or neurological conditions like headaches, seizures, mononeuropathies, and polyneuropathies, in which brain imaging did not reveal gross abnormalities nor underlying neurodegenerative diseases, were classified as control subjects with other neurological diseases (OND) (71 patients).

Samples Collection and Analysis

Lumbar puncture was performed according to international guidelines (Teunissen et al., 2009); 10–12 mL of CSF was collected in sterile polypropylene tubes (Sarstedt® tubes, code: 62.610.210) and centrifuged for 10 min (2000 × g), at room temperature. Aliquots of 0.5 mL were frozen at −80°C in polypropylene tubes (Sarstedt® tubes, code: 72.730.007). CSF samples were analyzed on the fully automated chemiluminescent platform Lumipulse G600-II (Fujirebio Inc) for β-amyloid 1-42 (Aβ42), β-amyloid 1-40 (Aβ40), t-tau and p-tau (Thr181) levels. For cohort 1, all the CSF samples were analyzed directly in their 0.5 mL storage tubes, while for cohort 2 samples were analyzed by transferring them in Hitachi® polystyrene sample caps (code: 80351). Throughout this work, Aβ42/Aβ40 ratio was used since it represents a more robust marker of amyloidosis with respect to the sole Aβ42 (Biscetti et al., 2019). Moreover, the use of Aβ42/Aβ40 can also partially compensate the above-mentioned methodological difference between the two centers, since the Aβ absorption due to tube transfer (Toombs et al., 2014) is thought to act similarly for the 1-40 and 1-42 Aβ isoforms (Lewczuk et al., 2006).

Cohorts Merging

A preliminary experiment was carried out to assess the inter-center variability of the CSF biomarkers and the possibility to merge the two cohorts. A total of 40 CSF samples (20 from each center) were measured with Lumipulse-G automated platforms in the two laboratories by using kits originating from the same batches. The composition of this validation cohort is reported in **Supplementary Table 1**. The concordance between the measurements was assessed by correlation analysis.

Statistical Methods

The data analysis was performed by using R software v 3.6 (R Core Team, 2013).

Correlation Analysis

Because of the known non-optimal normality of biomarker data (Bellomo et al., 2020a), Passing Bablok regressions (Passing and Bablok, 1983) were preferred to parametric least squares regressions. Confidence intervals (CI) for the fitted parameters were calculated with the bootstrap method (Carpenter and Bithell, 2000). Principal component analysis (PCA) was then applied to the whole dataset to graphically show the absence of a significant separation among samples belonging to different cohorts. The R-package *mcr* was used for these calculations [mcr package | R Documentation (2021)].

Cluster Analysis

The GMM algorithm (Figueiredo and Jain, 2002) of the machine learning Python package Scikit-learn v 0.23.2 (Pedregosa et al., 2011) was used for the cluster analysis. Biomarker values were all z-scored prior to the analysis. The optimal number of clusters was chosen minimizing the Bayesian information criterion (BIC) function (Schwarz, 1978). After the clustering, all the samples biomarker values were back-transformed into the original dimensions. Median biomarker values together with the 95% data range were calculated for each cluster. The prevalences of diagnostic categories in each cluster were represented in percentages in a heatmap. Both diagnostic groups and clusters were grouped according to a hierarchical clustering (Rokach and Maimon, 2005; Gu, 2021). Euclidean distance and average linkage were used as parameters for clustering.

Calculation of Cut-Off Values

Cut-off values were calculated for OND vs. AD and among clusters by maximizing Youden's index with the p-ROC package in R v3.6 (Robin et al., 2011). Cut-off CI were calculated by using 2000 bootstrap replicates.

Calculation of Cut-Off Values on Age- and Gender-Matched Subsets

Age histogram matching was performed by random exclusion of subjects within bins of 5 years width. Exclusion of samples according to gender was subsequently conducted until p-values > 0.25 were obtained by logistic regression (Dobson and Barnett, 2018; glm function | R Documentation, 2021) both for age and gender. Recalculation of cut-off values for the age- and gender-matched subsets was performed as described in section "Calculation of Cut-Off Values."
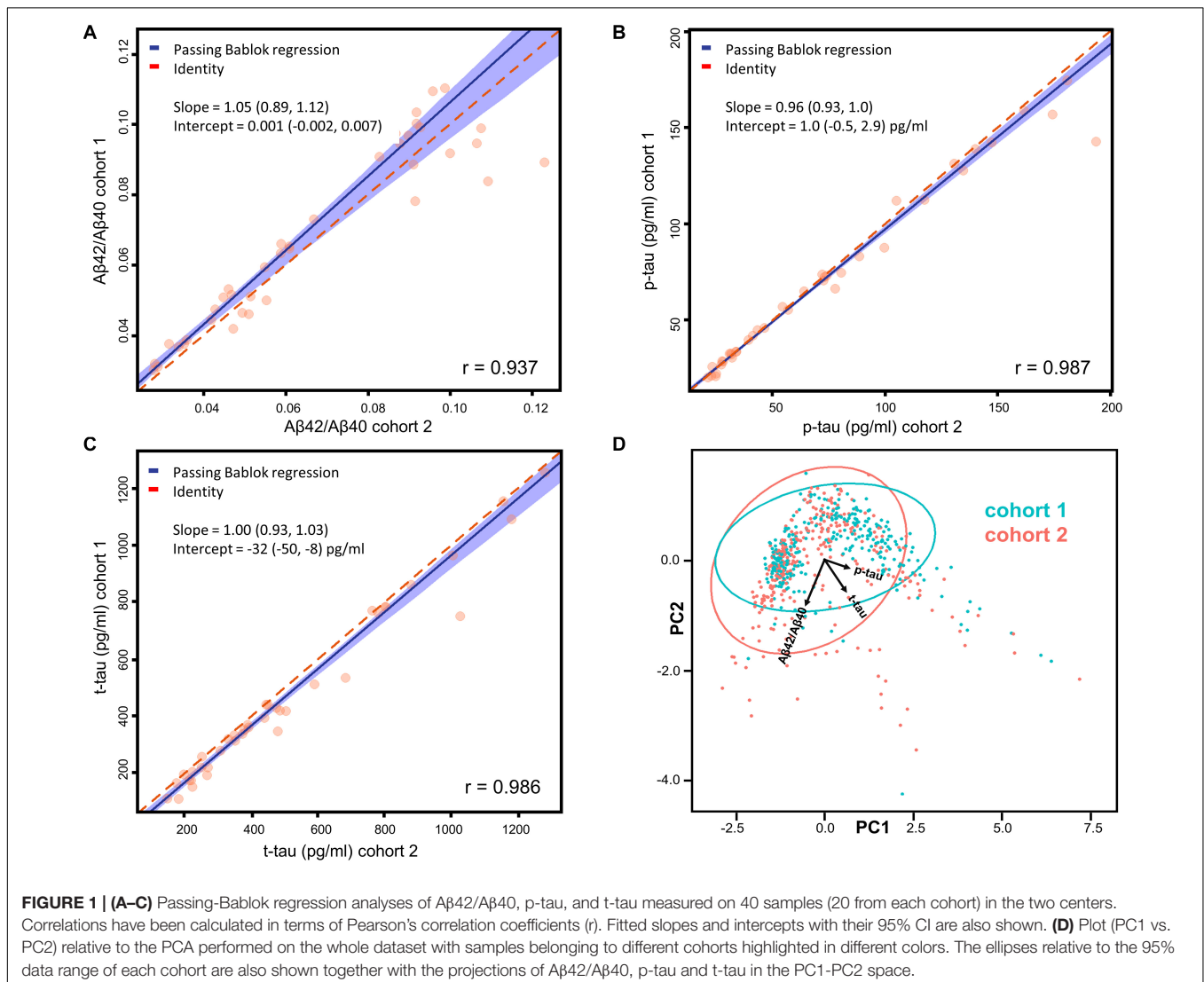
RESULTS

Patients Demographical Data

A total of 616 patients whose CSF samples were tested for AD biomarker by Lumipulse-G, were included in the study regardless of age and clinical diagnoses. Among them, 257 were clinically diagnosed as AD and 71 cognitively unimpaired subjects affected

by minor neurological non-neurodegenerative disorders were classified as OND. The whole cohort originated by merging two sub-cohorts: 303 subjects referred to the biobank of the Neurology clinic of the University of Perugia (cohort 1), while 313 referred to the biobank of the Carlo Besta Neurological Institute of Milan (cohort 2). The demographical details of the subjects included are reported in **Supplementary Table 3**.

Merging of the Two Cohorts

Assessment of CSF AD biomarkers with Lumipulse-G CLEIA technology showed very low inter-center variability in external quality control programs (Leitão et al., 2019; Paciotti et al., 2019). However, we performed a small scale inter-center variability study, measuring a total of 40 samples across the two centers involved, to assess the possibility to merge the cohorts from Perugia (cohort 1, 303 subjects) and Milan (cohort 2, 313 subjects). Patients' diagnoses, mean biomarker values and inter-assay/inter-laboratory coefficients of variations (CV) are reported for all of these samples in **Supplementary Table 1**. As expected,

the mean inter-assay CV of Aβ42/Aβ40 (7%) was lower compared to the ones of Aβ42 (12%) and Aβ40 (9%). Mean inter-assay CV of p-tau (4%) and t-tau (9%) were also relatively low. Correlation and Passing-Bablok linear regression analyses showed a good agreement between the measurements performed in the two centers (**Figures 1A–C**). A Pearson's correlation coefficient above 0.9 was found for each of the tested biomarkers and all the measured slopes were equal to 1 within their 95% CI. Intercepts for Aβ42/Aβ40 and p-tau were null within their 95% CI, whereas a non-null negative intercept was obtained for t-tau, although being small compared to usual nominal t-tau values. The results of Passing-Bablok linear regression analysis for Aβ42 and Aβ40 are shown in **Supplementary Table 2**. As expected (Lewczuk et al., 2006), we found a greater deviance from identity between the measurements of these two peptides performed in the two laboratories with respect to their ratio. We subsequently analyzed all the 616 samples included in the study by means of PCA. Projection of the data into the principal components space (**Figure 1D**)

FIGURE 1 | (A–C) Passing-Bablok regression analyses of Aβ42/Aβ40, p-tau, and t-tau measured on 40 samples (20 from each cohort) in the two centers. Correlations have been calculated in terms of Pearson's correlation coefficients (r). Fitted slopes and intercepts with their 95% CI are also shown. **(D)** Plot (PC1 vs. PC2) relative to the PCA performed on the whole dataset with samples belonging to different cohorts highlighted in different colors. The ellipses relative to the 95% data range of each cohort are also shown together with the projections of Aβ42/Aβ40, p-tau and t-tau in the PC1-PC2 space.

showed that the measurement from the two centers did not show any significant grouping related to the site of analysis, as shown by the ellipses representative of the 95% data range of the two cohorts.
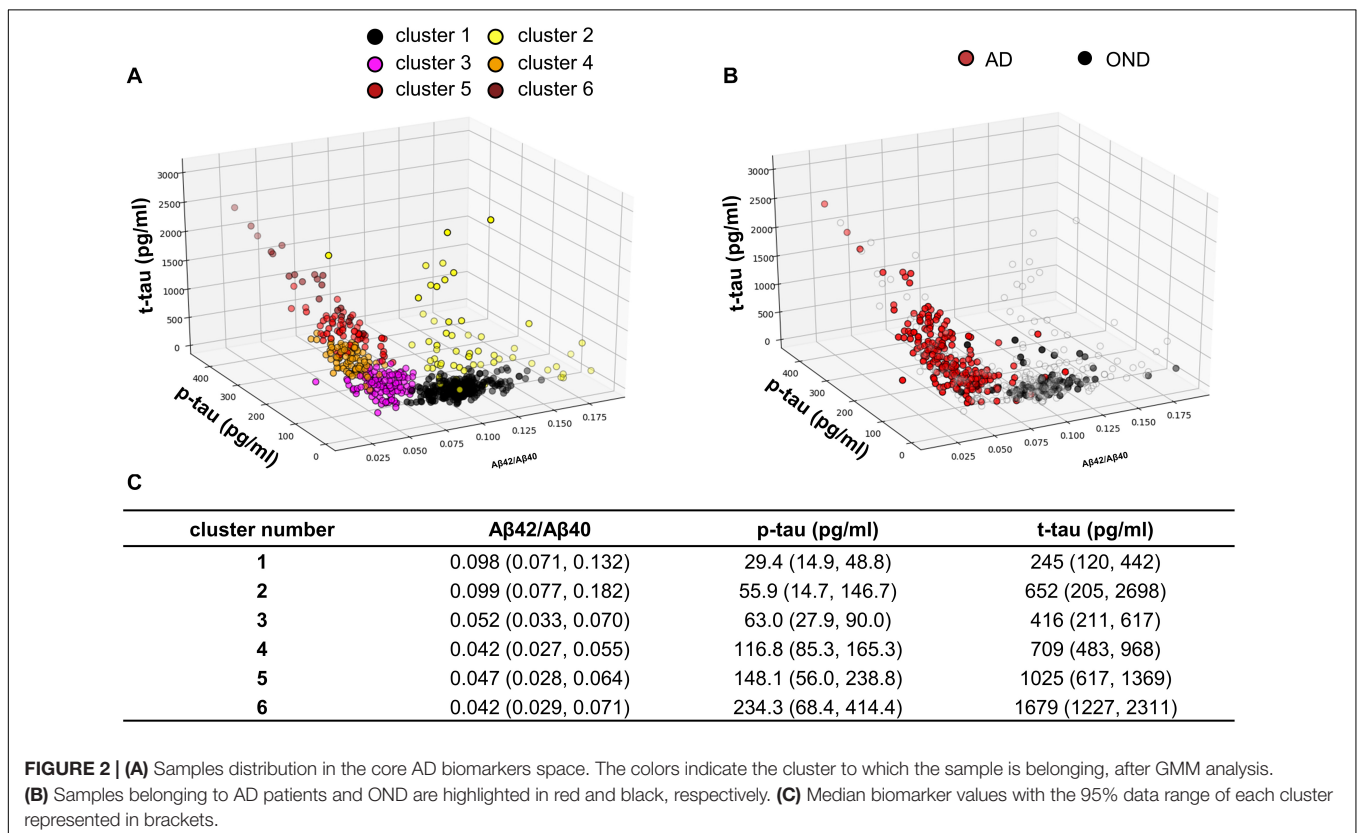
Both the regression and the PCA analysis showed that the inter-center variability was negligible for Aβ42/Aβ40, p-tau and t-tau. Cut-off values for Aβ42, Aβ42/Aβ40, p-tau and t-tau for cohort 1, cohort 2 and for the two cohorts merged are also reported in the **Supplementary Table 4**. The calculated cut-off values did not significantly differ within their 95% CI between cohort 1 and 2. Thus, we proceeded to merge the two cohorts for subsequent analyses, without applying any correction factor for Aβ42/Aβ40, p-tau and t-tau.

Clustering

For the unsupervised cluster analysis, we included all the subjects who consecutively underwent LP in the two centers and whose CSF was assayed for Aβ42/Aβ40, p-tau and t-tau with Lumipulse-G ($N = 616$). Considering the results of the PCA plotted in **Figure 1D**, we decided to apply GMM as clustering algorithm (Pedregosa et al., 2011). Other clustering algorithms such as K-means are known for not providing good fittings for anisotropic data. In order to standardize the dimensions of the three biomarkers considered, biomarker values were substituted with the corresponding Z-scores before the analysis and then back-transformed for plotting and data interpretation. The optimal number of clusters was decided according to BIC (Schwarz, 1978). A plot relative to the BIC

function is shown in **Supplementary Figure 1**. Accordingly, the optimal number of clusters turned out to be 6. The results of the unsupervised clustering analysis are shown in **Figure 2A**, GMM centroids and covariance matrices are reported in **Supplementary Table 5**. As it can be seen by comparing the 3D scatter plots in **Figures 2A,B**, most AD and OND samples were assigned to different clusters. In particular, most OND were included in cluster 1 while most AD subjects (95%) were comprised in clusters 3, 4, 5, and 6. Considering the biomarker 95% data ranges of each cluster (**Figure 2C**), clusters 3–6 corresponded to low Aβ42/Aβ40 values and high values of p-tau and t-tau, which, according to the A/T/(N) criteria (Jack et al., 2018), correspond to the presence of amyloidosis, tauopathy and tau-related neurodegeneration, respectively. These four clusters mainly differed in p-tau and t-tau values (**Figure 2C**). Among AD patients, for all the AD clusters, the prevalences of demented subjects ($N = 253$) did not significantly differ from the prevalences of subjects in the MCI phase ($N = 53$) by applying Fisher's exact test for count data. Cluster 1 was instead characterized by higher values of Aβ42/Aβ40 and small values of p-tau and t-tau. Biomarker values in cluster 2 were instead highly variable with respect to the other clusters (wide 95% data ranges for all the three biomarkers). This cluster was characterized by smaller Aβ42/Aβ40 median values with respect to clusters 3–6 and higher p-tau and t-tau median values compared to cluster 1.

The percentages of AD patients, OND and other sufficiently represented ($N > 6$) clinical conditions in each cluster are

cluster number	Aβ42/Aβ40	p-tau (pg/ml)	t-tau (pg/ml)
1	0.098 (0.071, 0.132)	29.4 (14.9, 48.8)	245 (120, 442)
2	0.099 (0.077, 0.182)	55.9 (14.7, 146.7)	652 (205, 2698)
3	0.052 (0.033, 0.070)	63.0 (27.9, 90.0)	416 (211, 617)
4	0.042 (0.027, 0.055)	116.8 (85.3, 165.3)	709 (483, 968)
5	0.047 (0.028, 0.064)	148.1 (56.0, 238.8)	1025 (617, 1369)
6	0.042 (0.029, 0.071)	234.3 (68.4, 414.4)	1679 (1227, 2311)

FIGURE 2 | (A) Samples distribution in the core AD biomarkers space. The colors indicate the cluster to which the sample is belonging, after GMM analysis. **(B)** Samples belonging to AD patients and OND are highlighted in red and black, respectively. **(C)** Median biomarker values with the 95% data range of each cluster represented in brackets.

reported in **Figure 3**. Interestingly, not only AD but also PDD, DLB, and CBD had a relevant presence (>20%) in clusters 3–6. Conversely, MS, PSP, sMCI, and synucleinopathies without dementia (PD and MSA) majorly colocalize with OND in cluster 1. The composition of cluster 2 was instead highly variable, consisting in the totality of CJD subjects and relevant percentages (≥20%) of NPH, ENC, FTD, and VAD. Considering the distribution of clinically defined OND and AD subjects in the clusters together with the clusters data ranges with respect to the calculated biomarkers cut-off values, we decided to indicate cluster 1 as the "control" cluster and clusters 3–6 as "AD-clusters". Grouping of the clusters and of the clinical diagnoses showed that cluster 1, the control cluster, was the most distant from the others, followed by cluster 2 which was included in a separated branch with respect to cluster 3–6. The clinical diagnosis showed also a peculiar grouping, with AD, PDD, and DLB in the same branch of the dendrogram, while the other conditions (mostly included in clusters 1 and 2) in a second large branch. CJD clustered separately from all the other conditions being characterized by high $A\beta42/A\beta40$ ratios and high t-tau values.

According to the prevalence of AD and OND in clusters 1–6, we calculated cut-off values for each biomarker considering AD vs. OND clinical diagnoses, control cluster vs. AD clusters and cluster 2 vs. AD clusters. The results are shown in **Table 1**.

We noticed that the cut-offs of the comparison control cluster vs. AD clusters were relatively similar to what obtained using the clinical diagnosis grouping (OND vs. AD), being mostly comprised within the 95% CI. Cut-off values for control cluster vs. AD clusters remained unchanged also by considering an age- and gender-matched subsets of the population (**Supplementary Figure 2**).
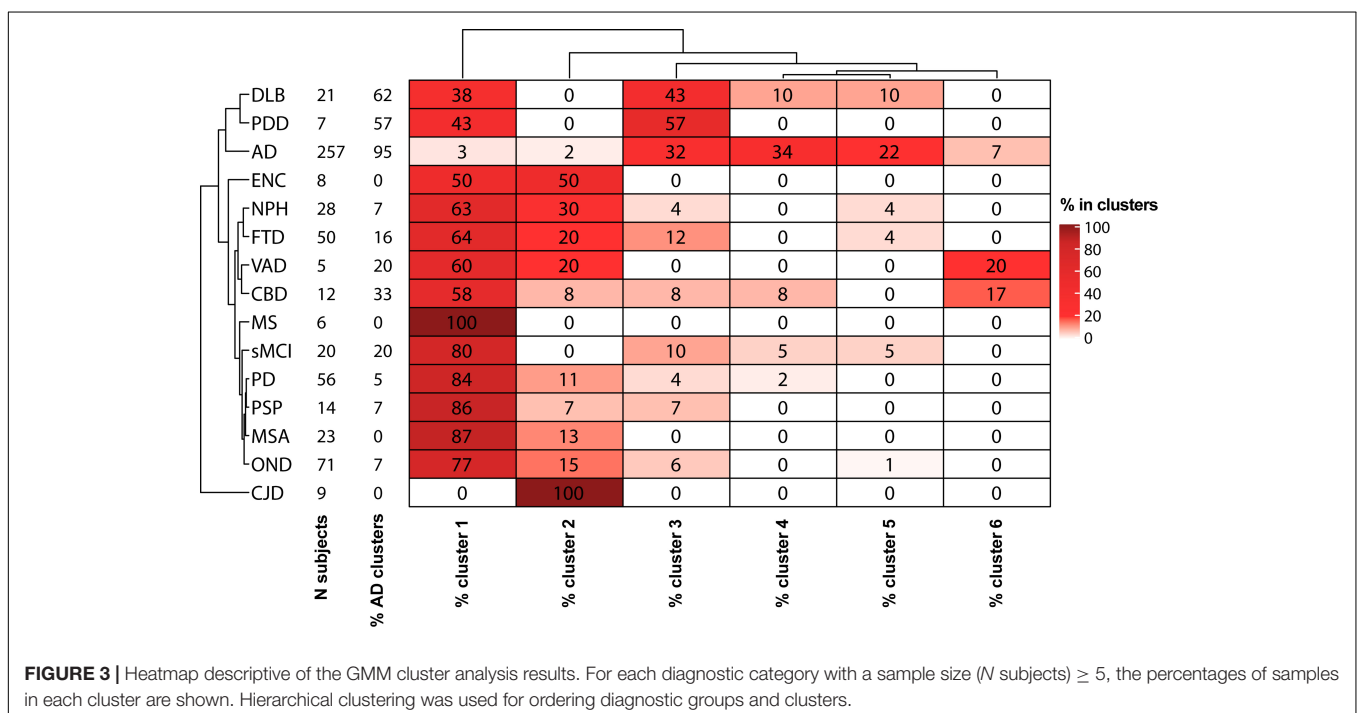
TABLE 1 | Cut-off values for the three core AD biomarkers with their 95% CI were calculated by maximizing the Youden's index for AD vs. OND, between samples belonging to the AD clusters (cluster 3, 4, 5, and 6) and "control" cluster (cluster 1) and between samples belonging to the AD clusters and cluster 2.

	Aβ42/Aβ40	p-tau (pg/ml)	t-tau (pg/ml)
OND vs. AD	0.073 (0.063, 0.079)	53.5 (47.2, 57.5)	371 (332, 393)
Control cluster vs. AD clusters	0.072 (0.070, 0.074)	50.0 (46.2, 52.3)	392 (359, 396)
Cluster 2 vs. AD clusters	0.073 (0.072, 0.078)	71.6 (50.6, 82.8)	1403 (485, 1999)

On the other hand, the cluster 2 vs. AD clusters comparison showed significant differences in the absolute values of cut-offs for t-tau and p-tau, while the $A\beta42/A\beta40$ ratio did not change significantly.

DISCUSSION

In the last decade, CSF $A\beta42/A\beta40$ ratio, p-tau and t-tau emerged as reliable markers of brain amyloidosis, tauopathy and tau-related neurodegeneration. The introduction of these markers into clinical practice has substantially helped the neurologist to change the definition of AD from a syndromal to a molecular construct (Jack et al., 2018). In particular, considering the well-established A/T/(N) system (Jack et al., 2018), AD is now defined by the presence of both brain amyloidosis (A+) and tauopathy (T+), with neurodegeneration (N+) being a non-necessary condition. The recent advent of automated platforms for core AD biomarker assessment in CSF, has been of substantial help in limiting both intra and inter-assay variability with

FIGURE 3 | Heatmap descriptive of the GMM cluster analysis results. For each diagnostic category with a sample size (N subjects) ≥ 5, the percentages of samples in each cluster are shown. Hierarchical clustering was used for ordering diagnostic groups and clusters.

respect to manual ELISA (Le Bastard et al., 2015). In this work, we used the reproducibility of automated CSF core AD biomarker determination to compare two large cohorts of patients diagnosed with various neurological disorders and enrolled in two centers with proven expertise in AD biomarkers. After a small round-robin validation step on 40 CSF samples, we were able to confirm the good reproducibility of the determinations between the two centers using the automated procedure. In order to overcome the diagnostic heterogeneity of both AD and control groups, we applied unsupervised GMMs to cluster the patients (n = 616) according to their CSF biomarker profile and investigate the degree of overlap between the clinical diagnosis and the data-driven classification of the subjects. The data spontaneously grouped in six clusters, 4 of these (clusters 3–6) contained the 95% of clinically defined AD patients (**Figure 3**), characterized by low Aβ42/Aβ40 values and high values of p-tau and t-tau (**Figure 2C**). Interestingly, high percentages (>50%) of synucleinopathies with dementia (DLB and PDD) and CBD (33%) fell in these clusters. This fact is not surprising, since the presence of brain amyloidosis and tauopathy is a feature of both DLB and PDD (Irwin et al., 2013; Irwin and Hurtig, 2018; Bellomo et al., 2020b). Cluster 1 instead contained the majority of OND subjects (77%), and PD (84%), MSA (87%), PSP (86%), and MS (100%) patients, suggesting that AD pathology is not frequent in these conditions. This result, together with the high concordance between cut-off values for the OND vs. AD and control cluster vs. AD clusters comparisons, suggest that these conditions may be treated as controls with respect to core AD biomarkers. The inclusion of NPH in control groups should be instead avoided since neurodegeneration and dilution effects may significantly alter the concentration of t-tau and Aβ peptides (Graff-Radford, 2014), in a way that is not fully compensated by computing their ratio. By considering the clusters with highest and lowest frequencies of AD and OND subjects, we were able to compute cluster-defined cut-off values. In our work, the cut-off values calculated by the clustering method were concordant to the ones calculated by relying on clinical diagnoses for AD vs. OND, but this approach may be of more substantial help while facing low numbers of well clinically characterized OND and/or AD subjects. The cut-off values calculated for AD clusters vs. control cluster were substantially unchanged considering age- and gender-matched subsets of the clusters (**Supplementary Figure 2**), thus reinforcing the reliability of the approach used. Because of these advantages, unsupervised and partially supervised machine-learning algorithms (like the one we applied) have recently started to be applied in neurodegenerative diseases diagnostics (Skillbäck et al., 2015; Racine et al., 2016; Toschi et al., 2019). Moreover, these approaches represent a best choice while dealing with a large number of biomarkers or candidate biomarkers (Solorio-Fernández et al., 2020), e.g., in omics studies (Lopez et al., 2018). The inclusion of a wide panel of CSF markers, possibly linked to different biological pathways, may help in differentiating synucleinopathies with dementia and FTD from AD and synucleinopathies without dementia and PSP from controls. As a limitation of our study, we must report the small sample size of some diagnostic categories (e.g., MS, VAD and PDD). Thus, the frequencies of these categories in clusters may

potentially be biased. Another limitation is the lack of amyloid PET, which is of substantial help in identifying brain amyloidosis and well correlates with brain amyloidosis markers such as CSF Aβ42 and Aβ42/Aβ40 (Alcolea et al., 2019; Leitão et al., 2019). However, this partially supervised approach is insensitive to the presence of hidden interfering pathologies in control subjects and to the presence of few AD misdiagnosis, which may occur when the diagnosis is made prevalently by the examination of clinical features. Since the definition of AD and control clusters depended only on the prevalences of clinically defined AD and OND subjects within each cluster, we expect that this approach may provide reliable results as long as the diagnosis/exclusion of AD is correct in the majority of the cases. As an example, considering the biomarker distributions presented in this study, to misclassify the control cluster it would have required at least 28 out of 71 OND subjects (40%) misdiagnosed for AD.

Overall, our findings suggest that automated assays are amenable for large-scale biomarker studies across centers. Furthermore, the use of unsupervised (or partially supervised) machine learning approaches may help the biochemical phenotyping of neurological disorders, being also a robust option for the definition of cut-off values. The implementation of such approaches in biomarker research could substantially improve the development of adequate diagnostic protocols and increase the quality of diagnostic tools for complex and heterogeneous disorders presenting with overlapping clinical syndromes, like dementias.

DATA AVAILABILITY STATEMENT

Raw data were generated at University of Perugia and Carlo Besta Neurological Institute. Derived data supporting the findings of this study are available from the corresponding authors (LP and GDF) on request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee from Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, and University of Perugia. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

GB analyzed the data and wrote the first draft. AI analyzed the CSF samples at Carlo Besta Neurological Institute and revised the manuscript. DC conceptualized the data analysis and contributed to writing the first draft of the manuscript. EM contributed in analyzing the CSF samples at Carlo Besta Neurological Institute. FPP contributed to the sample collection and assisted with data management of samples. LG contributed to the sample selection and collection. SP and MP analyzed the samples at the University of Perugia and revised the manuscript. FT and GG contributed in the conceptualization of the study

and revised the manuscript. GDF and LP designed the study and revised the final version of the manuscript. All the authors read and approved the final version of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2021.647783/full#supplementary-material

REFERENCES

Alcolea, D., Pegueroles, J., Muñoz, L., Camacho, V., López-Mora, D., Fernández-León, A., et al. (2019). Agreement of amyloid PET and CSF biomarkers for Alzheimer's disease on Lumipulse. *Ann. Clin. Transl. Neurol.* 6, 1815–1824. doi: 10.1002/acn3.50873

Armstrong, M. J., Litvan, I., Lang, A. E., Bak, T. H., Bhatia, K. P., Borroni, B., et al. (2013). Criteria for the diagnosis of corticobasal degeneration. *Neurology* 80, 496–503. doi: 10.1212/WNL.0b013e31827f0fd1

Bartlett, J. W., Frost, C., Mattsson, N., Skillbäck, T., Blennow, K., Zetterberg, H., et al. (2012). Determining cut-points for Alzheimer's disease biomarkers: statistical issues, methods and challenges. *Biomark. Med.* 6, 391–400. doi: 10.2217/bmm.12.49

Bellomo, G., Cataldi, S., Paciotti, S., Paolini Paoletti, F., Chiasserini, D., and Parnetti, L. (2020a). Measurement of CSF core Alzheimer disease biomarkers for routine clinical diagnosis: do fresh vs frozen samples differ? *Alzheimers Res. Ther.* 12:121. doi: 10.1186/s13195-020-00689-0

Bellomo, G., Paolini Paoletti, F., Chipi, E., Petricciuolo, M., Simoni, S., Tambasco, N., et al. (2020b). A/T/(N) Profile in Cerebrospinal Fluid of Parkinson's Disease with/without Cognitive Impairment and Dementia with Lewy Bodies. *Diagnostics* 10:1015. doi: 10.3390/diagnostics10121015

Biscetti, L., Salvadori, N., Farotti, L., Cataldi, S., Eusebi, P., Paciotti, S., et al. (2019). The added value of Aβ42/Aβ40 in the CSF signature for routine diagnostics of Alzheimer's disease. *Clin. Chim. Acta* 494, 71–73. doi: 10.1016/j.cca.2019.03.001

Carpenter, J., and Bithell, J. (2000). Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Statist. Med.* 19, 1141–1164. doi: 10.1002/(SICI)1097-0258(20000515)19:9<1141::AID-SIM479>3.0.CO;2-F

de Souza, L. C., Mariano, L. I., de Moraes, R. F., and Caramelli, P. (2019). Behavioral variant of frontotemporal dementia or frontal variant of Alzheimer's disease? A case study. *Dement. Neuropsychol.* 13, 356–360. doi: 10.1590/1980-57642018dn13-030015

Di Fede, G., Catania, M., Maderna, E., Ghidoni, R., Benussi, L., Tonoli, E., et al. (2018). Molecular subtypes of Alzheimer's disease. *Sci. Rep.* 8:3269. doi: 10.1038/s41598-018-21641-1

Dobson, A. J., and Barnett, A. G. (2018). *An Introduction to Generalized Linear Models*. Florida, FL: CRC Press.

Dubois, B., Feldman, H. H., Jacova, C., Dekosky, S. T., Barberger-Gateau, P., Cummings, J., et al. (2007). Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol.* 6, 734–746. doi: 10.1016/S1474-4422(07)70178-3

Emre, M., Aarsland, D., Brown, R., Burn, D. J., Duyckaerts, C., Mizuno, Y., et al. (2007). Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Mov. Disord.* 22, 1689–1707. doi: 10.1002/mds.21507

Faber, R. (1999). Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology* 53:1159. doi: 10.1212/wnl.53.5.1158-b

Figueiredo, M. A. T., and Jain, A. K. (2002). Unsupervised learning of finite mixture models. *IEEE Transact. Pattern Anal. Machine Intellig.* 24, 381–396. doi: 10.1109/34.990138

Gaetani, L., Paoletti, F. P., Bellomo, G., Mancini, A., Simoni, S., Di Filippo, M., et al. (2020). CSF and Blood Biomarkers in Neuroinflammatory and Neurodegenerative Diseases: Implications for Treatment. *Trends Pharmacol. Sci.* 41, 1023–1037. doi: 10.1016/j.tips.2020.09.011

Gilman, S., Wenning, G. K., Low, P. A., Brooks, D. J., Mathias, C. J., Trojanowski, J. Q., et al. (2008). Second consensus statement on the diagnosis of multiple system atrophy. *Neurology* 71, 670–676. doi: 10.1212/01.wnl.0000324625.00404.15

glm function | R Documentation (2021). *Fitting Generalized Linear Models*. Available online at: https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/glm (accessed March 3, 2021)

Goedert, M., and Spillantini, M. G. (2006). A century of Alzheimer's disease. *Science* 314, 777–781. doi: 10.1126/science.1132814

Graff-Radford, N. R. (2014). Alzheimer CSF biomarkers may be misleading in normal-pressure hydrocephalus. *Neurology* 83, 1573–1575. doi: 10.1212/WNL.0000000000000916

Graus, F., Titulaer, M. J., Balu, R., Benseler, S., Bien, C. G., Cellucci, T., et al. (2016). A clinical approach to diagnosis of autoimmune encephalitis. *Lancet Neurol.* 15, 391–404. doi: 10.1016/S1474-4422(15)00401-9

Gu, Z. (2021). *ComplexHeatmap: Make Complex Heatmaps. Bioconductor version: Release (3.12)*. doi: 10.18129/B9.bioc.ComplexHeatmap

Höglinger, G. U., Respondek, G., Stamelou, M., Kurz, C., Josephs, K. A., Lang, A. E., et al. (2017). Clinical diagnosis of progressive supranuclear palsy: The movement disorder society criteria. *Mov. Disord.* 32, 853–864. doi: 10.1002/mds.26987

Irwin, D. J., and Hurtig, H. I. (2018). The Contribution of Tau, Amyloid-Beta and Alpha-Synuclein Pathology to Dementia in Lewy Body Disorders. *J. Alzheimers Dis. Parkinsonism* 8:1000444. doi: 10.4172/2161-0460.1000444

Irwin, D. J., Lee, V. M.-Y., and Trojanowski, J. Q. (2013). Parkinson's disease dementia: convergence of α-synuclein, tau and amyloid-β pathologies. *Nat. Rev. Neurosci.* 14, 626–636. doi: 10.1038/nrn3549

Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., et al. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement.* 14, 535–562. doi: 10.1016/j.jalz.2018.02.018

Kollhoff, A. L., Howell, J. C., and Hu, W. T. (2018). Automation vs. Experience: Measuring Alzheimer's Beta-Amyloid 1-42 Peptide in the CSF. *Front. Aging Neurosci.* 10:253. doi: 10.3389/fnagi.2018.00253

Le Bastard, N., De Deyn, P. P., and Engelborghs, S. (2015). Importance and impact of preanalytical variables on Alzheimer disease biomarker concentrations in cerebrospinal fluid. *Clin. Chem.* 61, 734–743. doi: 10.1373/clinchem.2014.236679

Leitão, M. J., Silva-Spínola, A., Santana, I., Olmedo, V., Nadal, A., Le Bastard, N., et al. (2019). Clinical validation of the Lumipulse G cerebrospinal fluid assays for routine diagnosis of Alzheimer's disease. *Alzheimers Res. Ther.* 11:91. doi: 10.1186/s13195-019-0550-8

Lewczuk, P., Beck, G., Esselmann, H., Bruckmoser, R., Zimmermann, R., Fiszer, M., et al. (2006). Effect of Sample Collection Tubes on Cerebrospinal Fluid Concentrations of Tau Proteins and Amyloid β Peptides. *Clin. Chem.* 52, 332–334. doi: 10.1373/clinchem.2005.058776

Lopez, C., Tucker, S., Salameh, T., and Tucker, C. (2018). An unsupervised machine learning method for discovering patient clusters based on genetic signatures. *J. Biomed. Informat.* 85, 30–39. doi: 10.1016/j.jbi.2018.07.004

Manix, M., Kalakoti, P., Henry, M., Thakur, J., Menger, R., Guthikonda, B., et al. (2015). Creutzfeldt-Jakob disease: updated diagnostic criteria, treatment algorithm, and the utility of brain biopsy. *Neurosurg. Focus* 39:E2. doi: 10.3171/2015.8.FOCUS15328

Mattsson, N., Zegers, I., Andreasson, U., Bjerke, M., Blankenstein, M. A., Bowser, R., et al. (2012). Reference measurement procedures for Alzheimer's disease cerebrospinal fluid biomarkers: definitions and approaches with focus on amyloid β42. *Biomark. Med.* 6, 409–417. doi: 10.2217/bmm.12.39

Mattsson-Carlgren, N., Palmqvist, S., Blennow, K., and Hansson, O. (2020). Increasing the reproducibility of fluid biomarker studies in neurodegenerative studies. *Nat. Commun.* 11:6252. doi: 10.1038/s41467-020-19957-6

McKeith, I. G., Boeve, B. F., Dickson, D. W., Halliday, G., Taylor, J.-P., Weintraub, D., et al. (2017). Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB Consortium. *Neurology* 89, 88–100. doi: 10.1212/WNL.0000000000004058

mcr package | R Documentation (2021). *Method Comparison Regression v1.2.1*. Available online at: https://www.rdocumentation.org/packages/mcr/versions/1.2.1 (accessed March 3, 2021)

Molinuevo, J. L., Blennow, K., Dubois, B., Engelborghs, S., Lewczuk, P., Perret-Liaudet, A., et al. (2014). The clinical use of cerebrospinal fluid biomarker testing for Alzheimer's disease diagnosis: A consensus paper from the Alzheimer's Biomarkers Standardization Initiative. *Alzheimer's Dement.* 10, 808–817. doi: 10.1016/j.jalz.2014.03.003

Paciotti, S., Sepe, F. N., Eusebi, P., Farotti, L., Cataldi, S., Gatticchi, L., et al. (2019). Diagnostic performance of a fully automated chemiluminescent enzyme immunoassay for Alzheimer's disease diagnosis. *Clin. Chim. Acta* 494, 74–78. doi: 10.1016/j.cca.2019.03.1612

Passing, H., and Bablok. (1983). A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I. *J. Clin. Chem. Clin. Biochem.* 21, 709–720. doi: 10.1515/cclm.1983.21.11.709

Paterson, R. W., Slattery, C. F., Poole, T., Nicholas, J. M., Magdalinou, N. K., Toombs, J., et al. (2018). Cerebrospinal fluid in the differential diagnosis of Alzheimer's disease: clinical utility of an extended panel of biomarkers in a specialist cognitive clinic. *Alzheimers Res. Ther.* 10:32. doi: 10.1186/s13195-018-0361-3

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine Learning in Python. *J. Machine Learning Res.* 12, 2825–2830.

Postuma, R. B., Berg, D., Stern, M., Poewe, W., Olanow, C. W., Oertel, W., et al. (2015). MDS clinical diagnostic criteria for Parkinson's disease. *Mov. Disord.* 30, 1591–1601. doi: 10.1002/mds.26424

R Core Team (2013). *R: A language and environment for statistical computing.* Vienna: R Core Team.

Racine, A. M., Koscik, R. L., Berman, S. E., Nicholas, C. R., Clark, L. R., Okonkwo, O. C., et al. (2016). Biomarker clusters are differentially associated with longitudinal cognitive decline in late midlife. *Brain* 139, 2261–2274. doi: 10.1093/brain/aww142

Rektor, I., Bohnen, N. I., Korczyn, A. D., Gryb, V., Kumar, H., Kramberger, M. G., et al. (2018). An updated diagnostic approach to subtype definition of vascular parkinsonism - Recommendations from an expert working group. *Parkinson. Relat. Disord.* 49, 9–16. doi: 10.1016/j.parkreldis.2017.12.030

Relkin, N., Marmarou, A., Klinge, P., Bergsneider, M., and Black, P. M. (2005). Diagnosing idiopathic normal-pressure hydrocephalus. *Neurosurgery* 57, S4–S16. doi: 10.1227/01.neu.0000168185.29659.c5

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., et al. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12:77. doi: 10.1186/1471-2105-12-77

Rokach, L., and Maimon, O. (2005). "Clustering Methods," in *Data Mining and Knowledge Discovery Handbook*, eds O. Maimon and L. Rokach (Boston, MA: Springer), 321–352. doi: 10.1007/0-387-25465-X_15

Román, G. C., Tatemichi, T. K., Erkinjuntti, T., Cummings, J. L., Masdeu, J. C., Garcia, J. H., et al. (1993). Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology* 43, 250–260. doi: 10.1212/wnl.43.2.250

Sawyer, R. P., Rodriguez-Porcel, F., Hagen, M., Shatz, R., and Espay, A. J. (2017). Diagnosing the frontal variant of Alzheimer's disease: a clinician's yellow brick road. *J. Clin. Mov. Disord.* 4:2. doi: 10.1186/s40734-017-0052-4

Scheffer, I. E., Berkovic, S., Capovilla, G., Connolly, M. B., French, J., Guilhoto, L., et al. (2017). ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. *Epilepsia* 58, 512–521. doi: 10.1111/epi.13709

Schwarz, G. (1978). Estimating the Dimension of a Model. *Ann. Statist.* 6, 461–464. doi: 10.1214/aos/1176344136

Simrén, J., Ashton, N. J., Blennow, K., and Zetterberg, H. (2020). An update on fluid biomarkers for neurodegenerative diseases: recent success and challenges ahead. *Curr. Opin. Neurobiol.* 61, 29–39. doi: 10.1016/j.conb.2019.11.019

Skillbäck, T., Farahmand, B. Y., Rosén, C., Mattsson, N., Nägga, K., Kilander, L., et al. (2015). Cerebrospinal fluid tau and amyloid-β1-42 in patients with dementia. *Brain* 138, 2716–2731. doi: 10.1093/brain/awv181

Smith, E. E., and Greenberg, S. M. (2003). Clinical diagnosis of cerebral amyloid angiopathy: validation of the Boston criteria. *Curr. Atheroscler. Rep.* 5, 260–266. doi: 10.1007/s11883-003-0048-4

Solorio-Fernández, S., Carrasco-Ochoa, J. A., Martínez-Trinidad, J., and Fco. (2020). A review of unsupervised feature selection methods. *Artif. Intell. Rev.* 53, 907–948. doi: 10.1007/s10462-019-09682-y

Teunissen, C. E., Petzold, A., Bennett, J. L., Berven, F. S., Brundin, L., Comabella, M., et al. (2009). A consensus protocol for the standardization of cerebrospinal fluid collection and biobanking. *Neurology* 73, 1914–1922. doi: 10.1212/WNL.0b013e3181c47cc2

Thompson, A. J., Banwell, B. L., Barkhof, F., Carroll, W. M., Coetzee, T., Comi, G., et al. (2018). Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol.* 17, 162–173. doi: 10.1016/S1474-4422(17)30470-2

Toombs, J., Paterson, R. W., Schott, J. M., and Zetterberg, H. (2014). Amyloid-beta 42 adsorption following serial tube transfer. *Alzheimers Res. Ther.* 6:5. doi: 10.1186/alzrt236

Toschi, N., Lista, S., Baldacci, F., Cavedo, E., Zetterberg, H., Blennow, K., et al. (2019). Biomarker-guided clustering of Alzheimer's disease clinical syndromes. *Neurobiol. Aging* 83, 42–53. doi: 10.1016/j.neurobiolaging.2019.08.032

Traynor, B. J., Codd, M. B., Corr, B., Forde, C., Frost, E., and Hardiman, O. M. (2000). Clinical features of amyotrophic lateral sclerosis according to the El Escorial and Airlie House diagnostic criteria: A population-based study. *Arch. Neurol.* 57, 1171–1176. doi: 10.1001/archneur.57.8.1171

Villain, N., and Dubois, B. (2019). Alzheimer's Disease Including Focal Presentations. *Semin. Neurol.* 39, 213–226. doi: 10.1055/s-0039-1681041