frontiers in Dementia

TYPE Original Research
PUBLISHED 31 August 2026
DOI 10.3389/frdem.2026.1757680

Check for updates

OPEN ACCESS

EDITED BY
Simon Fristed Eskildsen,
Aarhus University, Denmark

REVIEWED BY
Benedictor Alexander Nguchu,
University of Science and Technology of China, China
Aya Galal,
American University in Cairo, Egypt

*CORRESPONDENCE
Marie Söntgerath
Marie.Soentgerath@uk-halle.de
mariesoentgerath@gmail.com

RECEIVED 30 November 2025
REVISED 10 August 2026
ACCEPTED 11 August 2026
PUBLISHED 31 August 2026
CORRECTED 11 September 2026

CITATION
Söntgerath M, Cutini S, Thöne-Otto AIT, Engel A, Regenbrecht F, Polyakova M, Mueller K, Scherf N, Boehm F, Anderl-Straub S, Landwehrmeyer B, Danek A, Diehl-Schmid J, Fassbender K, Fliessbach K, Jahn H, Kornhuber H, Lauer M, Prudlo J, Synofzik M, Wiltfang J, Otto M and Schroeter ML (2026) Evaluating the correspondence of neuropsychological profiles with clinical diagnostic criteria for frontotemporal lobar degeneration and Alzheimer's disease: insights from *k*-means clustering. *Front. Dement.* 5:1757680. doi: 10.3389/frdem.2026.1757680

COPYRIGHT
© 2026 Söntgerath, Cutini, Thöne-Otto, Engel, Regenbrecht, Polyakova, Mueller, Scherf, Boehm, Anderl-Straub, Landwehrmeyer, Danek, Diehl-Schmid, Fassbender, Fliessbach, Jahn, Kornhuber, Lauer, Prudlo, Synofzik, Wiltfang, Otto and Schroeter. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Evaluating the correspondence of neuropsychological profiles with clinical diagnostic criteria for frontotemporal lobar degeneration and Alzheimer's disease: insights from *k*-means clustering

Marie Söntgerath[1,2,3,4]*, Simone Cutini[4], Angelika I. T. Thöne-Otto[1,2], Annerose Engel[2], Frank Regenbrecht[2], Maryna Polyakova[1], Karsten Mueller[1,5], Nico Scherf[1], Fabiola Boehm[1,2,3], Sarah Anderl-Straub[6], Bernhard Landwehrmeyer[6], Adrian Danek[7,8], Janine Diehl-Schmid[9,10], Klaus Fassbender[11], Klaus Fliessbach[12], Holger Jahn[13], Holger Kornhuber[14], Martin Lauer[15], Johannes Prudlo[16,17], Matthis Synofzik[18], Jens Wiltfang[19,20,21], Markus Otto[3] and Matthias L. Schroeter[1,2]

[1]Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, [2]Department of Cognitive Neurology, University Clinic, University of Leipzig, Leipzig, Germany, [3]Department of Neurology, University Clinic, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany, [4]Department of Developmental Psychology, University of Padova, Padua, Italy, [5]Department of Neurology, Charles University, First Faculty of Medicine and General University Hospital, Prague, Czechia, [6]Department of Neurology, University Clinic, University Ulm, Ulm, Germany, [7]Department of Neurology, University Clinic, Ludwig-Maximilians-University Munich, Munich, Germany, [8]German Center for Neurodegenerative Diseases (DZNE), Research Site Munich, Munich, Germany, [9]School of Medicine, Department of Psychiatry and Psychotherapy, Technical University of Munich, Munich, Germany, [10]kbo-Inn-Salzach-Klinikum, Clinical Center for Psychiatry, Psychotherapy, Psychosomatic Medicine, Geriatrics and Neurology, Wasserburg am Inn, Germany, [11]Department of Neurology, Saarland University, Homburg, Germany, [12]Department of Psychiatry and Psychotherapy, University of Bonn, Bonn, Germany, [13]Department of Psychiatry and Psychotherapy, University Clinic Hamburg-Eppendorf, Hamburg, Germany, [14]Department of Psychiatry and Psychotherapy, University Clinic, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany, [15]Department of Psychiatry, Psychosomatic and Psychotherapy, University of Würzburg, Würzburg, Germany, [16]Department of Neurology, Rostock University Medical Center, Rostock, Germany, [17]German Center for Neurodegenerative Diseases (DZNE), Rostock-Greifswald, Rostock, Germany, [18]Department of Neurodegeneration, Hertie Institute for Clinical Brain Research (HIH), University of Tübingen, Tübingen, Germany, [19]Department of Psychiatry and Psychotherapy, University Medical Center Göttingen, Göttingen, Germany, [20]German Center for Neurodegenerative Diseases (DZNE), Göttingen, Germany, [21]Department of Medical Sciences, Neurosciences and Signaling Group, Institute of Biomedicine (iBiMED), University of Aveiro, Aveiro, Portugal

Introduction: Diagnosis of frontotemporal lobar degeneration (FTLD) is complicated by high heterogeneity within diagnostic groups and overlap between them. Using an unsupervised machine learning technique this study aimed to increase understanding of neuropsychological profiles in the spectrum of FTLD syndromes and to assess the consistency of these neuropsychological profiles with diagnostic criteria.

Methods: Data from the German FTLD Consortium (FTLD-c) including behavioral variant of frontotemporal dementia (bvFTD), primary progressive aphasia (PPA) syndromes, i.e., semantic variant (svPPA), and non-fluent variant (nfvPPA), healthy controls, and amnestic Alzheimer's disease (AD), and logopenic variant (lvPPA) was used. *K*-means clustering was applied on neuropsychological and behavioral data. Analysis was repeated for varying cluster numbers ($k = 2–9$). Resulting clusters were inspected on the correspondence with current diagnostic groups and neuropsychological profiles of the clusters were explored.
Results: The results indicated that, based on neuropsychological profiles, particularly homogeneous clusters emerged for healthy controls, svPPA and bvFTD. In contrast, amnestic AD did not separate well from other diagnostic groups, and nfvPPA and lvPPA clustered together. Inspecting the neuropsychological profiles of the clusters showed that segregation of bvFTD from other patient groups was most supported by companion-rated questionnaires focused on behavior. Patients with svPPA showed impairment in both language and verbal memory tasks. LvPPA and nfvPPA showed impairment in processing speed as well as short term and working memory.
Discussion: Using clustering this study provides insight into data-driven grouping of participants with FTLD based on neuropsychological deep-phenotyping. Whereas our approach confirmed svPPA and bvFTD to have comparatively homogeneous neuropsychological profiles, results suggest that the non-fluent PPA syndromes–lvPPA and nfvPPA–are difficult to segregate clinically, at least with the tests applied in our cohort. Our findings may suggest the need for further specification of clinical diagnostic criteria for these diseases.

KEYWORDS

Alzheimer's dementia, differential diagnosis, frontotemporal dementia, frontotemporal lobar degeneration, *K*-means clustering, neuropsychological tests, primary progressive aphasia

1 Introduction

Dementia is a clinical syndrome characterized by progressive cognitive decline that can manifest in different ways, depending on the underlying pathology and the brain areas affected. Frontotemporal lobar degeneration (FTLD) is the second most common cause for early-onset dementia, following amnestic Alzheimer's disease (AD) (Bang et al., 2015; Rascovsky et al., 2011; Gorno-Tempini et al., 2011). Here, we use FTLD as the umbrella term for the clinical syndromes associated with FTLD. The pre-dominant clinical phenotype of FTLD is behavioral variant frontotemporal dementia (bvFTD). BvFTD is characterized by changes in behavior, most commonly involving apathy and disinhibition but also mental rigidity and loss of empathy. With progression of the disease, dietary changes, such as overeating, and decreased hygiene are observed (Bang et al., 2015). Additional to behavioral changes, diagnostic criteria include as supportive features executive dysfunction with relatively intact memory and visuospatial skills (Rascovsky et al., 2011).

Beyond changes in personality and behavior, FTLD can also manifest as language-predominant dementias, known as primary progressive aphasias (PPA). Three main PPA variants have been described. The non-fluent variant PPA (nfvPPA) is characterized by effortful speech due to speech apraxia or agrammatism making the speech seem telegraphic (Gorno-Tempini et al., 2011; Warren et al., 2013). Over time, patients may develop complete mutism (Grossman, 2010) while comprehension and semantic knowledge are usually spared. In contrast, core symptoms of the semantic variant PPA (svPPA) are loss of semantic knowledge resulting in problems in confrontation naming (i.e., anomia) and word comprehension (Gorno-Tempini et al., 2011; Henry and Grasso, 2018). Speech typically remains fluid while becoming progressively empty in content, reducing to platitudes. During reading, patients may make regularization errors in pronunciation due to loss of associated meaning (Macoir et al., 2021).

A third variant of PPA, the logopenic variant (lvPPA), is commonly considered an atypical AD, as it has been related to underlying AD pathology (Irwin et al., 2013). The main characteristic of lvPPA is disturbed phonemic processing leading to phonological errors in both spontaneous speech and structured tasks of repetition or naming (Gorno-Tempini et al., 2011). Difficulties in repetition may be intensified by phonological short-term memory deficits, with greater impairments for longer sentences (Gil-Navarro et al., 2013). Phonological dyslexia and dysgraphia are present for both words and non-words (Henry and Grasso, 2018).

In clinical practice, correct differential diagnosis of the FTLD syndromes remains difficult often requiring more than 3 years (Beber and Chaves, 2013; van Vliet et al., 2013), particularly in young-onset cases. Due to changes in personality and behavior, a considerable proportion of patients may initially receive a diagnosis of a primary psychiatric disorder (Woolley et al., 2011). For example, apathetic symptoms or changes in eating behavior have been mentioned as reasons for confusing bvFTD with major

depressive disorder. Additional challenges arise due to significant heterogeneity within (Beber and Chaves, 2013; Harris et al., 2013) and considerable overlap between neuropsychological profiles of the dementias described (Bang et al., 2015; Coyle-Gilchrist et al., 2016; Kamath et al., 2019). Deficits in episodic memory function, which are considered a hallmark of amnestic AD, may be more pronounced in FTLD syndromes than previously thought (Hornberger and Piguet, 2012). In bvFTD retrieving problems may be related to executive dysfunction, while memory problems in lvPPA may be due to impaired phonological loop functioning (Hornberger and Piguet, 2012; Piguet et al., 2011; Foxe et al., 2021). Additionally, correct differential diagnosis may be hampered by behavioral symptoms observed in svPPA, such as mental rigidity and loss of empathy, which may resemble those observed in bvFTD (Bang et al., 2015). Lastly, nfvPPA and lvPPA may be particularly likely to be confounded due to similarity in slowed, halting speech (Bürger et al., 2017; Gorno-Tempini et al., 2008; Harris et al., 2019).

As a consequence of diagnostic difficulties observed in clinical settings, calls for revisions of diagnostic criteria have been expressed (Hornberger and Piguet, 2012; Piguet et al., 2011; Harris et al., 2019; Sajjadi et al., 2012). Previous articles have shown that a considerable proportion of patients with a PPA may not meet the diagnostic criteria for one of the three variants and the existence of a logopenic variant has been questioned (Harris et al., 2013; Sajjadi et al., 2012). Concerning bvFTD, disturbances in executive function may not be specific to differentiate bvFTD from other dementia syndromes (Walker et al., 2005; Schroeter et al., 2014; Foran et al., 2021; Reul et al., 2017; Overbeek et al., 2020). More recent literature on bvFTD has shifted the focus toward the study of social and emotional functions (Schroeter et al., 2018; Gossink et al., 2018; Rankin, 2021). Addressing the gap of diagnostic precision is an ongoing objective. It is relevant to psychological and social wellbeing of both patients and caregivers (Musa et al., 2020; Weder et al., 2007) and is necessary to power research on therapy and treatment, especially in the age of new disease-specific drugs (van Dyck et al., 2023).

Given these challenges in differential diagnosis, novel computational approaches such as machine learning (ML) techniques may be promising to improve diagnostic accuracy. This is particularly advantageous in cases where multidimensional data is available and high levels of diagnostic uncertainty persist. By being inherently multivariate ML can aid in finding patterns within data rather than looking for single predictive tests (Yanase and Triantaphyllou, 2019). Unsupervised ML can be used to compute similarity of unlabeled data in a multidimensional space (Fan et al., 2020) in order to delineate subgroups or detect outliers in a data-driven way (Bhardwaj et al., 2017; Habes et al., 2020). By uncovering patient similarities and differences, unsupervised ML may bring new insights to the existing diagnostic challenges. For example, a previous study, applying unsupervised ML on neuroimaging data from patients diagnosed with a PPA, determined five clusters (Matias-Guiu et al., 2019). In turn, neuroimaging-based clustering studies in bvFTD indicated the presence of several subgroups with distinct atrophy patterns (Bruun et al., 2019; Ranasinghe et al., 2016). So far, applying ML on a breadth of neuropsychological variables from patients with FTLD has been rarely explored. Instead, studies have focused on data from language assessments and various neuroimaging techniques for PPA (Fan et al., 2020; Leyton et al., 2014) and bvFTD (Cerami et al., 2016; Josephs et al., 2009), respectively. Further, only few studies have evaluated multiclass analysis (e.g., Zimmerer et al., 2020; Kim et al., 2019) and those that have, often grouped FTLD phenotypes together to separate them from AD, Lewy body dementia, Parkinson's disease dementia or vascular dementia (Lin et al., 2020; Garn et al., 2017).

This study aims to address this gap, applying unsupervised ML to neuropsychological data from a large cohort of participants with diagnoses from the FTLD and AD spectrum. An analysis of deep-phenotyping neuropsychological data may yield new insights for each one of the patient groups as well as for their distinction. The main goal of the current study was to use a data-driven approach to investigate similarities in neuropsychological profiles of patients diagnosed with a FTLD type and to evaluate how well the results correspond to the current diagnostic criteria. For this purpose, *k*-means clustering, an unsupervised ML algorithm, was used to explore behavioral and neuropsychological data from patients diagnosed with bvFTD, svPPA, nfvPPA, lvPPA, amnestic AD and healthy controls. This exploratory analysis may inform differential diagnosis which persists to be difficult. Neuropsychological examinations being a standard part of diagnoses in memory clinics (Schild et al., 2023; Deuschl et al., 2016), results from this study may inform their refinement when differentiating dementias from the FTLD and AD spectrum.

Exploration of the data in this study followed three major goals: (1) Investigating correspondence of clusters with diagnostic groups; (2) Investigating neuropsychological profiles of the emerging clusters; and (3) Comparing the results with current diagnostic standards. To the best of our knowledge, this is the first study to analyze all five diagnostic groups (i.e., bvFTD, svPPA, nfvPPA, lvPPA, amnestic AD) using clustering approaches on neuropsychological data. Focus was put on the exploration of the following contrasts thought to be particularly difficult but clinically relevant: (1) bvFTD vs. amnestic AD, (2) the three PPA variants and (3) svPPA vs. bvFTD.

2 Methods

2.1 Patient data

Patient data was provided by the German FTLD Consortium collected from 13 University clinics or research centers across Germany (Otto et al., 2011). We focused on data from the first participant visit to ensure early disease stage and proximity to diagnostic decision. From the 1,036 participants originally in the dataset, 254 were excluded for not belonging to one of the diagnostic groups of interest (Figure 1). In total 782 participants remained for further analysis. To address missing values, a combination of list-wise deletion and imputation was used. First, participants with more than 20% of missing variables were removed, thus leaving us with participants that had followed relatively complete visits. Next, remaining missing values were imputed using multiple imputation by chained equations (MICE) (van Buuren, 2019). The specific approach of handling missing data is detailed in the Supplementary materials.

The analysis was performed twice, firstly with neuropsychological and clinical data only, secondly additionally

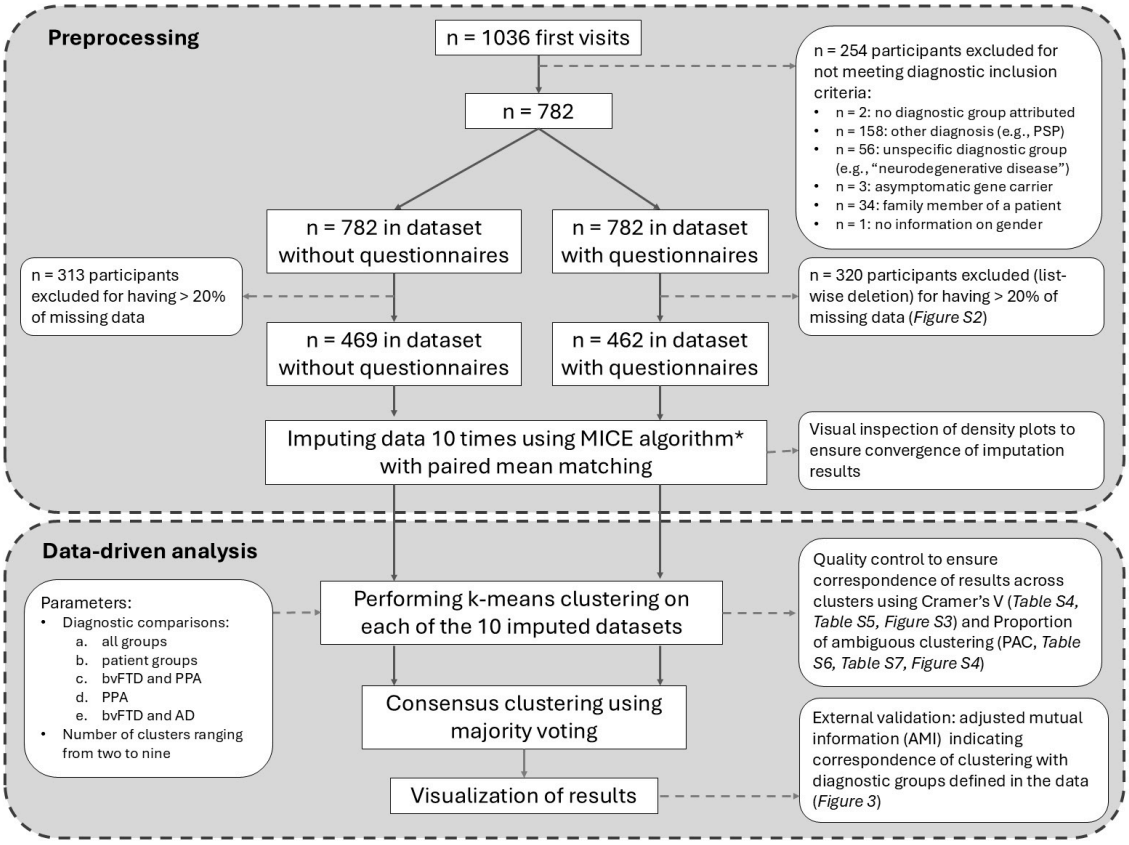


FIGURE 1
Flowchart of participant selection and analysis. *van Buuren and Groothuis-Oudshoorn (2011). AD, amnestic Alzheimer's Disease; bvFTD, behavioral variant frontotemporal dementia; MICE, multiple imputation by chained equations; PPA, primary progressive aphasia; PSP, progressive supranuclear palsy.

including questionnaires. The first analysis without questionnaires was focused on cognitive and language functions to investigate their potential for differential diagnosis. The second analysis took additionally behavioral and personality changes into account, which is particularly relevant for bvFTD. This allowed investigating the additional value of the behavioral questionnaires beyond cognitive assessments. Additionally, due to comparatively high levels of missingness in the behavioral questionnaires, robustness of results could be inspected by comparing them with the analysis without behavioral questionnaires. A total of 469 and 462 participants remained for the analysis without and with questionnaires, respectively (Table 1).

The participants mean age varied between 62.3 and 68.9 years for patients with bvFTD and nfvPPA, respectively. As measure of disease severity, the sum of boxes of the clinical dementia rating (FTLD-CDR SOB) (Morris, 1993; Knopman et al., 2008) was used. As expected, healthy controls differed significantly on the FTLD-CDR SOB compared to all other groups. Additionally, the bvFTD group had significantly higher scores on the FTLD-CDR SOB compared to the three PPA variants. The mean disease durationat study visit ranged from 1.9 to 4.0 years in nfvPPA and lvPPA, respectively, thus indicating significantly later participation in the study in lvPPA compared to nfvPPA (Table 1).

Diagnostic decisions were made by experienced neurologists or psychiatrists based on comprehensive medical, neuropsychological and speech therapeutic examination. Medical information included brain imaging and degenerative markers in the cerebrospinal fluid assessed by lumbar puncture. Of the included participants diagnosed with bvFTD or PPA, 119 (34%) had a clinical, 204 (59%) additional image-supported, and 23 (7%) a neuropathologically confirmed diagnosis in the analysis excluding questionnaires. In the analysis including questionnaires, 124 (36%) patients with bvFTD or PPA had a clinical, 197 (57%) had additional image-supported and 22 (6%) a neuropathologically confirmed diagnosis. Diagnosis was defined according to established international diagnostic criteria (Rascovsky et al., 2011; Gorno-Tempini et al., 2011). Control groups included healthy participants and patients with amnestic AD, as defined by McKhann and colleagues (McKhann et al., 2011) including biomarker data. Participants provided written informed consent. The study was approved by the ethics committees of all contributing Universities (ethics committee University Leipzig ID 137–11–18042011 and University Ulm ID #39/11) and was in accordance with the latest version of the Declaration of Helsinki.

2.2 Neuropsychological test scores

Demographic variables included for further analysis were participants' age and years of education. The neuropsychological

TABLE 1 Demographics and clinical characteristics for included groups and split by analysis excluding and including questionnaires.

Group	N	Total %	Male %	Mean age in years (SD)	Mean years of education in years (SD)	Mean FTLD-CDR SOB (SD)	Mean disease duration in years (SD)
Without questionnaires							
HC	49	10.4	44.9	64.5 (9.3)	15.0 (3.1)	0.1 (0.3)	—
Amnestic AD	74	15.8	54.1	66.1 (9.7)	13.5 (3.3)	5.8 (3.4)	2.8 (2.8)
bvFTD	171	36.5	62.6	62.3 (9.3)	13.5 (3.0)	6.8 (3.9)	3.4 (3.8)
svPPA	52	11.1	46.2	62.9 (7.9)	14.7 (3.2)	4.9 (2.4)	2.6 (1.9)
nfvPPA	76	16.2	48.7	68.9 (8.3)	13.1 (3.3)	4.2 (2.4)	1.9 (1.2)
lvPPA	47	10	48.9	68.7 (6.1)	13.3 (3.5)	4.7 (2.8)	4.0 (4.0)
Chi-square test			n. s.				
Kruskal-Wallis rank sum test				$X^2 = 43.299$ $p = 3.214e{-}08$	$X^2 = 20.139$ $p = 0.001177$	$X^2 = 158.79$ $p < 2.2e{-}16$	$X^2 = 11.09$ $p = 0.02557$
Pairwise comparison using Wilcoxon rank sum test (Bonferroni corrected)				bvFTD–nfvPPA*** bvFTD–lvPPA*** svPPA–nfvPPA*** svPPA–lvPPA** bvFTD–AD*	HC–bvFTD* HC–nfvPPA*	HC and all other groups*** bvFTD–nfvPPA*** bvFTD–lvPPA** bvFTD–svPPA**	nfvPPA–lvPPA*
With questionnaires							
HC	49	10.6	44.9	64.5 (9.3)	15.0 (3.1)	0.1 (0.3)	—
Amnestic AD	70	15.2	52.9	66.4 (9.7)	13.4 (3.4)	5.7 (3.3)	2.8 (2.8)
bvFTD	173	37.4	62.4	62.3 (9.5)	13.4 (2.9)	6.9 (4.0)	3.4 (3.5)
svPPA	50	10.8	46.0	63.2 (7.9)	14.6 (3.3)	4.8 (2.4)	2.6 (1.9)
nfvPPA	72	15.6	47.2	68.7 (8.6)	13.1 (3.4)	4.2 (2.5)	2.0 (1.3)
lvPPA	48	10.4	47.9	68.4 (6.2)	13.3 (3.4)	4.7 (2.8)	4.0 (3.9)
Chi-square test			n. s.				
Kruskal-Wallis rank sum test				$X^2 = 39.28$ $p = 2.085e{-}07$	$X^2 = 18.414$ $p = 0.00247$	$X^2 = 159.14$ $p < 2.2e\text{-}16$	$X^2 = 9.8962$ $p = 0.04221$
Pairwise comparison using Wilcoxon rank sum test (Bonferroni corrected)				bvFTD–nfvPPA*** bvFTD–lvPPA*** svPPA–nfvPPA** svPPA–lvPPA** bvFTD–AD*	HC–bvFTD* HC–nfvPPA*	HC and all other groups*** bvFTD–nfvPPA*** bvFTD–lvPPA** bvFTD–svPPA**	nfvPPA–lvPPA*

AD, Alzheimer's disease; bvFTD, behavioral variant frontotemporal dementia; FTLD CDR SOB, frontotemporal lobar degeneration clinical dementia rating sum of boxes; HC, healthy control; lvPPA, logopenic variant primary progressive aphasia; nfvPPA, non-fluent variant primary progressive aphasia; svPPA, semantic variant primary progressive aphasia; n. s., non-significant; *significant at $p < 0.05$, **significant at $p < 0.01$, ***significant at $p < 0.001$.

10.3389/frdem.2026.1757680

test scores used can broadly be attributed to the six neurocognitive domains defined in the DSM-5 aiming to cover the full spectrum of cognition (American Psychiatric Association, 2013). The FTLD-CDR with its eight sub-domains was used as a clinical staging instrument (Morris, 1993; Knopman et al., 2008). Further, the CERAD-NAB-Plus battery (Moms et al., 1989; Fillenbaum et al., 2008) developed for AD and standardly applied in German memory clinics (Schild et al., 2023) was used. The CERAD-NAB consists pre-dominantly of memory tests (besides BNT, verbal fluency and visuoconstruction), in accordance with the encoding deficits of episodic memory expected in amnestic AD. Additional tests were added specifically for the language, executive functioning and social cognition domains, thought to be more impacted in bvFTD and the PPAs (Rascovsky et al., 2011; Gorno-Tempini et al., 2011; Schroeter et al., 2018). Behavioral and psychiatric changes, as well as impairment in activities of daily living were assessed using questionnaires. All variables included in the analysis are summarized in Table 2. Descriptive statistics for the variables included in the clustering implementation are shown in Supplementary Tables S1, S2 for the analysis without and with questionnaires, respectively.

TABLE 2 List of variables included in the analysis. In brackets the range of test scores is specified where possible.

Category	Variables used for analysis (range)
Demographics	
	Age
	Years of education
Screening tools	
(FTLD-)CDR	FTLD-CDR sum of boxes (0–18)
	CDR sum of boxes (0–12)
	Memory (0–3)
	Orientation (0–3)
	Judgment and problem solving (0–3)
	Community affairs (0–3)
	Home and hobbies (0–3)
	Personal care (0–3)
	Behavior, comportment and personality (0–3)
	Language (0–3)
MMSE	CERAD MMSE total score (0–30)
Neurocognitive domain	
Complex attention	CERAD TMT A (0–180)
	Stroop task-color naming
	Stroop task-word reading
Perceptual-motor function	CERAD visuo-constructional praxis copy (0–11)
Language	CERAD Boston naming test (0–15)
	Repeat and point test–point (0–10)
	Repeat and point test–repeat (0–10)
	CERAD semantic fluency
	CERAD phonemic fluency
	Cookie theft task (0–20)
	AAT token test (0–50)
	AAT written language test (0–90)
Learning and memory	CERAD word list immediate recall 1 (0–10)
	CERAD word list immediate recall 2 (0–10)
	CERAD word list immediate recall 3 (0–10)
	CERAD word list immediate recall total (0–30)
	CERAD word list delayed recall (0–10)
	CERAD word list delayed savings
	CERAD word list recognition (0–100)
	WMS-R digit span forward (0–12)
	WMS-R visual memory span forward (0–14)
	CERAD visuo-constructional praxis delayed recall (0–11)
	CERAD visuo-constructional praxis delayed savings
Executive function	CERAD TMT B (0–300)
	CERAD TMT error scores
	CERAD TMT ratio B/A

(Continued)

2.3 Statistical analysis

Statistical analysis was performed entirely using RStudio (RStudio Team, 2020) version 4.2.0. Analysis was restricted to the first visit of participants with a focus on earliest disease stages. Diagnostic groups were inspected on differences in demographic variables (Table 1). For the subsequent analysis, the two numerical demographic variables age and years of education were included. Each neuropsychological variable was checked for implausible values. Nine datapoints fell outside the defined test range of the respective variables and thus were coded as missing value.

2.3.1 Multiple imputation

Multiple imputation was performed using the mice package (van Buuren and Groothuis-Oudshoorn, 2011) for multiple imputation by chained equations. In case both sub- and total scores were not present in the data, sub-scores were imputed, and total scores derived from the imputations *post hoc*. Imputation was performed 10 times and for a maximum of 100 iterations. To ensure that results of imputation would yield realistic values within the ranges of the variables used, paired mean matching was defined as imputation method. Convergence was ensured by inspecting mean and standard deviation across datasets and iterations as described by van Buuren (2019). Next, density distributions of the original and the imputed data were visually inspected.

2.3.2 Clustering analysis

Prior to *k*-means clustering analysis, variables that were inversely scored were reversed for higher scores to consistently indicate better performance. In the next step the scores were rescaled to vary between −1 and 1 to avoid range effects across

TABLE 2 (Continued)

Category	Variables used for analysis (range)
	WMS-R digit span backward (0–12)
	WMS-R visual memory span backward (0–12)
	Stroop color-word interference
	Stroop error score
	H5PT total
	H5PT total correct
	H5PT percent correct (0–100)
	Cognitive estimation (0–16)
	Applause test (0–6)
Social cognition	RMET (0–24)
Questionnaires	
AES	AES total score (0–54)
B.ADL	B.ADL total score (1–10)
FrSBe	Executive dysfunction frequency (8–40)
	Disinhibition frequency (8–40)
	Apathy frequency (8–40)
GDS	GDS total score (0–15)

AAT, Aachen aphasia test; AES, apathy evaluation scale; B.ADL, Bayer activities of daily living; BNT, Boston naming test; CDR, clinical dementia rating; CERAD, consortium to establish a registry for Alzheimer's disease; FrSBe, Frontal Systems Behavior Scale; FTLD-CDR, frontotemporal lobar degeneration clinical dementia rating; GDS, geriatric depression scale; H5PT, Hamasch five point test; MMSE, mini-mental state examination; RMET, reading the mind in the eyes test; TMT, trail making test; WMS-R, Wechsler memory scale revised.

variables (e.g., Brigadoi et al., 2017). *K*-means clustering was implemented using the *kmeans* function from the *stats* package (Core Team R, 2021). The *k*-means algorithm defined by Hartigan and Wong (Hartigan and Wong, 1979) was used. The number of random initializations was set to nstart = 100 with a maximum number of iterations set to iter-max = 400. In this paper a dimensional approach was taken, in which diagnostic decisions on a spectrum consider external factors, such as available treatment options (Murley et al., 2020). Previous clustering studies have used an internal validation metric such as the elbow method [e.g., (Alexander et al., 2021; Hoffman et al., 2017)] to define the number of clusters best representing the data. In this paper internal validation was not calculated. Instead, analyses were systematically repeated with a varying number of clusters. This allowed for flexible exploration of the data focusing on consistent results across analyses (von Luxburg et al., 2012). Results were evaluated with an external validation metric.

Number of clusters was set to vary between $k = 2$ and $k = 9$. Groups included in the analyses were (a) all groups, (b) only the patient groups, (c) bvFTD and PPAs, (d) only PPAs and (e) bvFTD and amnestic AD. Analyses were performed on each one of the 10 imputed datasets before combining the results using majority voting. Quality checks and their results are detailed in the Supplementary materials. For external validation, correspondence of clustering results with the pre-existing diagnostic labels was calculated using adjusted mutual information (AMI). Due to the unsupervised nature of *k*-means clustering the diagnostic labels had not been included in the clustering analysis, thus serving as reference clinical diagnosis for external validation. AMI evaluates results based on homogeneity and completeness while correcting for chance level [e.g., (Ficiarà et al., 2021)]. Clustering results were further evaluated visually on whether clusters separated between different diagnostic groups and how this changed with increasing number of clusters. Next, results were inspected more in detail to find neuropsychological profiles explaining the clustering observed.

Based on the exploratory nature of the clustering analysis, results will be summarized in a descriptive way as was performed in previous studies [e.g., (Ficiarà et al., 2021; Maruta et al., 2015)]. Focus is on consistent findings observed across cluster sizes and comparisons.

3 Results

3.1 Consensus clustering

3.1.1 Varying the number of clusters and diagnostic groups in the analysis

By varying both the patient groups included and the number of clusters defined in the analysis (see section 2.3.2, second paragraph), we observed that increasing the number of clusters has an effect of "zooming in" to the data structure. Similar results were obtained when including fewer groups or by increasing the number of clusters. Analysis allowed us to observe that some groups separate more easily than others. A two-dimensional representation of the data based on Principal Component Analysis (PCA) can be found in the Supplementary Figure S6.

3.1.2 Cluster results

To evaluate the results from the cluster analysis, the distribution of diagnostic groups within each cluster was visualized and is shown in Figure 2 for the analysis including behavioral questionnaire scores. The corresponding results from the analysis without questionnaires are shown in Supplementary Figure S5. Overall, results were considerably stable across various analyses. While no cluster showed complete correspondence with one diagnostic group, coherence between patient's diagnosis and the cluster results was partly supported. This was particularly true for participants diagnosed with bvFTD or svPPA, who, to a great proportion clustered together and separately from other diagnostic groups. Healthy controls clustered together but not separately from patients. By increasing *k*, healthy controls progressively separated better from patient groups reaching maximum homogeneity at $k = 9$ (excluding questionnaires: 75%, including questionnaires: 80%). Participants diagnosed with amnestic AD did not separate well from other diagnostic groups. Amnestic AD separated well from healthy controls while mixing to a similar extent with all other patient groups. In the analysis including only bvFTD and amnestic AD, maximum homogeneity reached by an amnestic AD cluster was 74% and 85% excluding and including questionnaire scores, respectively. Further, nfvPPA and lvPPA consistently clustered together while clustering apart from other groups. At low number of clusters, the three PPAs grouped together but separately from

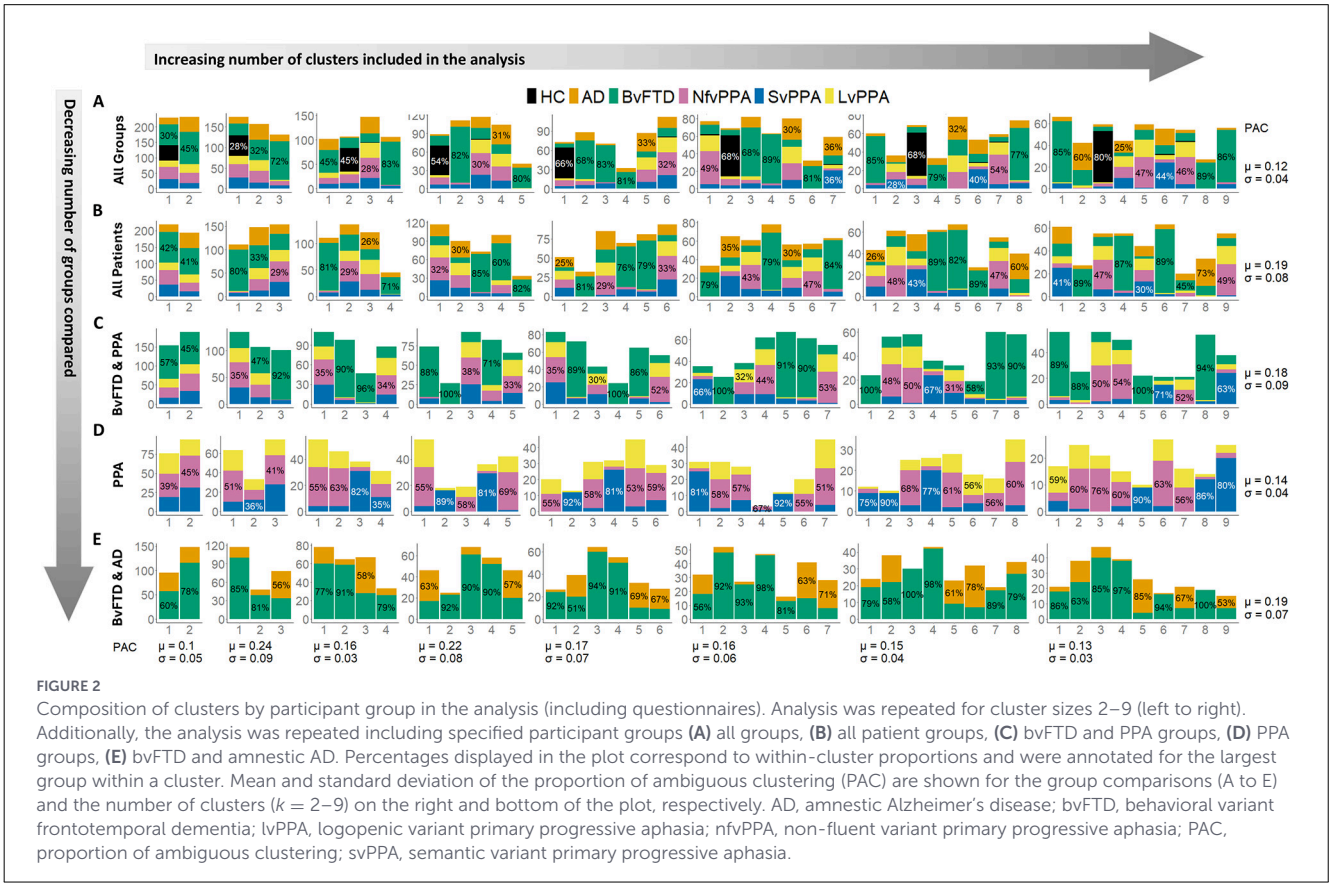


FIGURE 2
Composition of clusters by participant group in the analysis (including questionnaires). Analysis was repeated for cluster sizes 2–9 (left to right). Additionally, the analysis was repeated including specified participant groups **(A)** all groups, **(B)** all patient groups, **(C)** bvFTD and PPA groups, **(D)** PPA groups, **(E)** bvFTD and amnestic AD. Percentages displayed in the plot correspond to within-cluster proportions and were annotated for the largest group within a cluster. Mean and standard deviation of the proportion of ambiguous clustering (PAC) are shown for the group comparisons (A to E) and the number of clusters ($k = 2–9$) on the right and bottom of the plot, respectively. AD, amnestic Alzheimer's disease; bvFTD, behavioral variant frontotemporal dementia; lvPPA, logopenic variant primary progressive aphasia; nfvPPA, non-fluent variant primary progressive aphasia; PAC, proportion of ambiguous clustering; svPPA, semantic variant primary progressive aphasia.

other groups. At a higher number of clusters, single clusters emerged that consisted mainly of svPPA or a combination of nfvPPA and lvPPA.

External evaluation using AMI is shown in Figure 3. This score varies between zero and one and indicates correspondence of cluster results with the diagnosis registered in the data set. By increasing the number of groups included in the analyses AMI leveled at a higher number of clusters: leveling off at 2–3 clusters when including amnestic AD and bvFTD, 4–5 clusters when including PPA and 4–6 clusters when including PPA and bvFTD. When including more groups (all patients or all groups), no clear leveling off was observed, possibly suggesting that more clusters would be needed to capture possible groups in the data.

3.2 Inclusion of questionnaire scores

Systematically varying the inclusion of questionnaire scores in the analysis showed that scores of companion-rated questionnaires focused on behavior helped in distinguishing bvFTD from other groups, particularly from PPAs and healthy controls (Figure 2 and Supplementary Figure S5). Including questionnaires enabled a separation of bvFTD from other groups at lower number of clusters. When including PPA and bvFTD, inclusion of questionnaires seemed to favor separation of bvFTD while making separation of PPA groups more difficult, requiring larger number of clusters for svPPA to cluster apart. In the analysis with all groups, excluding questionnaires resulted in a greater proportion of bvFTD patients clustering with healthy controls also at a high number of clusters (compare Figure 2 and Supplementary Figure S5 part A at $k = 6$ and $k = 7$). Questionnaire scores that seemed of special relevance were the Frontal Systems Behavior Scale (FrSBe) and the Apathy Evaluation Scale (AES; e.g., see Figure 4). The usefulness of these questionnaires for differential diagnoses between non-bvFTD groups seemed limited as expected.

3.3 Neuropsychological profile

Next, radar plots visualizing scores of clusters across the range of neuropsychological tests, were inspected. Large heterogeneity was observed both within and between diagnostic groups and explains the fuzzy separation of diagnostic groups observed previously in Figure 2 (also, see Supplementary Figure S5). The large majority of healthy controls grouped in a single cluster which corresponded to the best performing cluster on virtually all tests. Interestingly, the cluster containing healthy control participants was partly outperformed on the FrSBe Disinhibition scale by clusters containing most nfvPPA and lvPPA patients. Further, one consistent finding was a binary influence of symptom severity on clustering results. At low number of clusters ($k = 2$ and $k = 3$), symptom severity seemed to be driving the cluster separation. When several clusters emerged that mainly consisted of the same diagnostic group, the difference between clusters seemed in many cases attributable to symptom severity

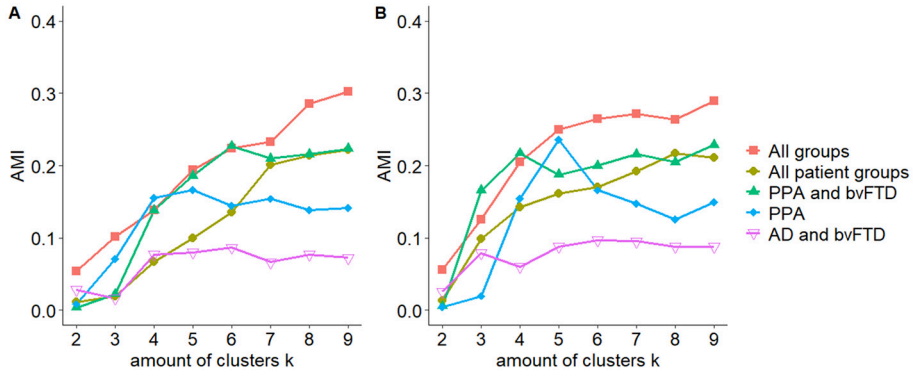


FIGURE 3
Adjusted mutual information (AMI) was calculated to quantitatively evaluate correspondence of clustering results with attributed diagnostic groups. **(A)** AMI in the analysis excluding questionnaires. **(B)** AMI in the analysis including questionnaires. Generally, by increasing number of diagnostic groups included, maximum AMI was reached with higher number of clusters. Comparing graphs A (without questionnaires) and B (with questionnaires) may demonstrate particular importance of questionnaires in discriminating patients with bvFTD. In the analyses including bvFTD a maximum AMI was reached with a lower number of clusters. AD, amnestic Alzheimer's disease; AMI, adjusted mutual information; bvFTD, behavioral variant frontotemporal dementia; PPA, primary progressive aphasia.

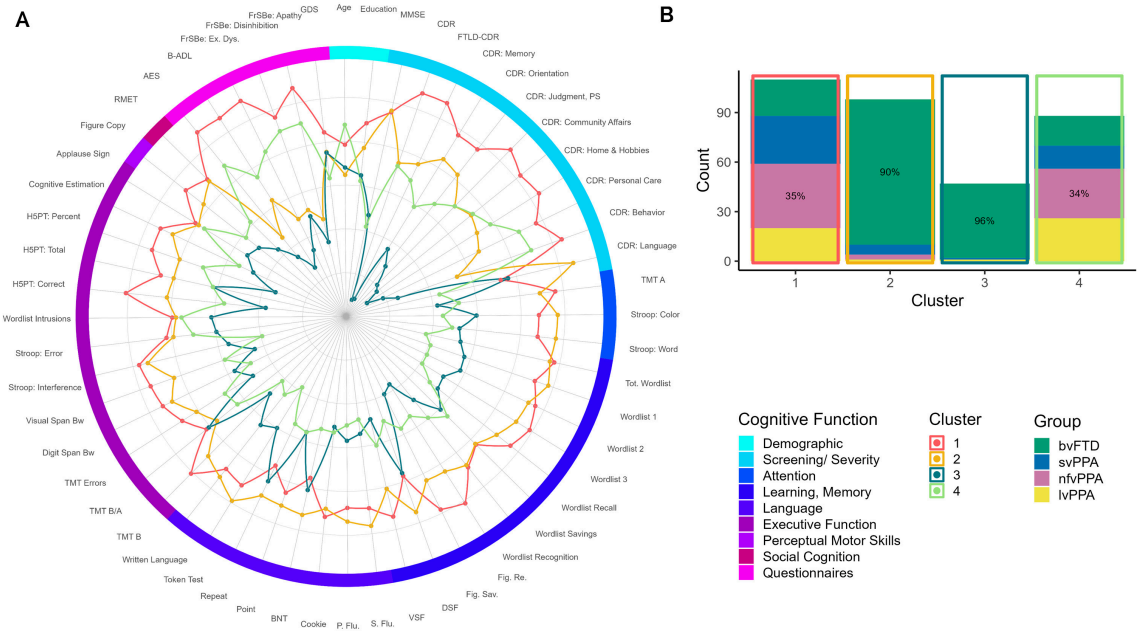


FIGURE 4
Neuropsychological profiles of cluster results (including questionnaires, analysis including PPA and bvFTD, $k = 4$). Example plot representing neuropsychological profiles for the analysis including patients with PPA and bvFTD, including questionnaires and limiting the analysis to $k = 4$ clusters. For the analysis scores were scaled to vary between −1 and 1. For visualization purposes only, cluster centers were z-standardized. Scores more at the inside of the radar plot correspond to worse performance. Comparing clusters 2 (yellow) and 3 (dark green) that both consist mainly of bvFTD patients, cluster 2 performs better on virtually all assessments indicating an influence of symptom severity on clustering results. Compared to clusters 1 (red) and 4 (light green) these two clusters perform particularly bad on the questionnaires, the FTLD-CDR subscores "Home and Hobbies", "Personal Care" and "Behavior, comportment and personality". **(A)** Radar Plot indicating z-standardized cluster centers for clusters 1–4. **(B)** Box plot indicating patient groups within clusters 1–4. Percentages displayed in the plot correspond to within-cluster proportions and were annotated for the largest group within a cluster. AES, Apathy Evaluation Scale; B-ADL, Bayer activities of daily living; BNT, Boston naming test; BvFTD, behavioral variant frontotemporal dementia; CDR, Clinical Dementia Rating Scale; DSF, digit span forward; Ex. Dys, executive dysfunction; Fig. Re, figure recall (from CERAD battery); Fig. Sav, Figure savings (from CERAD battery); FrSBe, Frontal Systems Behavior Scale; FTLD-CDR, Frontotemporal Lobar Degeneration Clinical Dementia Rating Scale; GDS, Geriatric Depression Scale; H5PT, Hamasch five point Test; LvPPA, logopenic variant primary progressive aphasia; MMSE, mini-mental state examination; NfvPPA, non-fluent variant primary progressive aphasia; TMT, trail making test; P. Flu, phonematic fluency, RMET, reading the mind in the eyes test; S. Flu, semantic fluency; SvPPA, semantic variant primary progressive aphasia; VSF, visual span forward.

(see as example clusters 2 and 3 in Figure 4). At higher number of clusters, disease-specific patterns were observed. In the following, consistent disease-specific patterns are described for the FTLD subgroups. Exemplary figures illustrating the findings are shown below (Figure 4) and in the Supplementary Figures S7–S15).

3.3.1 BvFTD

Clusters with mainly bvFTD patients scored worse than other clusters on FTLD-CDR subscores, particularly on "Behavior, comportment, & personality", "personal care", "Home & hobbies" and "personal affairs" (Figure 4). Further, a considerable difference between the "Language" and the "Behavior, comportment, & personality" subscores of the FTLD-CDR was observed in comparison to PPA clusters. While bvFTD scored considerably better on the "Language" subscore, the opposite was the case for PPA clusters. Additionally, lower scores on companion-rated questionnaires, especially on AES and FrSBe, separated bvFTD from other participant groups. This was the case also when cognitive performance was similar to healthy controls. In contrast, at high levels of cognitive impairment, bvFTD clusters maintained good performance on the BNT and repeat and point task.

Separation of amnestic AD and bvFTD in different clusters was incomplete, complicating interpretation of results. Particularly scores on executive and memory assessments did not seem to differ meaningfully between these two groups. However, amnestic AD clusters mirrored performance of the most impaired bvFTD clusters on memory scores, suggesting greater relative rather than absolute impairment on these tests.

3.3.2 PPA

Generally, PPA clusters showed a great drop from the "Behavior, comportment and personality" subscore to the "Language" subscore. One cluster that in many cases consisted mainly of lvPPA and nfvPPA patients, showed lowest performance on the MMSE while the cluster performing worst on the FTLD-CDR usually consisted mainly of bvFTD patients. SvPPA clusters performed worst on BNT and the point part of the repeat and point task compared to other clusters.

Comparing svPPA and nfvPPA/lvPPA clusters, a different pattern emerged on the repeat and point task for clusters containing svPPA patients compared to clusters containing patients diagnosed with lvPPA and nfvPPA (e.g., see Supplementary Figure S12). While absolute scores varied largely in both types of clusters, lvPPA and nfvPPA patients scored similarly on both repeat and point while svPPA patients showed better performance on the repeat compared to the point part of the task. Additionally, compared to lvPPA and nfvPPA, svPPA clusters scored lower on the FTLD-CDR subscore "Behavior, comportment, and personality". Generally, patients with svPPA scored low on BNT, Cookie Theft and verbal but not phonemic fluency assessments. LvPPA and nfvPPA patients instead showed greater impairment on tasks with smaller semantic but greater attention and executive components. For example, these clusters typically scored worse on the TMT A and B, digit span, block span and the Stroop, compared to svPPA clusters. This may indicate that these patients have greater problems in tasks assessing non-language domains. However, despite these tasks being developed to assess non-language domains, particularly the TMT B, digit span and Stroop, rely on language processing and speech production. Thus, one may question whether deficits observed in these tasks are a side-effect of underlying problems with language processing. One interesting question for future investigations will be whether these tasks could be more sensitive to differential diagnosis between svPPA and nfvPPA/lvPPA.

4 Discussion

The present study used an unsupervised ML algorithm to analyze neuropsychological and behavioral data from patients with FTLD. Comparison groups included in the study were healthy controls and patients with amnestic AD and associated lvPPA, to contain the clinically most relevant groups. The goal of the study was to explore neuropsychological profiles from these different groups and compare results with the current clinical diagnostic criteria. Differential diagnosis of these patient groups remains difficult, and misdiagnoses are common (Rascovsky and Grossman, 2013). To date, diagnostic certainty can be reached only via pathological confirmation post-mortem (Deuschl et al., 2016) and there is no gold-standard for diagnosis prior to death. Currently, biomarker profiles from imaging, cerebrospinal fluid and blood are being developed (Arevalo-Rodriguez et al., 2021; Hüper et al., 2024).

To summarize results, meaningful correspondence between patient's diagnosis and clustering results was observed across repeated analyses of varying patient groups, variables (including or excluding questionnaires), and numbers of clusters k defined for the analysis. Segregation in individual clusters was more pronounced for bvFTD and svPPA supporting the assumption of specific clinical profiles. Interestingly, amnestic AD patients mixed with all other patient groups to a similar extent. NfvPPA and lvPPA patients–both representing non-fluent PPA subtypes in contrast to fluent svPPA–grouped together but separately from the other groups. Remarkably, these results correspond to imaging studies applying ML to classify FTLD subtypes based on their different atrophy patterns (Lampe et al., 2022, 2023). While svPPA could be identified with high accuracy, the separation between lvPPA and nfvPPA was impossible, operating almost at chance level (Bisenius et al., 2017). This underlines the importance of fluid biomarkers for the differentiation of nfvPPA and lvPPA (Hüper et al., 2024). Alternatively, the used tests assessing cognitive and particularly language functioning may not be sufficiently sensitive for the distinction of lvPPA and nfvPPA. Particularly, no tasks assessing phonological working memory or speech apraxia were included in the current study.

Further, the findings of the current study supported the following aspects that are in accordance with former studies and disease concepts: (1) A hierarchical procedure of diagnosing PPA before defining the exact PPA variant is in accordance with current diagnostic criteria (Gorno-Tempini et al., 2011). A particularly high score on the Language item of the FTLD-CDR, indicating high levels of impairment, may be sensitive to PPA. A general instrument such as the FTLD-CDR, to a large extent based on informants' observations, may thus be used in a first step to understand impairment of language functioning above other cognitive/functional domains. In a second step, formal, quantitative neuropsychological testing of language deficits, particularly of semantic processing via the BNT and the Repeat and Point, may instead be useful and required to distinguish

svPPA from nfvPPA/lvPPA. A drop of performance from the repeat to the point task as indicative of svPPA symptomatology was supported by the current study (Henry and Grasso, 2018; Seckin et al., 2022). Apart from language assessments, results highlight the possible added value of including assessments of other cognitive domains, particularly of attention and executive function in the differential diagnosis of PPA variant. (2) In accordance with current diagnostic criteria, separation of bvFTD from other diagnostic groups seemed to be facilitated by including measures of changes in behavior rather than cognition. So far, the best way to detect these changes are companion-rated questionnaires, e.g., AES and FrSBe and the FTLD-CDR Behavior, comportment, & personality (Rascovsky et al., 2011). (3) Memory impairments in bvFTD may be as pronounced as in amnestic AD. Sparing of memory dysfunction as stated in the diagnostic criteria of bvFTD is only true relative to other dysfunction (Rascovsky et al., 2011). A dissociation of amnestic AD and bvFTD may only be observed when evaluating memory relative to general cognitive functioning. This was shown in the current study using the (FTLD-) CDR. BvFTD patients showed either relatively preserved memory as well as (FTLD-) CDR global scores or high impairment on both dimensions. Amnestic AD instead showed greater relative impairment on memory functions compared to their FTLD-CDR scores, serving as measure of disease severity.

Furthermore, the current study revealed findings contradicting former literature. While bvFTD was related to worse scores on the FrSBe Executive Dysfunction scale, standard assessments of executive functioning may not aid separation of bvFTD from the other patient groups. A review by Johnen and Bertoux (2019) previously highlighted the usefulness of the TMT B error score which did however not show meaningful distinction of bvFTD and other groups in the current study. Similarly, despite previous studies suggesting social cognition to differentiate bvFTD from healthy controls (Schroeter et al., 2018) and other neurodegenerative or psychiatric diagnoses (Gossink et al., 2018; Johnen and Bertoux, 2019; Ducharme et al., 2020) the current study did not find supporting evidence. Additional results suggest that cognitive involvement in the PPA variants may be larger than might have been expected and may exceed the language domain. This was found despite data in the current study being restricted to patient's first visit and average delay since symptom onset for PPA patients being less than 3 years. SvPPA and nfvPPA/lvPPA clusters differed on assessments of processing speed, short term and working memory. Previous studies indicated that patients with lvPPA show greater cognitive deficits compared to both nfvPPA and svPPA (Kamath et al., 2020). Additionally, impairment in working memory tasks was shown in both lvPPA and nfvPPA (Harris et al., 2019) while patients with svPPA may not be impaired in these tasks (Foxe et al., 2021; Kamath et al., 2020). Distinction of lvPPA from nfvPPA based on verbal and visuospatial working memory tasks could however not be replicated.

The findings of this study highlight the heterogeneity of neuropsychological profiles within diagnostic groups, as well as considerable overlap between them. This contrasts with clinical diagnostic criteria suggesting relatively circumscribed impairments in episodic memory, language or executive functioning for amnestic AD (McKhann et al., 2011), the PPAs (Gorno-Tempini et al., 2011) and bvFTD (Rascovsky et al., 2011), respectively. Importantly, our results suggest that the variability observed may largely be explained by differences in disease severity. Greater overlap in clinical symptoms between patient groups may be explained by expansion of neuropathology (Bürger et al., 2017). Consequently, differences in disease severity may partly obscure disease-specific cognitive profiles.

These findings highlight the importance of comprehensive neuropsychological examination across all six cognitive domains described in the DSM-5 (Sachdev et al., 2014) to interpret results from one cognitive domain with respect to results in other domains. Additionally, neuropsychological testing may benefit from interpretation relative to disease stage. While z-standardized scores remain essential for characterizing impairment compared to healthy controls, as well as for treatment planning, differential diagnosis may benefit from stage-specific comparisons. Future studies may therefore benefit from focusing on clearly defined disease stages, whether by stratifying patient groups based on staging instruments such as the FTLD-CDR or by applying ML tools such as subtype and stage inference (SuStaIn) (Young et al., 2018). Lastly, longitudinal follow-up of patients may provide insights into expected changes of cognitive profiles with disease progression and improve interpretation of neuropsychological results in the context of differential diagnosis.

4.1 Limitations and future studies

We did not find evidence suggesting that social cognition or executive function assessments aid the diagnosis of bvFTD. This may be linked to the tests used in the current study. For social cognition a single assessment, with high demands on semantic processing and a considerable proportion of missingness was included. Possible differences between participants on this task may have been washed out by algorithmic limitations and noise. Additionally, both executive and social function are complex constructs and relevant impairments may not have been reflected appropriately by the assessments used. The results encourage the investigation of alternative tests that minimize language barriers. Tests previously suggested for differential diagnosis include the Social Cognition and Emotional Assessment (SEA) (Bertoux et al., 2012), the ultimatum game (Hinterbuchinger et al., 2018), or tasks developed for the observation of spontaneous social behavior (Rankin et al., 2008). Concerning executive function, more recent studies suggest that the inclusion of a summed executive error index, calculated from perseverations, rule violations and stimulus-bound errors, may help to detect executive problems in bvFTD (Kamath et al., 2019; Kramer et al., 2003). Furthermore, executive function tests such as the Behavioral Assessment of the Dysexecutive Syndrome (BADS) might be applied that tap more in everyday life functions than classical tests (Schroeter et al., 2012). Future work is necessary to investigate sensitive and specific neuropsychological tests for bvFTD and diagnostic criteria need to be refined accordingly.

Biological sex was not included as demographic variable in the current analysis while recent studies suggest that sex has an impact on both clinical symptoms and extent of brain atrophy in dementia syndromes. Although results are not conclusive and may be limited if sex (and, in particular, gender) is regarded as binary, sexual dimorphism was shown for PPA (Rogalski et al., 2007; Sebastian et al., 2018), bvFTD (Illán-Gala et al., 2021) and AD (Gamberger et al., 2016).

Furthermore, one might criticize that surrogate biomarkers for histopathology from imaging, cerebrospinal fluid and blood were not included. These are particularly relevant for the identification of Alzheimer's pathology in amnestic AD and lvPPA (Hüper et al., 2024). However, our paper focused on behavioral and cognitive measures, and the results of these biomarkers were included in the diagnostic procedure to define diseases in our cohort (Hüper et al., 2024). Results from this clinical "gold-standard" were used as external validation of the clustering results.

In contrast to previous clustering studies (Cerami et al., 2016; Owens et al., 2018) the current study did not support the existence of robust subgroups within current diagnoses. However, diagnostic precision may always present a trade-off between simplifying the clinical picture of a patient while remaining meaningful for treatment or disease management. Instead of aiming to find the number of subgroups best explaining the data, as performed in previous studies using internal validation metrics, we intended to explore data structure flexibly. By evaluating results that we observed repeatedly across analyses, results may be independent of algorithmic choices, such as the number of clusters or the groups included for comparison and thus allow for generalization to other analyses. *K*-means clustering seemed very suitable for multiclass analysis including patient groups from various etiologies. Despite high clinical relevance, multiclass analysis is not common in unsupervised learning and to our knowledge the current study was the first one to jointly investigate amnestic AD, bvFTD and (the three subtypes of) PPA in this way. We propose to use these methods for exploration of patterns, heterogeneity, and overlap. Nonetheless, exploration remains limited by the method of clustering used and the means of data visualization at hand (von Luxburg et al., 2012). *K*-means clustering used in the current study is one of several clustering methods, each method relying on specific assumptions. *K*-means clustering aims to minimize clustering error, expressed as sum of the squared Euclidean distance, thereby favoring spherical clusters with similar cluster sizes, also referred to as the "uniform effect" (Zhou and Yang, 2020). Applying alternative clustering approaches may lead to slightly different results and may shed light on new aspects of the disease spectrum.

Lastly, some clusters emerged that remained mixed and were difficult to interpret. Missing pathological or neuroimaging information does not explain well the clusters that contained patients from all or most diagnostic groups. It may demonstrate the heterogeneity of cognitive symptoms observed for patients of the same diagnostic group. Additionally, it is conceivable that these clusters contain misdiagnoses. Pathological data existed only for a minority of participants included in the study. Future studies may investigate these mixed clusters more closely to understand whether additional, multimodal data could help the distinction of these groups.

5 Conclusion

Using a data-driven approach on a breadth of neuropsychological and behavioral data, this study demonstrated substantial heterogeneity in clinical profiles within diagnostic groups across the FTLD and AD spectrum, as well as overlap between them. Despite focusing on the baseline visit, heterogeneity within and overlap between groups could partly be explained by differences in disease severity. At higher number of clusters, patients with svPPA and bvFTD clustered separately from other patient groups, while grouping of the remaining patient groups included was more mixed. LvPPA and nfvPPA–as non-fluent PPA variants–did not separate well based on the cognitive and behavioral data included in the study.

Focusing on clusters with similar disease severity, similar syndrome-specific patterns were observed across analyses. High relevance of companion-rated questionnaires assessing apathy, executive dysfunction and disinhibition was shown for separating patients with bvFTD from other groups. Semantic impairment in svPPA was shown by low scores on both language and memory tests. NfvPPA and lvPPA patients instead showed impairment in processing speed and short term and working memory assessments. Results from this study support the need for revising the neuropsychological profile included in the clinical diagnostic criteria of bvFTD. Differential diagnosis of PPA variants may instead benefit from including additionally specifications of expected impairments on assessments that do not primarily focus on language functioning.

In conclusion, these findings–partial homogeneous clustering and syndrome-specific profiles–underscore the relevance of a patient-centered understanding of patients' neuropsychological and behavioral profiles. Data-driven approaches that are independent of prior labeling of the data, may therefore prove particularly useful for analyzing data from highly heterogeneous diagnostic groups or unclassifiable patients.

Data availability statement

The datasets presented in this article are not readily available because it contains sensible patient information. Data in anonymized form will be made available on reasonable request. Requests to access the datasets should be directed to Marie.Soentgerath@uk-halle.de.

Ethics statement

The studies involving humans were approved by Ethics Committees of all contributing Universities and was in accordance with the latest version of the Declaration of Helsinki (ethics committee University Leipzig ID 137-11-18042011 and University Ulm ID #39/11). The studies were conducted in accordance with the local legislation and institutional requirements. Where possible, informed consent was provided by participants themselves. When participants were not able to give informed consent (e.g., due to

severe cognitive impairment), it was provided by participants' legal guardians/next of kin.

Author contributions

MS: Writing – review & editing, Conceptualization, Formal analysis, Writing – original draft, Methodology, Data curation, Visualization, Investigation. SC: Supervision, Methodology, Writing – review & editing. AT-O: Investigation, Writing – review & editing. AE: Investigation, Writing – review & editing. FR: Writing – review & editing, Investigation. MP: Conceptualization, Writing – review & editing. KM: Conceptualization, Writing – review & editing. NS: Data curation, Writing – review & editing, Methodology. FB: Conceptualization, Writing – review & editing. SA-S: Investigation, Writing – review & editing. BL: Writing – review & editing, Funding acquisition. AD: Investigation, Writing – review & editing. JD-S: Investigation, Writing – review & editing. KFa: Investigation, Writing – review & editing. KFl: Writing – review & editing, Investigation. HJ: Writing – review & editing, Investigation. JK: Writing – review & editing, Investigation. ML: Investigation, Writing – review & editing. JP: Writing – review & editing, Investigation. MSy: Investigation, Writing – review & editing. JW: Investigation, Writing – review & editing. MO: Investigation, Writing – review & editing. MSc: Conceptualization, Supervision, Writing – review & editing, Investigation, Methodology, Funding acquisition, Writing – original draft, Data curation.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This study has been supported by grants from the German Research Foundation (DFG; SCHR 774/5-1 to MSc), the German Federal Ministry of Education and Research (project FTLDc 01GI1007A), the Sächsische Aufbaubank (SAB; eHealthSax Initiative, project TelDem) and the European Union (EFRE InfraProNet, 100757914, NeuroTrace) to MSc. Accordingly, this study is co-financed with tax revenue based on the budget approved by the Saxon state parliament.

Acknowledgments

We thank all participants and the team of the FTLD Consortium, who made this study possible.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The authors KM, NS and MSc declared that they were an editorial board member of Frontiers at the time of submission. This had no impact on the peer review process and the final decision.

Correction note

This article has been corrected with minor changes. These changes do not impact the scientific content of the article.

Generative AI statement

The author(s) declared that Generative AI was used in the creation of this manuscript. During major revision of the manuscript, ChatGPT (OpenAI) was used solely to improve readability of the text. All suggestions were carefully reviewed and edited by the author(s) who take full responsibility for the content of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frdem.2026.1757680/full#supplementary-material

References

Alexander, N., Alexander, D. C., Barkhof, F., and Denaxas, S. (2021). Identifying and evaluating clinical subtypes of Alzheimer's disease in care electronic health records using unsupervised machine learning. *BMC Med Inform Decis Mak*. 21:343. doi: 10.1186/s12911-021-01693-6

American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders. 5th edn. (Washington, DC). doi: 10.1176/appi.books.9780890425596

Arevalo-Rodriguez, I., Smailagic, N., Roqué-Figuls, M., Ciapponi, A., Sanchez-Perez, E., Giannakou, A., et al. (2021). Mini-mental state examination (MMSE) for the early detection of dementia in people with mild cognitive impairment

(MCI). *Cochr. Database Syst Rev.* 2021:CD010783. doi: 10.1002/14651858.CD010783.pub3

Bang, J., Spina, S., and Miller, B. L. (2015). Non-Alzheimer's dementia 1 frontotemporal dementia. *The Lancet*. 386, 1672–1682. doi: 10.1016/S0140-6736(15)00461-4

Beber, B. C., and Chaves, M. L. F. (2013). Evaluation of patients with behavioral and cognitive complaints: Misdiagnosis in frontotemporal dementia and Alzheimer's disease. *Dement Neuropsychol*. 7, 60–65. doi: 10.1590/S1980-57642013DN70100010

Bertoux, M., Delavest, M., de Souza, L. C., Funkiewiez, A., Lépine, J,-P., Fossati, P., et al. (2012). Social cognition and emotional assessment differentiates frontotemporal dementia from depression. *J. Neurol. Neurosurg. Psychiatry*. 83, 411–416. doi: 10.1136/jnnp-2011-301849

Bhardwaj, R., Nambiar, A. R., and Dutta, D. A. (2017). Study of machine learning in healthcare. *2017 IEEE 41st annual computer software and applications conference (COMPSAC)*. IEEE. 2, 236–41. doi: 10.1109/COMPSAC.2017.164

Bisenius, S., Mueller, K., Diehl-Schmid, J., Fassbender, K., Grimmer, T., Jessen, F., et al. (2017). Predicting primary progressive aphasias with support vector machine approaches in structural MRI data. *Neuroimag. Clin.* 14, 334–343. doi: 10.1016/j.nicl.2017.02.003

Brigadoi, S., Cutini, S., Meconi, F., Castellaro, M., Sessa, P., Marangon, M., et al. (2017). On the role of the inferior intraparietal sulcus in visual working memory for lateralized single-feature objects. *J. Cogn. Neurosci*. 29, 337–351. doi: 10.1162/jocn_a_01042

Bruun, M., Koikkalainen, J., Rhodius-Meester, H. F. M., Baroni, M., Gjerum, L., van Gils, M., et al. (2019). Detecting frontotemporal dementia syndromes using MRI biomarkers. *Neuroimage Clin*. 22:101711. doi: 10.1016/j.nicl.2019.101711

Bürger, K., Arzberger, T., Stephan, J., Levin, J., and Edbauer, D. (2017). Pathomechanismen und klinische Aspekte der frontotemporalen Lobärdegeneration. *Nervenarzt*. 88, 163–172. doi: 10.1007/s00115-016-0259-x

Cerami, C., Dodich, A., Lettieri, G., Iannaccone, S., Magnani, G., Marcone, A., et al. (2016). Different FDG-PET metabolic patterns at single-subject level in the behavioral variant of fronto-temporal dementia. *Cortex*. 83, 101–112. doi: 10.1016/j.cortex.2016.07.008

Core Team R (2021). *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.

Coyle-Gilchrist, I. T. S., Dick, K. M., Patterson, K., Vázquez Rodríquez, P., Wehmann, E., Wilcox, A., et al. (2016). Prevalence, characteristics, and survival of frontotemporal lobar degeneration syndromes. *Neurology*. 86, 1736–1743. doi: 10.1212/WNL.0000000000002638

Deuschl, G., Maier, W., Jessen, F., and Spottke, A. (2016). S3-Leitlinie Demenzen. *Deutsche Gesellschaft Für Neurologie, Hrsg Leitlinien Für Diagnostik Und Therapie in Der, Neurologie*. Available online at: www.dgn.org/leitlinien (Accessed June 15, 2021).

Ducharme, S., Dols, A., Laforce, R., Devenney, E., Kumfor, F., van den Stock, J., et al. (2020). Recommendations to distinguish behavioural variant frontotemporal dementia from psychiatric disorders. *Brain*. 143, 1632–1650. doi: 10.1093/brain/awaa018

Fan, J. M., Gorno-Tempini, M. L., Dronkers, N. F., Miller, B. L., Berger, M. S., Chang, E. F., et al. (2020). Data-driven, visual framework for the characterization of aphasias across stroke, post-resective, and neurodegenerative disorders over time. *Front Neurol*. 11, 1–10. doi: 10.3389/fneur.2020.616764

Ficiarà, E., Boschi, S., Ansari, S., D'Agata, F., Abollino, O., Caroppo, P., et al. (2021). Machine learning profiling of Alzheimer's disease patients based on current cerebrospinal fluid markers and iron content in biofluids. *Front. Aging Neurosci*. 13:607858. doi: 10.3389/fnagi.2021.607858

Fillenbaum, G. G., van Belle, G., Morris, J. C., Mohs, R. C., Mirra, S. S., Davis, P. C., et al. (2008). Consortium to establish a registry for Alzheimer's disease (CERAD): the first twenty years. *Alzheimer's and Dementia*. 4, 96–109. doi: 10.1016/j.jalz.2007.08.005

Foran, A., Mathias, J., and Bowden, S. (2021). Effectiveness of sorting tests for detecting cognitive decline in older adults with dementia and other common neurodegenerative disorders: A meta-analysis. *Neurosci Biobehav. Rev*. 120, 442–454. doi: 10.1016/j.neubiorev.2020.10.013

Foxe, D., Cheung, S. C., Cordato, N. J., Burrell, J. R., Ahmed, R. M., Taylor-Rubin, C., et al. (2021). Verbal short-term memory disturbance in the primary progressive aphasias: challenges and distinctions in a clinical setting. *Brain Sci*. 11, 1060. doi: 10.3390/brainsci11081060

Gamberger, D., Ženko, B., Mitelpunkt, A., Shachar, N., and Lavrač, N. (2016). Clusters of male and female Alzheimer's disease patients in the Alzheimer's disease neuroimaging initiative (ADNI) database. *Brain Inform*. 3, 169–179. doi: 10.1007/s40708-016-0035-5

Garn, H., Coronel, C., Waser, M., Caravias, G., and Ransmayr, G. (2017). Differential diagnosis between patients with probable Alzheimer's disease, Parkinson's disease dementia, or dementia with Lewy bodies and frontotemporal dementia, behavioral variant, using quantitative electroencephalographic features. *J. Neural. Transm. (Vienna)*. 124, 569–581. doi: 10.1007/s00702-017-1699-6

Gil-Navarro, S., Lladó, A., Rami, L., Castellví, M., Bosch, B., Bargalló, N., et al. (2013). Neuroimaging and biochemical markers in the three variants of primary progressive Aphasia. *Dement Geriatr. Cogn. Disord*. 35, 106–117. doi: 10.1159/000346289

Gorno-Tempini, M. L., Brambati, S. M., Ginex, V., Ogar, J., Dronkers, N. F., Marcone, A., et al. (2008). The logopenic/phonological variant of primary progressive aphasia. *Neurology*. 71, 1227–1234. doi: 10.1212/01.wnl.0000320506.79811.da

Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., et al. (2011). Classification of primary progressive aphasia and its variants. *Neurology*. 76, 1006–1014. doi: 10.1212/WNL.0b013e31821103e6

Gossink, F., Schouws, S., Krudop, W., Scheltens, P., Stek, M., Pijnenburg, Y., et al. (2018). Social cognition differentiates behavioral variant frontotemporal dementia from other neurodegenerative diseases and psychiatric disorders. *Am. J. Geriatric Psychiatr*. 26, 569–579. doi: 10.1016/j.jagp.2017.12.008

Grossman, M. (2010). Primary progressive aphasia: clinicopathological correlations. *Nat Rev Neurol*. 6, 88–97. doi: 10.1038/nrneurol.2009.216

Habes, M., Grothe, M. J., Tunc, B., McMillan, C., Wolk, D. A., Davatzikos, C., et al. (2020). Disentangling heterogeneity in Alzheimer's disease and related dementias using data-driven methods. *Biol. Psychiatry*. 88, 70–82. doi: 10.1016/j.biopsych.2020.01.016

Harris, J. M., Gall, C., Thompson, J. C., Richardson, A. M. T., Neary, D., du Plessis, D., et al. (2013). Classification and pathology of primary progressive aphasia. *Neurology*. 81, 1832–1839. doi: 10.1212/01.wnl.0000436070.28137.7b

Harris, J. M., Saxon, J. A., Jones, M., Snowden, J. S., and Thompson, J. C. (2019). Neuropsychological differentiation of progressive aphasic disorders. *J. Neuropsychol*. 13, 214–239. doi: 10.1111/jnp.12149

Hartigan, J. A., and Wong, M. A. (1979). Algorithm AS 136: a k-means clustering algorithm. *J. Royal Stat. Society*. 28, 100–108. doi: 10.2307/2346830

Henry, M., and Grasso, S. (2018). Assessment of individuals with primary progressive aphasia. *Semin Speech Lang*. 39, 231–241. doi: 10.1055/s-0038-1660782

Hinterbuchinger, B., Kaltenboeck, A., Baumgartner, J. S., Mossaheb, N., and Friedrich, F. (2018). Do patients with different psychiatric disorders show altered social decision-making? A systematic review of ultimatum game experiments in clinical populations. *Cogn. Neuropsychiatry*. 23, 117–141. doi: 10.1080/13546805.2018.1453791

Hoffman, P., Sajjadi, S. A., Patterson, K., and Nestor, P. J. (2017). Data-driven classification of patients with primary progressive aphasia. *Brain Lang*. 174, 86–93. doi: 10.1016/j.bandl.2017.08.001

Hornberger, M., and Piguet, O. (2012). Episodic memory in frontotemporal dementia: a critical review. *Brain*. 135, 678–692. doi: 10.1093/brain/aws011

Hüper, L., Steinacker, P., Polyakova, M., Mueller, K., Godulla, J., Herzig, S., et al. (2024). Neurofilaments and progranulin are related to atrophy in frontotemporal lobar degeneration – A transdiagnostic study cross-validating atrophy and fluid biomarkers. *Alzheimer. Dementia*. 20, 4461–4475. doi: 10.1002/alz.13863

Illán-Gala, I., Casaletto, K. B., Borrego-Écija, S., Arenaza-Urquijo, E. M., Wolf, A., Cobigo, Y., et al. (2021). Sex differences in the behavioral variant of frontotemporal dementia: A new window to executive and behavioral reserve. *Alzheimers Dement*. 17, 1329–1341. doi: 10.1002/alz.12299

Irwin, D. J., Trojanowski, J. Q., and Grossman, M. (2013). Cerebrospinal fluid biomarkers for differentiation of frontotemporal lobar degeneration from Alzheimer's disease. *Front Aging Neurosci*. 5:6. doi: 10.3389/fnagi.2013.00006

Johnen, A., and Bertoux, M. (2019). Psychological and cognitive markers of behavioral variant frontotemporal dementia–a clinical neuropsychologist's view on diagnostic criteria and beyond. *Front Neurol*. 10, 1–24. doi: 10.3389/fneur.2019.00594

Josephs, K. A., Whitwell, J. L., Knopman, D. S., Boeve, B. F., Vemuri, P., Senjem, M. L., et al. (2009). Two distinct subtypes of right temporal variant frontotemporal dementia. *Neurology*. 73, 1443–1450. doi: 10.1212/WNL.0b013e3181bf9945

Kamath, V., Chaney, G,-A. S., DeRight, J., and Onyike, C. U. (2019). A meta-analysis of neuropsychological, social cognitive, and olfactory functioning in the behavioral and language variants of frontotemporal dementia. *Psychol. Med*. 49, 2669–2680. doi: 10.1017/S0033291718003604

Kamath, V., Sutherland, E. R., and Chaney, G- A. A. (2020). Meta-analysis of neuropsychological functioning in the logopenic variant of primary progressive aphasia: comparison with the semantic and non-fluent variants. *J. Int. Neuropsychol. Society*. 26, 322–330. doi: 10.1017/S1355617719001115

Kim, Y. J., Cho, S.-K., Kim, H. J., Lee, J. S., Lee, J., Jang, Y. K., et al. (2019). Data-driven prognostic features of cognitive trajectories in patients with amnestic mild cognitive impairments. *Alzheimers Res. Ther*. 11, 10. doi: 10.1186/s13195-018-0462-z

Knopman, D. S., Kramer, J. H., Boeve, B. F., Caselli, R. J., Graff-Radford, N. R., Mendez, M. F., et al. (2008). Development of methodology for conducting clinical trials in frontotemporal lobar degeneration. *Brain*. 131, 2957–2968. doi: 10.1093/brain/awn234

Kramer, J. H., Jurik, J., Sha, S. J., Rankin, K. P., Rosen, H. J., Johnson, J. K., et al. (2003). Distinctive neuropsychological patterns in frontotemporal dementia, semantic dementia, and Alzheimer disease. *Cognitiv. Behav. Neurol*. 16, 211–218. doi: 10.1097/00146965-200312000-00002

Lampe, L., Huppertz, H,-J., Anderl-Straub, S., Albrecht, F., Ballarini, T., Bisenius, S., et al. (2023). Multiclass prediction of different dementia syndromes based on multi-centric volumetric MRI imaging. *Neuroimage Clin*. 37:103320. doi: 10.1016/j.nicl.2023.103320

Lampe, L., Niehaus, S., Huppertz, H,-J., Merola, A., Reinelt, J., Mueller, K., et al. (2022). Comparative analysis of machine learning algorithms for multi-syndrome classification of neurodegenerative syndromes. *Alzheimers Res Ther*. 14:62. doi: 10.21203/rs.3.rs-178965/v1

Leyton, C. E., Ballard, K. J., Piguet, O., and Hodges, J. R. (2014). Phonologic errors as a clinical marker of the logopenic variant of PPA. *Neurology*. 82, 1620–1627. doi: 10.1212/WNL.0000000000000387

Lin, C,-H., Chiu, S,-I., Chen, T,-F., Jang, J,-S. R., and Chiu, M,-J. (2020). Classifications of neurodegenerative disorders using a multiplex blood biomarkers-based machine learning model. *Int J Mol Sci*. 21:6914. doi: 10.3390/ijms21186914

Macoir, J., Légaré, A., and Lavoie, M. (2021). Contribution of the cognitive approach to language assessment to the differential diagnosis of primary progressive aphasia. *Brain Sci*. 11:815. doi: 10.3390/brainsci11060815

Maruta, C., Pereira, T., Madeira, S. C., De Mendonça, A., and Guerreiro, M. (2015). Classification of primary progressive aphasia: Do unsupervised data mining methods support a logopenic variant? *Amyotroph. Lateral Scler Frontotemporal. Degener*. 16, 147–159. doi: 10.3109/21678421.2015.1026266

Matias-Guiu, J. A., Díaz-Alvarez, J., Cuetos, F., Cabrera-Martín, M. N., Segovia-Ríos, I., Pytel, V., et al. (2019). Machine learning in the clinical and language characterisation of primary progressive aphasia variants. *Cortex*. 119, 312–323. doi: doi: 10.1016/j.cortex.2019.05.007

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., et al. (2011). 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. *Alzheimer's and Dementia*. 7, 263–269. doi: 10.1016/j.jalz.2011.03.005

Moms, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., et al. (1989). The consortium to establish a registry for Alzheimer's disease (CERAD). Part I. Clinical and neuropsychological assesment of Alzheimer's disease. *Neurology*. 39, 1159–1159. doi: 10.1212/WNL.39.9.1159

Morris, J. C. (1993). The clinical dementia rating (CDR) Current version and scoring rules. *Neurology*. 43, 2412.2-2412-a. doi: 10.1212/WNL.43.11.2412-a

Murley, A. G., Coyle-Gilchrist, I., Rouse, M. A., Jones, P. S., Li, W., Wiggins, J., et al. (2020). Redefining the multidimensional clinical phenotypes of frontotemporal lobar degeneration syndromes. *Brain*. 143, 1555–1571. doi: 10.1093/brain/awaa097

Musa, G., Slachevsky, A., Muñoz-Neira, C., Méndez-Orellana, C., Villagra, R., González-Billault, C., et al. (2020). Alzheimer's disease or behavioral variant frontotemporal dementia? review of key points toward an accurate clinical and neuropsychological diagnosis. *J. Alzheimer's Dis*. 73, 833–848. doi: 10.3233/JAD-190924

Otto, M., Ludolph, A. C., Landwehrmeyer, B., Förstl, H., Diehl-Schmid, J., Neumann, M., et al. (2011). Konsortium zur Erforschung der frontotemporalen Lobärdegeneration. *Nervenarzt*. 82, 1002–1005. doi: 10.1007/s00115-011-3261-3

Overbeek, J. M., Korten, N., Gossink, F., Fieldhouse, J., van de Beek, M., Reus, L., et al. (2020). The value of neuropsychological assessment in the differentiation between behavioral variant frontotemporal dementia and late-onset psychiatric disorders. *J. Clin. Psychiatry*. 81, 1–8. doi: 10.4088/JCP.19m12811

Owens, T. E., Machulda, M. M., Duffy, J. R., Strand, E. A., Clark, H. M., Boland, S., et al. (2018). Patterns of neuropsychological dysfunction and cortical volume changes in logopenic aphasia. *J. Alzheimer's Dis*. 66, 1015–1025. doi: 10.3233/JAD-171175

Piguet, O., Hornberger, M., Mioshi, E., and Hodges, J. R. (2011). Behavioural-variant frontotemporal dementia: diagnosis, clinical staging, and management. *Lancet Neurol*. 10, 162–172. doi: 10.1016/S1474-4422(10)70299-4

Ranasinghe, K. G., Rankin, K. P., Pressman, P. S., Perry, D. C., Lobach, I. V., Seeley, W. W., et al. (2016). Distinct subtypes of behavioral variant frontotemporal dementia based on patterns of network degeneration. *JAMA Neurol*. 73:1078. doi: 10.1001/jamaneurol.2016.2016

Rankin, K. P. (2021). Measuring behavior and social cognition in FTLD. *Adv. Exp. Med. Biol*. 51–65. doi: 10.1007/978-3-030-51140-1_4

Rankin, K. P., Santos-Modesitt, W., Kramer, J. H., Pavlic, D., Beckman, V., Miller, B. L., et al. (2008). Spontaneous social behaviors discriminate behavioral dementias from psychiatric disorders and other dementias. *J. Clin. Psychiatry*. 69, 60–73. doi: 10.4088/JCP.v69n0109

Rascovsky, K., and Grossman, M. (2013). Clinical diagnostic criteria and classification controversies in frontotemporal lobar degeneration. *Int. Rev. Psychiatry*. 25, 145–158. doi: 10.3109/09540261.2013.763341

Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., et al. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*. 134, 2456–2477. doi: 10.1093/brain/awr179

Reul, S., Lohmann, H., Wiendl, H., Duning, T., and Johnen, A. (2017). Can cognitive assessment really discriminate early stages of Alzheimer's and behavioural variant frontotemporal dementia at initial clinical presentation? *Alzheimers Res. Ther*. 9:61. doi: 10.1186/s13195-017-0287-1

Rogalski, E., Rademaker, A., and Weintraub, S. (2007). Primary progressive aphasia: relationship between gender and severity of language impairment. *Cognitiv. Behav. Neurol*. 20, 38–43. doi: 10.1097/WNN.0b013e31802e3bae

RStudio Team (2020). *RStudio: Integrated Development for R*. Boston, MA: RStudio, PBC.

Sachdev, P. S., Blacker, D., Blazer, D. G., Ganguli, M., Jeste, D. V., Paulsen, J. S., et al. (2014). Classifying neurocognitive disorders: the DSM-5 approach. *Nat Rev Neurol*. 10, 634–642. doi: 10.1038/nrneurol.2014.181

Sajjadi, S. A., Patterson, K., Arnold, R. J., Watson, P. C., and Nestor, P. J. (2012). Primary progressive aphasia: a tale of two syndromes and the rest. *Neurology*. 78, 1670–1677. doi: 10.1212/WNL.0b013e3182574f79

Schild, A., Meiberth, D., Frommann, I., Wagner, M., Belz, M., Bartels, C., et al. (2023). Neuropsychologische Demenzdiagnostik: Umfrage zum klinischen Standardvorgehen im Rahmen des Deutschen Netzwerks Gedächtnisambulanzen e. V. *Zeitschrift Für Neuropsychologie*. 34, 183–202. doi: 10.1024/1016-264X/a000382

Schroeter, M. L., Laird, A. R., Chwiesko, C., Deuschl, C., Schneider, E., Bzdok, D., et al. (2014). Conceptualizing neuropsychiatric diseases with multimodal data-driven meta-analyses–The case of behavioral variant frontotemporal dementia. *Cortex*. 57, 22–37. doi: 10.1016/j.cortex.2014.02.022

Schroeter, M. L., Pawelke, S., Bisenius, S., Kynast, J., Schuemberg, K., Polyakova, M., et al. (2018). A modified reading the mind in the eyes test predicts behavioral variant frontotemporal dementia better than executive function tests. *Front. Aging Neurosci*. 10, 1–11. doi: 10.3389/fnagi.2018.00011

Schroeter, M. L., Vogt, B., Frisch, S., Becker, G., Barthel, H., Mueller, K., et al. (2012). Executive deficits are related to the inferior frontal junction in early dementia. *Brain*. 135, 201–215. doi: 10.1093/brain/awr311

Sebastian, R., Thompson, C. B., Wang, N,-Y., Wright, A., Meyer, A., Friedman, R. B., et al. (2018). Patterns of decline in naming and semantic knowledge in primary progressive aphasia. *Aphasiology*. 32, 1010–1030. doi: 10.1080/02687038.2018.1490388

Seckin, M., Ricard, I., Raiser, T., Heitkamp, N., Ebert, A., Prix, C., et al. (2022). Utility of the repeat and point test for subtyping patients with primary progressive aphasia. *Alzheimer Dis. Assoc. Disord*. 1–8. doi: 10.1097/WAD.0000000000000482

van Buuren, S. (2019). *Flexible Imputation of Missing Data, Second Edition*. Second Edn. Boca, Raton, FL : CRC Press. | Chapman and Hall/CRC (2018). doi: 10.1201/9780429492259

van Buuren, S., Groothuis-Oudshoorn K. (2011). MICE : multivariate imputation by chained equations in R. *J. Stat. Softw*. 45, 1–67. doi: 10.18637/jss.v045.i03

van Dyck, C. H., Swanson, C. J., Aisen, P., Bateman, R. J., Chen, C., Gee, M., et al. (2023). Lecanemab in Early Alzheimer's Disease. *New England Journal of Medicine*. 388, 9–21. doi: 10.1056/NEJMoa2212948

van Vliet, D., de Vugt, M. E., Bakker, C., Pijnenburg, Y. A. L., Vernooij-Dassen, M. J. F. J., Koopmans, R. T. C. M., et al. (2013). Time to diagnosis in young-onset dementia as compared with late-onset dementia. *Psychol Med*. 43, 423–432. doi: 10.1017/S0033291712001122

von Luxburg, U., Williamson, R. C., and Guyon, I. (2012). Clustering: science or art? *JMLR Workshop Conf Proc*. 27:6579.

Walker, A. J., Meares, S., Sachdev, P. S., and Brodaty, H. (2005). The differentiation of mild frontotemporal dementia from Alzheimer's disease and healthy aging by neuropsychological tests. *Int. Psychogeriatr*. 17, 57–68. doi: 10.1017/S1041610204000778

Warren, J. D., Rohrer, J. D., and Rossor, M. N. (2013). Frontotemporal dementia. *BMJ*. 347, 1–9. doi: 10.1136/bmj.f4827

Weder, N. D., Aziz, R., Wilkins, K., and Tampi, R. R. (2007). Frontotemporal dementias: a review. *Ann Gen Psychiatry*. 6:15. doi: 10.1186/1744-859X-6-15

Woolley, J. D., Khan, B. K., Murthy, N. K., Miller, B. L., and Rankin, K. P. (2011). The diagnostic challenge of psychiatric symptoms in neurodegenerative disease. *J. Clin. Psychiatry*. 72, 126–133. doi: 10.4088/JCP.10m06382oli

Yanase, J., and Triantaphyllou, E. A. (2019). systematic survey of computer-aided diagnosis in medicine: past and present developments. *Expert Syst. Appl*. 138:112821. doi: 10.1016/j.eswa.2019.112821

Young, A. L., Marinescu, R. V., Oxtoby, N. P., Bocchetta, M., Yong, K., Firth, N. C., et al. (2018). Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with subtype and stage inference. *Nat Commun*. 9, 4273. doi: 10.1038/s41467-018-05892-0

Zhou, K., and Yang, S. (2020). Effect of cluster size distribution on clustering: a comparative study of k-means and fuzzy c-means clustering. *Pattern Analy. Appl*. 23, 455–466. doi: 10.1007/s10044-019-00783-6

Zimmerer, V. C., Hardy, C. J. D. D., Eastman, J., Dutta, S., Varnet, L., Bond, R. L., et al. (2020). Automated profiling of spontaneous speech in primary progressive aphasia and behavioral-variant frontotemporal dementia: an approach based on usage-frequency. *Cortex*. 133, 103–119. doi: 10.1016/j.cortex.2020.08.027