11-2020

Distinct Clinicopathologic Clusters of Persons with TDP-43 Proteinopathy

Yuriko Katsumata
University of Kentucky, katsumata.yuriko@uky.edu

Erin L. Abner
University of Kentucky, erin.abner@uky.edu

Shama Karanth
University of Kentucky, shama.karanth@uky.edu

Merilee A. Teylan
University of Washington

Charles N. Mock
University of Washington

See next page for additional authors

Distinct Clinicopathologic Clusters of Persons with TDP-43 Proteinopathy

Digital Object Identifier (DOI)
https://doi.org/10.1007/s00401-020-02211-0

Notes/Citation Information
Published in *Acta Neuropathologica*, v. 140, issue 5.

Authors
Yuriko Katsumata, Erin L. Abner, Shama Karanth, Merilee A. Teylan, Charles N. Mock, Matthew D. Cykowski, Edward B. Lee, Kevin L. Boehme, Shubhabrata Mukherjee, John S. K. Kauwe, Richard J. Kryscio, Frederick A. Schmitt, David W. Fardo, and Peter T. Nelson

HHS Public Access

Author manuscript

Acta Neuropathol. Author manuscript; available in PMC 2020 November 01.

Published in final edited form as:

Acta Neuropathol. 2020 November ; 140(5): 659–674. doi:10.1007/s00401-020-02211-0.

Distinct clinicopathologic clusters of persons with TDP-43 proteinopathy

Yuriko Katsumata, PhD,
Department of Biostatistics, University of Kentucky, Lexington, KY 40536, USA

Sanders-Brown Center on Aging, University of Kentucky, Lexington, KY 40536, USA

Erin L. Abner, PhD,
Department of Epidemiology, University of Kentucky, Lexington, KY 40536, USA

Sanders-Brown Center on Aging, University of Kentucky, Lexington, KY 40536, USA

Shama Karanth, MDS,
Department of Epidemiology, University of Kentucky, Lexington, KY 40536, USA

Merilee A. Teylan, MPH,
National Alzheimer's Coordinating Center, Department of Epidemiology, University of Washington, Seattle, WA 98105, USA

Charles N. Mock, MD, PhD,
Department of Epidemiology, University of Washington, Seattle, WA 98105, USA

National Alzheimer's Coordinating Center, Department of Epidemiology, University of Washington, Seattle, WA 98105, USA

Matthew D. Cykowski, MD,
Department of Pathology and Genomic Medicine, Houston Methodist Hospital, Houston, TX 77030, USA

Edward B Lee, MD, PhD,
Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

Kevin L. Boehme, BS,
Department of Biology, Brigham Young University, Provo, UT 84602, USA

Shubhabrata Mukherjee, PhD,
Department of Medicine, University of Washington, Seattle, WA 98104, USA

Corresponding Author: Peter T. Nelson, MD, PhD, Department of Pathology, University of Kentucky, Rm 311 Sanders-Brown Center on Aging, 800 S. Limestone Avenue, Lexington, KY 40536, USA, TEL: +1-859-218-3862, peter.nelson@uky.edu.

Author contributions

YK was responsible for study conception, data analysis and interpretation, and drafting the manuscript. PTN was responsible for study conception, supervision, data interpretation, and drafting the manuscript. ELA was responsible for study conception, data interpretation, and drafting the manuscript. All other authors were involved in critical revision of the manuscript for important intellectual content.

Compliance with ethical standards

The authors declare that they have no conflict of interest.

John S.K. Kauwe, PhD,
Brigham Young University–Hawaii, Laie, HI 96762, USA

Biology Department, Brigham Young University, Provo, UT 84602, USA

Richard J. Kryscio, PhD,
Department of Statistics, University of Kentucky, Lexington, KY 40536, USA

Sanders-Brown Center on Aging, University of Kentucky, Lexington, KY 40536, USA

Frederick A. Schmitt, PhD,
Department of Neurology, University of Kentucky, Lexington, KY 40536, USA

Sanders-Brown Center on Aging, University of Kentucky, Lexington, KY 40536, USA

David W. Fardo, PhD,
Department of Biostatistics, University of Kentucky, Lexington, KY 40536, USA

Sanders-Brown Center on Aging, University of Kentucky, Lexington, KY 40536, USA

Peter T. Nelson, MD, PhD
Department of Pathology, University of Kentucky, Lexington, KY 40536, USA

Sanders-Brown Center on Aging, University of Kentucky, Lexington, KY 40536, USA

Abstract

To better understand clinical and neuropathological features of TDP-43 proteinopathies, data were analyzed from autopsied research volunteers who were followed in the National Alzheimer's Coordinating Center (NACC) data set. All subjects (n=495) had autopsy-proven TDP-43 proteinopathy as an inclusion criterion. Subjects underwent comprehensive longitudinal clinical evaluations yearly for 6.9 years before death on average. We tested whether an unsupervised clustering algorithm could detect coherent groups of TDP-43 immunopositive cases based on age at death and extensive neuropathologic data. Although many of the brains had mixed pathologies, four discernible clusters were identified. Key differentiating features were age at death and the severity of comorbid Alzheimer's disease neuropathologic changes (ADNC), particularly neuritic amyloid plaque densities. Cluster 1 contained mostly cases with a pathologic diagnosis of frontotemporal lobar degeneration (FTLD-TDP), consistent with enrichment of frontotemporal dementia clinical phenotypes including appetite/eating problems, disinhibition and primary progressive aphasia (PPA). Cluster 2 consisted of elderly limbic-predominant age-related TDP-43 encephalopathy (LATE-NC) subjects without severe neuritic amyloid plaques. Subjects in Cluster 2 had a relatively slow cognitive decline. Subjects in both Clusters 3 and 4 had severe ADNC + LATE-NC; however, Cluster 4 was distinguished by earlier disease onset, swifter disease course, more Lewy body pathology, less neocortical TDP-43 proteinopathy, and a suggestive trend in a subgroup analysis (n=114) for increased *C9orf72* risk SNP rs3849942 T allele (Fisher's exact test p-value = 0.095). Overall, clusters enriched with neocortical TDP-43 proteinopathy (Clusters 1 and 2) tended to have lower levels of neuritic amyloid plaques, and those dying older (Clusters 2 and 3) had far less PPA or disinhibition, but more apathy. Indeed, 98% of subjects dying past age 85 years lacked clinical features of the frontotemporal dementia syndrome. Our study revealed discernible subtypes of LATE-NC and underscored the importance of age of death for differentiating FTLD-TDP and LATE-NC.

Keywords

hallucinations; psychosis; anxiety; FTD; trajectory; DLB

Introduction

TAR-DNA binding protein 43kDa (TDP-43) proteinopathy is present in up to 50% of brains in advanced age and has a strong association with cognitive impairment [39]. A working group recently suggested a classification system for limbic-predominant age-related TDP-43 encephalopathy neuropathologic changes (LATE-NC) [39]. The terminology is parallel with the current classification of Alzheimer's disease (AD), where the neuropathologic changes are termed ADNC [33]. Cross-sectional data have been interpreted to indicate that TDP-43 proteinopathy occurs in a stereotypic, hierarchical spatiotemporal pattern in the brain [23, 36]. In this hypothetical schema, which is the basis for the proposed neuropathologic staging of LATE-NC, TDP-43 deposition first appears in the amygdala, then the hippocampal formation, and in ~15% of the elderly, it may develop in frontal neocortex and other brain structures [22, 23, 36].

There is some controversy in the field as to the specific definitions and distinguishing features of LATE-NC, ADNC, and frontotemporal lobar degeneration with TDP-43 proteinopathy (FTLD-TDP) [6, 21]. Compelling data indicate that some parallel or synergistic mechanisms occur in ADNC and LATE-NC, because the two pathologies are frequently comorbid [24]. However, many severe ADNC cases lack LATE-NC, and LATE-NC can occur without ADNC (~75% of aged brains harbor ADNC, with or without comorbid LATE-NC) [39]. Further, the presence of LATE-NC is clinicopathologically impactful—ADNC + LATE-NC has a more severe clinical phenotype than ADNC without LATE-NC [38]. Thus, the LATE Working Group suggested that the diagnosis of LATE-NC be applied whether or not comorbid ADNC is present. In LATE-NC cases that lack severe ADNC, there is an open question about how LATE-NC is differentiated from FTLD-TDP. There are clear epidemiologic differences – LATE-NC is far more common than FTLD-TDP and affects older persons than FTLD-TDP [39]. Yet more work is required to generate criteria to differentiate between these two conditions.

A key question is whether clinical and pathologic features occur in predictable patterns in aged persons' brains that would enable differentiation between LATE-NC, ADNC, and FTLD-TDP, or, alternatively, if the brain pathologies seem to occur in more random patterns, which would render classification problematic. These questions will become more important if future disease-specific therapeutic strategies are developed. Furthermore, the clinical features of cases with LATE-NC but lacking ADNC remain to be well described.

Here we examined if distinct and clinically relevant groups of TDP-43 cases could be identified using unsupervised clustering. We were interested to see how the pathology-defined groups associated with clinical course and disease manifestations including neuropsychiatric symptoms (NPS). Standardized data from research participants who died and consented to autopsy at Alzheimer's Disease Centers (ADCs) were compiled by the National Alzheimer's Coordinating Center (NACC). Cases with TDP-43 immunoreactive

inclusions were analyzed along with other detailed data. Before coming to autopsy, research volunteers were followed longitudinally with detailed clinical visits, that allowed us to compare the cognitive and NPS trajectories between persons grouped by clustering algorithms that were based on age at death and neuropathologic findings.

Materials and methods

Participants

Thirty-eight United States (U.S.) ADCs contributed data through the March 2020 data freeze (https://www.alz.washington.edu/). Participants were excluded if at least one of 19 rare brain diseases were diagnosed (Supplementary Table 1) and were included if TDP-43 pathology in at least one brain region was observed (Supplementary Table 2). Autopsies were performed within each of the contributory ADCs. Research activities at individual ADCs were approved by their local Institutional Review Boards (IRB). Informed consent was obtained from all participants at the individual ADCs. No additional IRB approval was needed for this secondary analysis of de-identified data.

Neuropathologies for clustering

For clustering, we investigated neurodegenerative disease-associated neuropathologies including amyloid-β (Aβ) plaques, tau neurofibrillary tangles (NFTs), and α-synuclein (α-syn) along with TDP-43 and autopsy-confirmed cerebrovascular pathologies. Data on brain region-specific TDP-43-immunoreactive inclusions were collected with response categories “no”, “yes”, “not assessed”, and “missing/unknown”. Each ADC used either phospho-specific or non-phospho-specific antibodies as described in detail previously [26].

The recently proposed LATE-NC staging system characterizes the anatomical distribution TDP-43 proteinopathy based on three brain regions: amygdala, hippocampus, and neocortex [36, 39]. Accordingly, we included participants who had data on TDP-43 pathology in at least one of these brain regions in the subsequent analyses. For ADNC, we used the consensus “A, B, C” system [33]. Tau neurofibrillary degeneration was represented by ADNC Braak NFT stage categories B score: B0 = stage 0 (none), B1 = stage I or II, B2 = stage III or IV, and B3 = stage V or VI [10] and Aβ plaques were represented by ADNC neocortical neuritic plaque density ratings (C score: C0 = none, C1 = sparse, C2 = moderate, and C3 = frequent) [30], and ADNC Thal phase ratings for Aβ distribution (A score: Thal Aβ phase A0 to A5) [48]. As a proxy of severity of α-syn pathology, we used Lewy body pathology data with three categories: none; present in neocortical region; or present in non-neocortical regions including brainstem, limbic, amygdala, and olfactory bulb. For cerebrovascular pathologies, data were available on cerebral amyloid angiopathy (none, mild, moderate, or severe), large infarcts or lacunar infarcts (no or yes), microinfarcts (no or yes), and arteriolosclerosis (none, mild, moderate, or severe). These parameters were described elsewhere [9] and the data dictionary is publicly available at https://www.alz.washington.edu/WEB/forms_np.html.

Other neuropathologies

Hippocampal sclerosis (HS) was determined by the variable of "hippocampal sclerosis of CA1 and/or subiculum (NPHIPSCL)" with four response categories: none, unilateral, bilateral, or present but laterality not assessed. We dichotomized the variable by collapsing unilateral, bilateral, and present but laterality not assessed. Presence of FTLD-TDP was determined by the variable NPFTDTDP (FTLD with TDP-43 pathology).

Cognitive tests

Cognitive data were drawn from the NACC Uniform Data Set (UDS) [51]. Mini Mental State Examination (MMSE) [17], and Montreal Cognitive Assessment (MoCA) [37] for global function, verbal fluency (Animal and Vegetable Naming) [34] for language/fluency function, Wechsler Memory Scale-Revised (WMS-R) Logical Memory – immediate and delayed [50] and Craft Story 21 Recall – immediate and delayed [12] for memory function (Supplementary Table 3). Since the MoCA and Craft Story 21 Recall – immediate and delayed were introduced in the NACC UDS version 3 from March 2015 (neuropsychological battery – form C2) instead of MMSE and WMS-R Logical Memory – immediate and delayed (neuropsychological battery – form C1), respectively, we transformed the new battery scores into equivalent old battery scores based on Monsell and colleagues' crosswalk study [32]. We also included the Clinical Dementia Rating Scale (CDR) Sum of Boxes ratings with the cognitive test measures.

Neuropsychiatric symptoms (NPS) and primary progressive aphasia (PPA)

Neuropsychiatric symptoms (NPS) were measured in the UDS using the Neuropsychiatric Inventory (NPI-Q) [13]. Study co-participants (defined as someone who knows the participant well, usually a caregiver for persons with dementia) were asked if the following specific NPS were present in the past month prior to the study visit: delusions, hallucinations, agitation or aggression, depression/dysphoria, anxiety, elation/euphoria, apathy/indifference, disinhibition, irritability/lability, motor disturbance, nighttime behaviors, and appetite and eating problems. Primary progressive aphasia (PPA) was evaluated by clinicians (the variable name in UDS: NACCPPA) with "no" or "yes" rating categories in participants with cognitive impairment.

Genetics

Genetic data were obtained from Alzheimer's Disease Genetics Consortium (ADGC), which were linked to clinical and neuropathological outcome data from the NACC data set. These data were from SNP platforms, and did not include tandem repeat expansion data. Moreover, those data were oriented toward downstream "AD vs non-AD" studies, and did not include substantial numbers of cases that were previously diagnosed as FTLD-TDP. We examined five putative risk single nucleotide polymorphisms (SNPs) that reported to be associated with TDP-43 related disease including FTLD-TDP, ALS, and HS: rs9637454 in *KCNMB2* located on chromosome 3q [8], rs1990622 in *TMEM106B* on chromosome 7p [5, 35, 44], rs3849942 in *C9orf72* on chromosome 9p associated with increased hexanucleotide GGGGCC repeats [14, 20], rs704180 in *ABCC9* on chromosome 12p [40], and rs5848 in *GRN* on chromosome 17q [15, 35, 41, 43].

Statistical analysis

We first performed clustering using uniform manifold approximation and projection (UMAP), which is a nonlinear dimensionality reduction technique to model high-dimensional data in a lower-dimensional space. Using data on the proteinopathies, cerebrovascular disease, and categorized age at death (< 65 years old (y.o.), ≥ 65 and < 85 y.o., and ≥ 85 y.o.) data, similar cases were clustered as low-dimensional representations that closely match the topological structure of the data. The ages of 65 and 85 as cutoffs were selected as in prior studies of "early-onset" (< 65 y.o.) and "late-onset" (> 85 y.o.) disease, and we confirmed that around 25% cases who were clinically diagnosed as FTD died by age of 65 years and that almost all cases with clinical FTD and/or PPA died by age of 85 years (Fig. 1a). Each of the variables was dummy-coded (Supplementary Table 4). We used the Python package for UMAP available at https://umap-learn.readthedocs.io/en/latest/ [29]. We set the parameters of UMAP: metric = "dice" (for binary data), number of neighbors = 10, and number of components = 2 (i.e., embedded into two dimensions). We calculated McFadden's pseudo coefficients of determination (R^2) [28] and Akaike information criterion (AIC) to evaluate the contribution of each variable to clustering using a multinomial logistic regression model with the cluster category as the outcome. This was implemented with the "PseudoR2" function in the DescTools R package [45]. We also quantified the relative importance of variables for clustering assignments based on the conditional mean decreases in accuracy using a random forest and bagging ensemble algorithm. This approach takes into account correlations between variables using the "cforest" function in the party R package [18, 46, 47].

After creating the clusters, we conducted cross-sectional and longitudinal analyses for cognitive functions and NPS/PPA. For cross-sectional analysis, we retrieved cognitive test scores (continuous) and NPS/PPA (binary) data measured at the clinical last visit within three years of death, and then we performed pairwise comparisons in the means (for cognitive test scores) and the proportions (for NPS/PPA) between the clusters generated by UMAP technique. In modelling the longitudinal change in cognitive test scores over years, we used a non-linear mixed effects regression model with a logistic function [27] implemented in lme4 R package [7] to take into account variability between and within subjects and floor and ceiling effects on the scores. To examine whether the clusters were associated with development of NPS/PPA, we constructed unadjusted Kaplan-Meier curves for each of the symptoms and performed pairwise log rank tests between the curves of the clusters. All p-values from pairwise comparisons were corrected by Bonferroni-Holm. All statistical analyses were performed with R version 3.6.1 [42].

Results

After applying exclusion and inclusion criteria, 514 autopsied participants had TDP-43 pathology in at least one of the regions: amygdala, hippocampus, and neocortex. Of these, 495 participants, who had no missing data on proteinopathies and cerebrovascular diseases as shown in Fig. 1b, were included in the UMAP dimensionality reduction and clustering analyses (Supplementary Fig. 1). Among the included subjects, mean age at death was 80.8 years (standard deviation (SD) = 10.3), 52.7% were women, and mean years of education

was 15.6 (SD = 3.1) (Supplementary Table 5). As expected, TDP-43 proteinopathy was most commonly observed in the amygdala (89.4% of cases), followed by hippocampus (84.3%) and neocortex (31.2%). ADNC pathology was also common: 60.0% of cases had frequent neuritic plaques (C3), 71.9% had Thal phase 5 (amyloid plaques in cerebellum; A3), and 65.7% had severe NFTs (B3). Neocortical Lewy bodies were less common (12.7%). FTLD-TDP was diagnosed in 21.6% of included cases, and 35.2% showed HS at autopsy (Supplementary Table 6).

We derived two UMAP embeddings for visualization, and four clusters were identified (Fig. 1c). We conducted univariate multinomial regression with cluster membership (1–4) as the categorical outcome. McFadden's pseudo R^2 and AIC of each of the variables used in the UMAP are shown in Supplementary Table 7, and their relative importance for clustering are shown in Fig 1d. The most contributed variables for clustering were age at death and neuritic plaques (C score), whereas cerebrovascular diseases and Lewy body pathology had less contribution to clustering. The outcomes distributions are shown in Fig. 2. The full distribution plots for the other variables used in the UMAP approach and are in Supplementary Fig. 3 and Table 1. We confirmed the robustness of the clustering using five times repeated random subsampling validation of UMAP with 80% of the analytic participants (i.e., n = 495). As shown in Supplementary Fig. 2, the UMAP consistently reproduced the distinct clusters and thus was quite robust.

Cluster 1 included 103 participants (red colored points in Fig. 1c). The majority of participants in Cluster 1 died between 65 and 85 years of age (70.9%), with no neuritic amyloid plaques (74.8%), no or mild NFT burden (25.2% for B0 and 58.3% for B1), no Lewy body pathology (85.4%), and no cerebral amyloid angiopathy (86.4%). Cluster 2 included 71 participants (blue colored points in Fig. 1c). In Cluster 2, 91.5% had sparse or moderate neuritic plaques (26.8% for sparse (C1) and 64.7% for moderate (C2)), 88.8% had moderate or severe NFT burden (46.5% for B2 or 42.3% for B3), and 77.5% had no Lewy body pathology. In both Clusters 3 and 4, more than 90% had frequent neuritic plaques (C3), and ~90% had severe NFT burden (B3). Almost all subjects in Cluster 3 died at 85 years or older, while the majority in Cluster 4 died between 65 and 85 years.

More than 90% in both Cluster 3 and 4 had TDP-43 in amygdala and more than 90% in both Cluster 1 and 2 had TDP-43 in hippocampus. The majority of participants in Cluster 1 had TDP-43 pathology in neocortex (76.7%) (Table 3 and Fig. 2b). We note that 77.8% in Cluster 1 did not have the *APOE* ε4 allele, whereas 72.7% in Cluster 4 had at least on *APOE* ε4 allele. The majority of Cluster 1 participants were diagnosed as FTLD-TDP (76.7%) and had lobar atrophy (70.6%) at autopsy. Although HS was present in 35.2% of the overall cases, there was no statistically significant difference in HS prevalence among the clusters (Table 2 and 3). Clustering distributions for HS, FTLD-TDP, and *APOE* ε4 allele were visualized and shown in Supplementary Fig. 4.

Next, we investigated cross-sectional differences and longitudinal changes over time in cognitive test scores. Supplementary Table 8 shows the crude mean and Fig. 3a and Supplementary Fig. 5 shows boxplots for cognitive test scores measured at the last visit within three years before death. Cluster 4 had the lowest scores for all the cognitive tests,

and global function (i.e., CDR Sum of Boxes and MMSE) scores were significantly lower than those in other clusters. Memory (as measured by Logical Memory – immediate and delayed) scores in Cluster 1 were higher than in other clusters. The scores of both Naming tests were higher in Cluster 1 compared to Cluster 4.

In participants who were diagnosed with cognitive impairment (impaired not MCI, MCI, or dementia) during the follow-up, non-linear mixed effects regression modelling was used to assess the longitudinal trajectory of change before and after the first diagnosis of cognitive impairment (Fig. 3b and Supplementary Fig. 6). Results showed that declines in cognitive function were underway prior to the first diagnosis of impairment. The declines were faster in Cluster 1, especially for MMSE and Naming tests.

Supplementary Table 9 and Supplementary Fig. 7 show the proportion of each of the neuropsychiatric symptoms observed at the last visit within three years before death by the clusters. The proportions of delusions and hallucinations were lower in Cluster 1 than in Cluster 4. Cluster 3 had a lower proportion of hallucinations than Cluster 4. The proportions of appetite and eating problems (a frequent symptom in ALS/FTD spectrum disorders [2]) and PPA in Cluster 1 were higher than in the other clusters. In terms of PPA subtypes, these data were mostly missing, so relatively few cases had specified subtypes of PPA. Among the 9 cases specified to have exhibited semantic subtype of PPA, 7 were in Cluster 1, whereas among the 7 cases with logopenic subtype of PPA, 6 were in Cluster 4 (data not shown).

We further compared the probability of symptom-free survival among the clusters using two approaches regarding the time scale: age at visit until age at death and years since first diagnosed with cognitive impairment. We used the variable “NACCUDSD,” which codes the clinical syndromic diagnosis (normal, impaired not MCI, MCI, or dementia) at each UDS visit to identify whether and for how long the participants had cognitive impairment. Fig. 4 and Supplementary Fig. 8 show Kaplan-Meier curves for each of the symptoms and the Bonferroni-Holm adjusted p-values for pairwise comparison between the clusters in the first approach (i.e., the x-axis indicates age at visit until age at death). The Kaplan- Meier curves for time to first appearance of delusions, hallucinations, agitation or aggression, elation or euphoria, disinhibition, irritability or lability, and appetite and eating problems in Cluster 4 were significantly different from those in other clusters. In addition to these differences, the Kaplan- Meier curves of agitation or aggression, elation or euphoria, disinhibition, irritability or lability, and appetite and eating problems in Cluster 1 were significantly different from those in Clusters 2 and 3. The significant differences between Clusters 2 and 3 were seen in anxiety, apathy or indifference, and nighttime behavior. All pairwise comparisons were significant in motor disturbance.

Fig. 5 and Supplementary Fig. 9 show Kaplan-Meier curves for each of the symptoms and the Bonferroni-Holm adjusted p-values for pairwise comparison between the clusters in the second approach to the timescale (i.e., the x-axis indicates years since first diagnosed as cognitive impairment). The Kaplan-Meier curve for hallucinations in Cluster 4 was significantly different from that in other clusters which was a similar result with the first approach. On the other hand, significant differences between Cluster 4 and Clusters 2 and 3 were not seen in delusions. There were significant differences between Cluster 1 and the

other clusters in agitation or aggression, elation or euphoria, disinhibition, motor disturbance, appetite and eating problems, and PPA. Only anxiety and apathy or indifference showed significant differences between Clusters 2 and 3.

Finally, we examined the genetic associations with the clusters as shown in Supplementary Table 10. Of 495 included participants with TDP-43 pathology, a total of 114 had ADGC genotype data. Although we observed no significant genetic association with the clusters because of the small sample size, there was suggestive association in a subgroup analysis for more risk allele of *C9orf72* SNP (rs3849942) in Cluster 4 compared to other clusters (two-tailed Fisher's exact test p-value was 0.095) (Supplementary Fig. 10).

Discussion

The clinical correlates of neuropathologically-defined groups were assessed among longitudinally followed research participants who had autopsy-proven TDP-43 proteinopathy (n = 495). Using a dimension reduction technique, four discernible clusters were resolved based on detailed neuropathology and age at death. Findings in the four clusters of cases are summarized in Fig. 6 and Table 4. These results indicated that there are neuropathologically and clinically differentiable subsets of persons with age-related TDP-43 proteinopathy.

Cluster 1 was highly enriched for cases that were clinically diagnosed with FTD syndrome, and were ultimately given a pathological diagnosis FTLD-TDP (Table 3). The participants in this cluster died at younger age and the majority had predominantly TDP-43 pathology without significant ADNC or Lewy body pathology (Fig. 6a). These patients also had a higher prevalence of appetite and eating problems, disinhibition, and PPA compared to other clusters (Fig. 6b). The association of appetite and eating problems in this cluster is intriguing given the association of these behaviors with FTD variants more so than in clinical AD [1, 19]. Cluster 2 was enriched for cases with lower neuritic amyloid plaques, like Cluster 1, but did not show clinical features of FTD syndrome. The participants in Cluster 2 died at an older age and showed a more gradual cognitive decline.

Two clusters (Clusters 3 and 4) were enriched for severe comorbid ADNC (Table 1). Cluster 4 showed a younger age of symptom onset and death. These patients tended to have more severe pathologies, including ADNC and Lewy body disease (Fig. 6a), and more NPS, especially higher proportion of hallucinations (Fig. 6b). We analyzed whether the higher proportion of hallucinations in Cluster 4 was attributable to neocortical Lewy body pathology -- there was not a statistically significant association between hallucinations and Lewy body pathology in Cluster 4 (17% with no Lewy body pathology reported hallucination, versus 27% with Lewy body pathology, p=0.38), although this conclusion was limited somewhat by sample size. The lack of a strong association between Lewy bodies and hallucinations in the Cluster 4 group may reflect a greater cortical and limbic density of neurofibrillary tangles, rather than differences in Lewy bodies, or overall Braak stage, as tangle density has been shown to associate with earlier onset of hallucinations in both AD and AD with co-morbid Lewy body disease [16].

Our study design incorporated longitudinal clinical data from both before and after the onset of cognitive impairment. As expected, subjects in a FTLD-TDP-enriched case cluster showed clinical features of motor disturbance, disinhibition, apathy, and eating/appetite problems. There was also a trend for this cluster to be associated with PPA and more broadly-defined language dysfunction. The findings of the clinical features of the syndrome of FTD in Cluster 1 are reassuring with regard to the validity of the clustering results, because we did not factor in FTD (clinical) or FTLD (final pathological diagnosis) into our clustering algorithms. By contrast, the common forms of LATE-NC, as in Clusters 2–4, were not associated with a FTD clinical phenotype.

An important and controversial topic area is in how LATE-NC overlaps with, and is different from, FTLD-TDP and ADNC. Clinical and epidemiologic features are often used to help discriminate between different conditions, although a given disease (e.g., brain infarcts) can manifest clinically in a variety of ways. The clinical and pathological criteria that can definitely discriminate LATE-NC from FTLD-TDP have not yet been developed [39]. Here we found, in line with previous studies [39], that age of death was a differentiating factor. Among persons dying after age 85 years, 98% of persons with TDP-43 proteinopathy lacked FTD clinical syndrome or PPA. Diagnosing FTLD-TDP at autopsy for persons in this age group may therefore cause confusion for clinicians and family members. Otherwise, the generation of pathology-based criteria to differentiate FTLD-TDP from LATE-NC will probably require tools that are sharper than the parameters currently available in the NACC NP data set.

Differentiating LATE-NC and ADNC cases is in a sense easier, because the presence and severity of ADNC are defined independently of TDP-43 proteinopathy. However, age-related TDP-43 proteinopathy is often accompanied by comorbid ADNC [24], may be a part of the AD neuropathologic spectrum [49], and/or may “reflect impaired cellular function in end-stage neurodegeneration” [21]. Further, a recently published study showed in a community-based cohort that there was a common neuropathologic phenotype with comorbid Tau, Aβ, TDP-43 and Lewy body pathologies, corresponding with a relatively aggressive disease course [25]. Here, we found that Cluster 4 cases were indeed enriched for subjects with comorbid Tau, Aβ, TDP-43 and Lewy body pathologies. Further, this cluster had a swift disease course and also had a trend for enrichment in the *C9orf72* risk allele in a subset analysis of 114 cases. This risk allele is associated with increased *C9orf72* repeats including intermediate repeats which have been recently associated with corticobasal degeneration and altered autophagic flux [11]. Whether this risk allele is truly associated with Cluster 4 requires additional analysis of larger cohorts.

Extensive data were factored into our clustering algorithm, including all “A” (Aβ Thal stages), “B” (Braak stages), and “C” (CERAD neuritic amyloid plaques, or NP densities) parameters. The clustering algorithm implicated NP densities as a key differentiating parameter. NPs were described by Alois Alzheimer as silver-impregnated “miliary foci” in the historic case report of Auguste Dieter [3] and remain a pathognomonic disease feature. In the present study, NP densities were a strong driver for the clustering algorithm, helping to differentiate Clusters 1 and 2 from the more ADNC-enriched Clusters 3 and 4. Cluster 2, for example, had no subject with high NP densities, although 84.5% of subjects in Cluster 2

had been given a diagnosis of Probable AD at final clinical examination. Cluster 2 also was enriched for neocortical TDP-43 proteinopathy, but longitudinal follow-up indicated a less severe clinical course and more subjects with final clinical diagnosis of MCI instead of dementia. This was an unexpected clue for neuropathologists about the importance of NP densities in aged brains. It also may provide clinicians with relevant information about clinical features associated with a lower burden of ADNC and a higher amount of TDP-43 proteinopathy. Such an outcome is what is hoped for from a clustering algorithm, elucidating patterns that would not necessarily be revealed by descriptive statistics alone. Collectively, these observations highlight clinical-pathologic patterns that may help guide future refinements of diagnostic classification.

There are some limitations in our study, related to the NACC data set [9]. The ADCs tend to recruit highly educated Caucasian/white people; therefore, the results should be interpreted with caution when generalizing to other populations. Many of the ADC cohorts recruit from dementia clinics, and thus are highly enriched for FTLD cases, with lower numbers of LATE-NC. Individual ADCs may apply exclusion criteria related to mental illness, substance abuse, physical disability, or other prevalent conditions that decrease the number of autopsied participants and limit the generalizability of our results. The lack of methodologic standardization between the ADCs in terms of TDP-43 IHC methods and data collection at the clinical visits may have affected our results, as previously discussed [26]. There also was not statistical power to show subtle differences in the clinical features of FTLD-TDP and LATE-NC; the unequal sample sizes for these groups affects the power to detect differences. An additional limitation is that direct *C9ORF72* repeat expansion data were not available for further analysis. Indeed, the observed trend for the *C9orf72* associated SNP in Cluster 4 should be interpreted with caution as this polymorphism is associated with a Finnish haplotype block [31] and so future genetic analyses in larger cohorts are required to adjust for confounds associated with uneven population structures.

Although HS is often comorbid with LATE-NC [4, 39], HS was not a focus of the present study. There are no widely applied consensus-based criteria for HS neuropathologic diagnosis, which is problematic for a study where cases were worked up at dozens of different research centers. Moreover, HS is a nonspecific pathologic endpoint that is neither necessary nor sufficient for the diagnosis of LATE-NC [39].

Key strengths of the current study are the large number of cognitive and non-cognitive domains tested, the longitudinal assessments prior to death, and the state-of-the-art neuropathologic assessments (all 2014 and after) performed at high quality academic research centers. Further, the sample sizes of FTLD-TDP and LATE-NC cases lacking ADNC are relatively large for a study with longitudinal clinical assessments and autopsy confirmation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are extremely grateful to the many patients, clinicians, and other colleagues, who have worked so hard to provide and organize these data. NIH grants R01 NS118584, P30 AG028383, R01 AG057187, R56 AG057191, and R21 AG061551 helped fund the project. The NACC database is funded by NIA/NIH Grant U01 AG016976. NACC data are contributed by the NIA-funded ADCs: P30 AG019610 (PI Eric Reiman, MD), P30 AG013846 (PI Neil Kowall, MD), P30 AG062428-01 (PI James Leverenz, MD) P50 AG008702 (PI Scott Small, MD), P50 AG025688 (PI Allan Levey, MD, PhD), P50 AG047266 (PI Todd Golde, MD, PhD), P30 AG010133 (PI Andrew Saykin, PsyD), P50 AG005146 (PI Marilyn Albert, PhD), P30 AG062421-01 (PI Bradley Hyman, MD, PhD), P30 AG062422-01 (PI Ronald Petersen, MD, PhD), P50 AG005138 (PI Mary Sano, PhD), P30 AG008051 (PI Thomas Wisniewski, MD), P30 AG013854 (PI Robert Vassar, PhD), P30 AG008017 (PI Jeffrey Kaye, MD), P30 AG010161 (PI David Bennett, MD), P50 AG047366 (PI Victor Henderson, MD, MS), P30 AG010129 (PI Charles DeCarli, MD), P50 AG016573 (PI Frank LaFerla, PhD), P30 AG062429-01(PI James Brewer, MD, PhD), P50 AG023501 (PI Bruce Miller, MD), P30 AG035982 (PI Russell Swerdlow, MD), P30 AG028383 (PI Linda Van Eldik, PhD), P30 AG053760 (PI Henry Paulson, MD, PhD), P30 AG010124 (PI John Trojanowski, MD, PhD), P50 AG005133 (PI Oscar Lopez, MD), P50 AG005142 (PI Helena Chui, MD), P30 AG012300 (PI Roger Rosenberg, MD), P30 AG049638 (PI Suzanne Craft, PhD), P50 AG005136 (PI Thomas Grabowski, MD), P30 AG062715-01 (PI Sanjay Asthana, MD, FRCP), P50 AG005681 (PI John Morris, MD), P50 AG047270 (PI Stephen Strittmatter, MD, PhD). The National Institutes of Health, National Institute on Aging (NIH-NIA) supported this work through the following grants: ADGC, U01 AG032984, RC2 AG036528; Samples from the National Cell Repository for Alzheimer's Disease (NCRAD), which receives government support under a cooperative agreement grant (U24 AG21886) awarded by the National Institute on Aging (NIA), were used in this study. Data for this study were prepared, archived, and distributed by the National Institute on Aging Alzheimer's Disease Data Storage Site (NIAGADS) at the University of Pennsylvania (U24-AG041689-01); NACC, U01 AG016976; NIA LOAD (Columbia University), U24 AG026395, U24 AG026390, R01AG041797; Banner Sun Health Research Institute P30 AG019610; Boston University, P30 AG013846, U01 AG10483, R01 CA129769, R01 MH080295, R01 AG017173, R01 AG025259, R01 AG048927, R01AG33193, R01 AG009029; Columbia University, P50 AG008702, R37 AG015473, R01 AG037212, R01 AG028786; Duke University, P30 AG028377, AG05128; Emory University, AG025688; Group Health Research Institute, UO1 AG006781, UO1 HG004610, UO1 HG006375, U01 HG008657; Indiana University, P30 AG10133, R01 AG009956, RC2 AG036650; Johns Hopkins University, P50 AG005146, R01 AG020688; Massachusetts General Hospital, P50 AG005134; Mayo Clinic, P50 AG016574, R01 AG032990, KL2 RR024151; Mount Sinai School of Medicine, P50 AG005138, P01 AG002219; New York University, P30 AG08051, UL1 RR029893, 5R01AG012101, 5R01AG022374, 5R01AG013616, 1RC2AG036502, 1R01AG035137; North Carolina A&T University, P20 MD000546, R01 AG28786-01A1; Northwestern University, P30 AG013854; Oregon Health & Science University, P30 AG008017, R01 AG026916; Rush University, P30 AG010161, R01 AG019085, R01 AG15819, R01 AG17917, R01 AG030146, R01 AG01101, RC2 AG036650, R01 AG22018; TGen, R01 NS059873; University of Alabama at Birmingham, P50 AG016582; University of Arizona, R01 AG031581; University of California, Davis, P30 AG010129; University of California, Irvine, P50 AG016573; University of California, Los Angeles, P50 AG016570; University of California, San Diego, P50 AG005131; University of California, San Francisco, P50 AG023501, P01 AG019724; University of Michigan, P50 AG008671; University of Pennsylvania, P30 AG010124; University of Pittsburgh, P50 AG005133, AG030653, AG041718, AG07562, AG02365; University of Southern California, P50 AG005142; University of Texas Southwestern, P30 AG012300; University of Miami, R01 AG027944, AG010491, AG027944, AG021547, AG019757; University of Washington, P50 AG005136, R01 AG042437; University of Wisconsin, P50 AG033514; Vanderbilt University, R01 AG019085; and Washington University, P50 AG005681, P01 AG03991, P01 AG026276. The Kathleen Price Bryan Brain Bank at Duke University Medical Center is funded by NINDS grant # NS39764, NIMH MH60451 and by Glaxo Smith Kline. Support was also from the Alzheimer's Association (LAF, IIRG-08-89720; MP-V, IIRG-05-14147), the US Department of Veterans Affairs Administration, Office of Research and Development, Biomedical Laboratory Research Program, and BrightFocus Foundation (MP-V, A2111048). P.S.G.-H. is supported by Wellcome Trust, Howard Hughes Medical Institute, and the Canadian Institute of Health Research. Genotyping of the TGEN2 cohort was supported by Kronos Science. The TGen series was also funded by NIA grant AG041232 to AJM and MJH, The Banner Alzheimer's Foundation, The Johnnie B. Byrd Sr. Alzheimer's Institute, the Medical Research Council, and the state of Arizona and also includes samples from the following sites: Newcastle Brain Tissue Resource, MRC London Brain Bank for Neurodegenerative Diseases (funding via the Medical Research Council),South West Dementia Brain Bank (funding via numerous sources including the Higher Education Funding Council for England (HEFCE), Alzheimer's Research Trust (ART), BRACE as well as North Bristol NHS Trust Research and Innovation Department and DeNDRoN), The Netherlands Brain Bank (funding via numerous sources including Stichting MS Research, Brain Net Europe, Hersenstichting Nederland Breinbrekend Werk, International Parkinson Fonds, Internationale Stiching Alzheimer Onderzoek), Institut de Neuropatologia, Servei Anatomia Patologica, Universitat de Barcelona. ADNI data collection and sharing was funded by the National Institutes of Health Grant U01 AG024904 and Department of Defense award number W81XWH-12-2-0012. Tissue also was provided by the Newcastle Brain Tissue Resource, which is funded in part by a grant from the UK Medical Research Council (G0400074), by Brains for Dementia research, a joint venture between Alzheimer's Society and Alzheimer's Research UK and by the NIHR Newcastle Biomedical Research Centre awarded to the Newcastle upon Tyne Hospitals NHS Foundation Trust and Newcastle University. ADNI is funded by the NIA, the NIBIB, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery

Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the NIH. The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the AD Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. We thank Drs. Dallas Anderson and Marilyn Miller from NIA who are *ex-officio* ADGC members.

Funding

This work was study was funded by grants R56AG057191, R01AG057187, R21AG061551, R01AG054060, and the UK-ADC P30AG028383 from the National Institute on Aging.

References

1. Ahmed RM, Irish M, Henning E, Dermody N, Bartley L, Kiernan MC, Piguet O, Farooqi S, Hodges JR (2016) Assessment of Eating Behavior Disturbance and Associated Neural Networks in Frontotemporal Dementia. JAMA neurology 73: 282–290 Doi 10.1001/jamaneurol.2015.4478 [PubMed: 26810632]
2. Ahmed RM, Irish M, Kam J, van Keizerswaard J, Bartley L, Samaras K, Hodges JR, Piguet O (2014) Quantifying the eating abnormalities in frontotemporal dementia. JAMA neurology 71: 1540–1546 Doi 10.1001/jamaneurol.2014.1931 [PubMed: 25329012]
3. Alzheimer A (1907) Über eine eigenartige Erkrankung der Hirnrinde. Allgemeine Z Psychiatrie Psychisch-Gerichtliche Med 64: 146–148
4. Amador-Ortiz C, Lin WL, Ahmed Z, Personett D, Davies P, Duara R, Graff-Radford NR, Hutton ML, Dickson DW (2007) TDP-43 immunoreactivity in hippocampal sclerosis and Alzheimer's disease. Ann Neurol 61: 435–445 Doi 10.1002/ana.21154 [PubMed: 17469117]
5. Aoki N, Murray ME, Ogaki K, Fujioka S, Rutherford NJ, Rademakers R, Ross OA, Dickson DW (2015) Hippocampal sclerosis in Lewy body disease is a TDP-43 proteinopathy similar to FTLD-TDP Type A. Acta Neuropathol 129: 53–64 Doi 10.1007/s00401-014-1358-z [PubMed: 25367383]
6. Baborie A, Griffiths TD, Jaros E, Momeni P, McKeith IG, Burn DJ, Keir G, Larner AJ, Mann DM, Perry R (2012) Frontotemporal dementia in elderly individuals. Arch Neurol 69: 1052–1060 Doi 10.1001/archneurol.2011.3323 [PubMed: 22529248]
7. Bates D, Maechler M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4. J Stat Softw 67: 1–48
8. Beecham GW, Hamilton K, Naj AC, Martin ER, Huentelman M, Myers AJ, Corneveaux JJ, Hardy J, Vonsattel JP, Younkin SGet al. (2014) Genome-wide association meta-analysis of neuropathologic features of Alzheimer's disease and related dementias. PLoS Genet 10: e1004606 Doi 10.1371/journal.pgen.1004606 [PubMed: 25188341]
9. Besser LM, Kukull WA, Teylan MA, Bigio EH, Cairns NJ, Kofler JK, Montine TJ, Schneider JA, Nelson PT (2018) The Revised National Alzheimer's Coordinating Center's Neuropathology Form-Available Data and New Analyses. J Neuropathol Exp Neurol 77: 717–726 Doi 10.1093/jnen/nly049 [PubMed: 29945202]
10. Braak H, Braak E (1991) Alzheimer's disease affects limbic nuclei of the thalamus. Acta Neuropathol 81: 261–268 [PubMed: 1711755]
11. Cali CP, Patino M, Tai YK, Ho WY, McLean CA, Morris CM, Seeley WW, Miller BL, Gaig C, Vonsattel JPGet al. (2019) C9orf72 intermediate repeats are associated with corticobasal degeneration, increased C9orf72 expression and disruption of autophagy. Acta neuropathologica 138: 795–811 Doi 10.1007/s00401-019-02045-5 [PubMed: 31327044]
12. Craft S, Newcomer J, Kanne S, Dagogo-Jack S, Cryer P, Sheline Y, Luby J, Dagogo-Jack A, Alderson A (1996) Memory improvement following induced hyperinsulinemia in Alzheimer's disease. Neurobiol Aging 17: 123–130 Doi 10.1016/0197-4580(95)02002-0 [PubMed: 8786794]

13. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J (1994) The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurology 44: 2308–2314 Doi 10.1212/wnl.44.12.2308 [PubMed: 7991117]

14. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, Nicholson AM, Finch NA, Flynn H, Adamson J et al. (2011) Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron 72: 245–256 Doi 10.1016/j.neuron.2011.09.011 [PubMed: 21944778]

15. Dickson DW, Baker M, Rademakers R (2010) Common variant in GRN is a genetic risk factor for hippocampal sclerosis in the elderly. Neuro-degenerative diseases 7: 170–174 Doi 10.1159/000289231 [PubMed: 20197700]

16. Ferman TJ, Arvanitakis Z, Fujishiro H, Duara R, Parfitt F, Purdy M, Waters C, Barker W, Graff-Radford NR, Dickson DW (2013) Pathology and temporal onset of visual hallucinations, misperceptions and family misidentification distinguishes dementia with Lewy bodies from Alzheimer's disease. Parkinsonism Relat Disord 19: 227–231 Doi 10.1016/j.parkreldis.2012.10.013 [PubMed: 23182311]

17. Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12: 189–198 [PubMed: 1202204]

18. Hothorn T, Buhlmann P, Dudoit S, Molinaro A, van der Laan MJ (2006) Survival ensembles. Biostatistics 7: 355–373 Doi 10.1093/biostatistics/kxj011 [PubMed: 16344280]

19. Ikeda M, Brown J, Holland AJ, Fukuhara R, Hodges JR (2002) Changes in appetite, food preference, and eating habits in frontotemporal dementia and Alzheimer's disease. J Neurol Neurosurg Psychiatry 73: 371–376 Doi 10.1136/jnnp.73.4.371 [PubMed: 12235302]

20. Jones AR, Woollacott I, Shatunov A, Cooper-Knock J, Buchman V, Sproviero W, Smith B, Scott KM, Balendra R, Abel O et al. (2013) Residual association at C9orf72 suggests an alternative amyotrophic lateral sclerosis-causing hexanucleotide repeat. Neurobiol Aging 34: 2234 e2231–2237 Doi 10.1016/j.neurobiolaging.2013.03.003

21. Josephs KA, Mackenzie I, Frosch MP, Bigio EH, Neumann M, Arai T, Dugger BN, Ghetti B, Grossman M, Hasegawa M et al. (2019) LATE to the PART-y. Brain 142: e47 Doi 10.1093/brain/awz224 [PubMed: 31359030]

22. Josephs KA, Murray ME, Whitwell JL, Parisi JE, Petrucelli L, Jack CR, Petersen RC, Dickson DW (2014) Staging TDP-43 pathology in Alzheimer's disease. Acta Neuropathol 127: 441–450 Doi 10.1007/s00401-013-1211-9 [PubMed: 24240737]

23. Josephs KA, Murray ME, Whitwell JL, Tosakulwong N, Weigand SD, Petrucelli L, Liesinger AM, Petersen RC, Parisi JE, Dickson DW (2016) Updated TDP-43 in Alzheimer's disease staging scheme. Acta Neuropathol 131: 571–585 Doi 10.1007/s00401-016-1537-1 [PubMed: 26810071]

24. Josephs KA, Whitwell JL, Weigand SD, Murray ME, Tosakulwong N, Liesinger AM, Petrucelli L, Senjem ML, Knopman DS, Boeve BFet al. (2014) TDP-43 is a key player in the clinical features associated with Alzheimer's disease. Acta Neuropathol 127: 811–824 Doi 10.1007/s00401-014-1269-z [PubMed: 24659241]

25. Karanth S, Nelson PT, Katsumata Y, Kryscio RJ, Schmitt FA, Fardo DW, Cykowski MD, Jicha GA, Van Eldik LJ, Abner EL (2020) Prevalence and Clinical Phenotype of Quadruple Misfolded Proteins in Older Adults. JAMA neurology: Doi 10.1001/jamaneurol.2020.1741

26. Katsumata Y, Fardo DW, Kukull WA, Nelson PT (2018) Dichotomous scoring of TDP-43 proteinopathy from specific brain regions in 27 academic research centers: associations with Alzheimer's disease and cerebrovascular disease pathologies. Acta Neuropathol Commun 6: 142 Doi 10.1186/s40478-018-0641-y [PubMed: 30567576]

27. Martins CA, Oulhaj A, de Jager CA, Williams JH (2005) APOE alleles predict the rate of cognitive decline in Alzheimer disease: a nonlinear model. Neurology 65: 1888–1893 Doi 10.1212/01.wnl.0000188871.74093.12 [PubMed: 16380608]

28. McFadden D (1974) Conditional logit analysis of qualitative choice behavior In: Zarembka (ed) Frontiers in Ecnometrics. Academic press, City

29. Mclnnes L, Healy J (2018) UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv: preprint arXiv:1802.03426

30. Mirra SS, Heyman A, McKeel D, Sumi SM, Crain BJ, Brownlee LM, Vogel FS, Hughes JP, van Belle G, Berg L (1991) The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part II. Standardization of the neuropathologic assessment of Alzheimer's disease. Neurology 41: 479–486 [PubMed: 2011243]

31. Mok K, Traynor BJ, Schymick J, Tienari PJ, Laaksovirta H, Peuralinna T, Myllykangas L, Chio A, Shatunov A, Boeve BF et al. (2012) Chromosome 9 ALS and FTD locus is probably derived from a single founder. Neurobiol Aging 33: 209 e203–208 Doi 10.1016/j.neurobiolaging.2011.08.005

32. Monsell SE, Dodge HH, Zhou XH, Bu Y, Besser LM, Mock C, Hawes SE, Kukull WA, Weintraub S, Neuropsychology Work Group Advisory to the Clinical Task F (2016) Results From the NACC Uniform Data Set Neuropsychological Battery Crosswalk Study. Alzheimer Dis Assoc Disord 30: 134–139 Doi 10.1097/WAD.0000000000000111 [PubMed: 26485498]

33. Montine TJ, Phelps CH, Beach TG, Bigio EH, Cairns NJ, Dickson DW, Duyckaerts C, Frosch MP, Masliah E, Mirra SS et al. (2012) National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease: a practical approach. Acta Neuropathol 123: 1–11 Doi 10.1007/s00401-011-0910-3 [PubMed: 22101365]

34. Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, Mellits ED, Clark C (1989) The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. Neurology 39: 1159–1165 [PubMed: 2771064]

35. Murray ME, Cannon A, Graff-Radford NR, Liesinger AM, Rutherford NJ, Ross OA, Duara R, Carrasquillo MM, Rademakers R, Dickson DW (2014) Differential clinicopathologic and genetic features of late-onset amnestic dementias. Acta Neuropathol 128: 411–421 Doi 10.1007/s00401-014-1302-2 [PubMed: 24899141]

36. Nag S, Yu L, Boyle PA, Leurgans SE, Bennett DA, Schneider JA (2018) TDP-43 pathology in anterior temporal pole cortex in aging and Alzheimer's disease. Acta Neuropathol Commun 6: 33 Doi 10.1186/s40478-018-0531-3 [PubMed: 29716643]

37. Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H (2005) The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 53: 695–699 Doi 10.1111/j.1532-5415.2005.53221.x [PubMed: 15817019]

38. Nelson PT, Abner EL, Schmitt FA, Kryscio RJ, Jicha GA, Smith CD, Davis DG, Poduska JW, Patel E, Mendiondo MS et al. (2010) Modeling the association between 43 different clinical and pathological variables and the severity of cognitive impairment in a large autopsy cohort of elderly persons. Brain Pathol 20: 66–79 Doi 10.1111/j.1750-3639.2008.00244.x [PubMed: 19021630]

39. Nelson PT, Dickson DW, Trojanowski JQ, Jack CR, Boyle PA, Arfanakis K, Rademakers R, Alafuzoff I, Attems J, Brayne Cet al. (2019) Limbic-predominant age-related TDP-43 encephalopathy (LATE): consensus working group report. Brain 142: 1503–1527 Doi 10.1093/brain/awz099 [PubMed: 31039256]

40. Nelson PT, Estus S, Abner EL, Parikh I, Malik M, Neltner JH, Ighodaro E, Wang WX, Wilfred BR, Wang LS et al. (2014) ABCC9 gene polymorphism is associated with hippocampal sclerosis of aging pathology. Acta Neuropathol 127: 825–843 Doi 10.1007/s00401-014-1282-2 [PubMed: 24770881]

41. Pickering-Brown SM, Rollinson S, Du Plessis D, Morrison KE, Varma A, Richardson AM, Neary D, Snowden JS, Mann DM (2008) Frequency and clinical characteristics of progranulin mutation carriers in the Manchester frontotemporal lobar degeneration cohort: comparison with patients with MAPT and no known mutations. Brain 131: 721–731 Doi 10.1093/brain/awm331 [PubMed: 18192287]

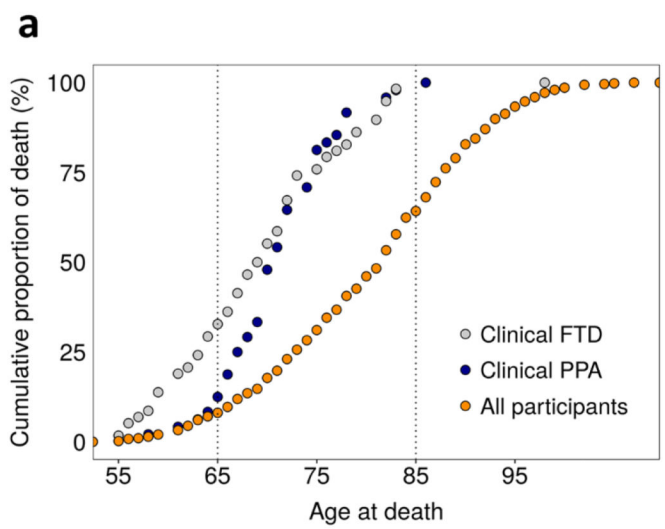
42. R Development Core Team (2013) R: A language and environment for statistical computing. 3.0.2 edn. R. Foundation for Statistical Computing, City

43. Rademakers R, Eriksen JL, Baker M, Robinson T, Ahmed Z, Lincoln SJ, Finch N, Rutherford NJ, Crook RJ, Josephs KAet al. (2008) Common variation in the miR-659 binding-site of GRN is a major risk factor for TDP43-positive frontotemporal dementia. Hum Mol Genet 17: 3631–3642 Doi 10.1093/hmg/ddn257 [PubMed: 18723524]

44. Rutherford NJ, Carrasquillo MM, Li M, Bisceglio G, Menke J, Josephs KA, Parisi JE, Petersen RC, Graff-Radford NR,Younkin SGet al. (2012) TMEM106B risk variant is implicated in the

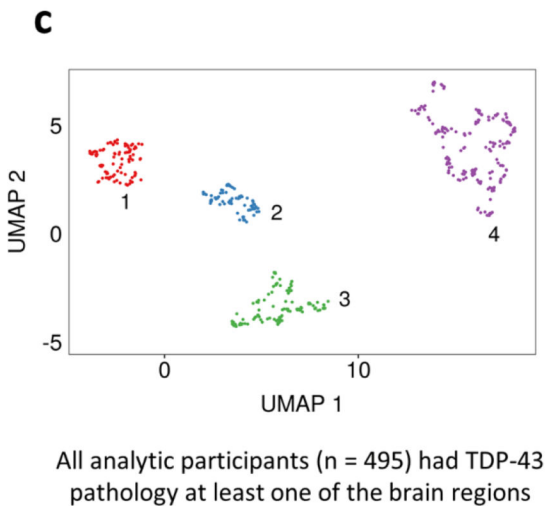
pathologic presentation of Alzheimer disease. Neurology 79: 717–718 Doi 10.1212/WNL.0b013e318264e3ac [PubMed: 22855871]

45. Signorell A et al. (2020) DescTools: tools for descriptive statistics R package version 0.99.36 City
46. Strobl C, Boulesteix AL, Kneib T, Augustin T, Zeileis A (2008) Conditional variable importance for random forests. BMC Bioinformatics 9: 307 Doi 10.1186/1471-2105-9-307 [PubMed: 18620558]
47. Strobl C, Boulesteix AL, Zeileis A, Hothorn T (2007) Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinformatics 8: 25 Doi 10.1186/1471-2105-8-25 [PubMed: 17254353]
48. Thal DR, Rub U, Orantes M, Braak H (2002) Phases of A beta-deposition in the human brain and its relevance for the development of AD. Neurology 58: 1791–1800 [PubMed: 12084879]
49. Tome SO, Vandenberghe R, Ospitalieri S, Van Schoor E, Tousseyn T, Otto M, von Arnim CAF, Thal DR (2020) Distinct molecular patterns of TDP-43 pathology in Alzheimer's disease: relationship with clinical phenotypes. Acta Neuropathol Commun 8: 61 Doi 10.1186/s40478-020-00934-5 [PubMed: 32349792]
50. Wechsler D (1987) Wechsler Memory Scale-Revised Psychological Corporation, City
51. Weintraub S, Salmon D, Mercaldo N, Ferris S, Graff-Radford NR, Chui H, Cummings J, DeCarli C, Foster NL, Galasko Det al. (2009) The Alzheimer's Disease Centers' Uniform Data Set (UDS): the neuropsychologic test battery. Alzheimer Dis Assoc Disord 23: 91–101 Doi 10.1097/WAD.0b013e318191c7dd [PubMed: 19474567]



b

Data used in UMAP dimensionality reduction and clustering
Age at death
Proteinopathy
Neuritic plaques (C score)
Thal phase (A score)
Braak NFT stage (B score)
Lewy body pathology
Cerebrovascular disease
Cerebral amyloid angiopathy
Infarct and lacunes
Microinfarcts
Arteriolosclerosis



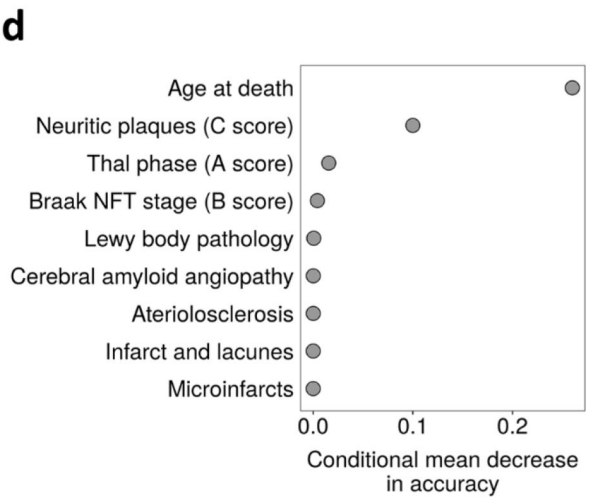


Fig. 1. A clustering analysis was performed with an algorithm based on age at death and neuropathologies, using data on subjects with autopsy-confirmed TDP-43 proteinopathy (n = 495).

As shown in panel **(a)**, ~98% of subjects with TDP-43 proteinopathy, who died after age 85 years, lacked clinical frontotemporal dementia (FTD) syndrome or primary progressive aphasia (PPA), so the 85 year old age cutoff was included in the clustering algorithm. The data used in the clustering are shown in panel (**b**). Uniform manifold approximation and projection (UMAP) visualization of clusters and relative importance of each variables for clustering was performed. Dimensionality reduction was performed based on age at death, proteinopathy, and cerebrovascular disease data as shown in panel (**c**). Numeric values (one to four) and colors (red, blue, green, purple) were assigned to each of the four clusters. The relative importance quantified by conditional mean decreases in accuracy which takes into account correlations between the variables shown in panel (**d**). NFT = neurofibrillary tangle

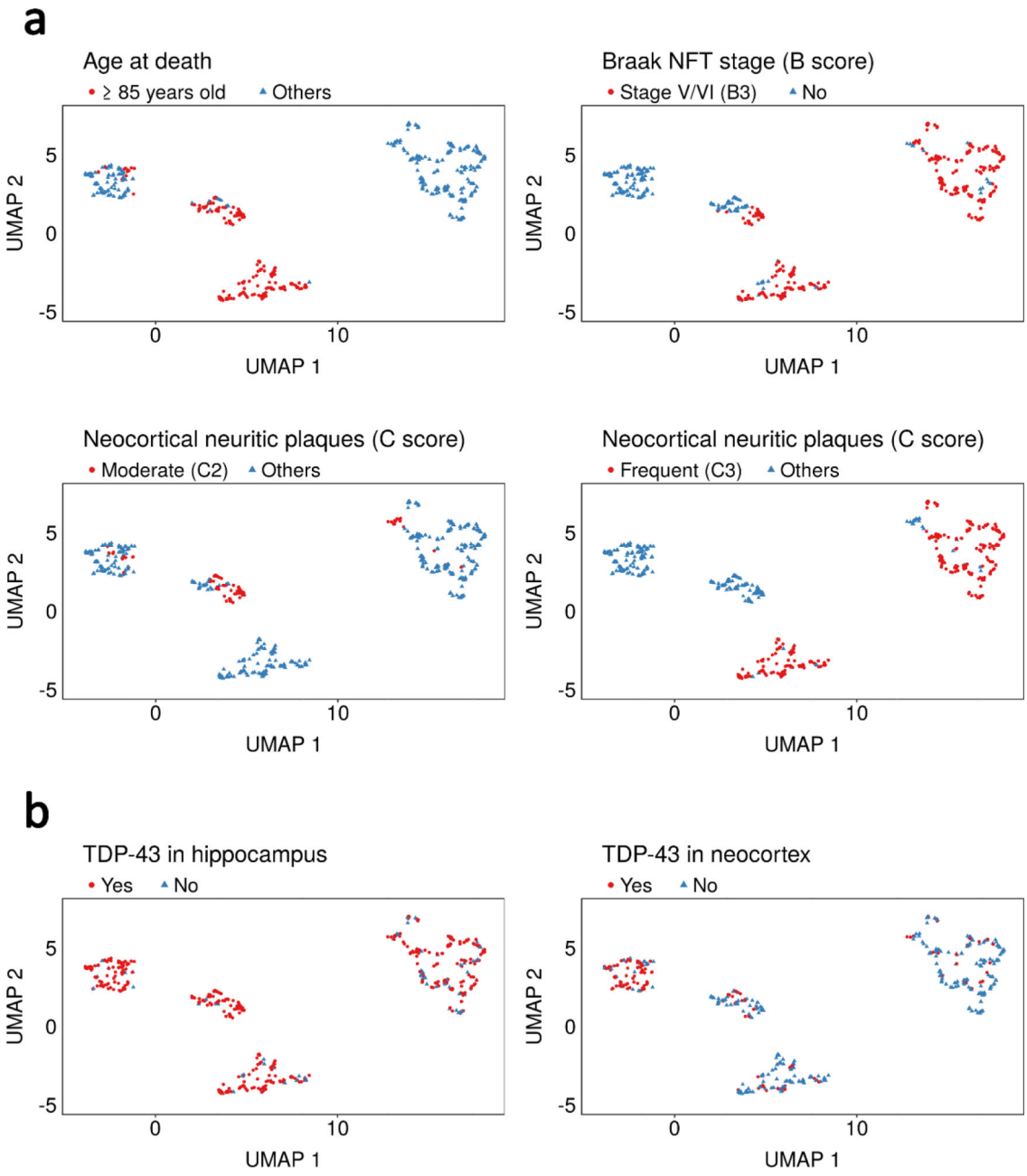


Fig. 2. Distributions to the clusters for age at death, ADNC, and TPD-43 proteinopathy anatomic regions.

Features for clustering included age, Braak stages (B0-B3), and neuritic plaque densities (C0-C3). Shown in panel (**a**) are distributions by age at death of ≥ 85 years, Braak NFT stage V/VI (B3), and moderate and frequent neocortical neuritic plaques (C2 and C3). Not included in the clustering algorithm, were the anatomic location of the TDP-43 proteinopathy (**b**). Note that a majority of cases had hippocampal TDP-43 pathology whereas neocortical TDP-43 pathology was enriched particularly in Cluster 1.

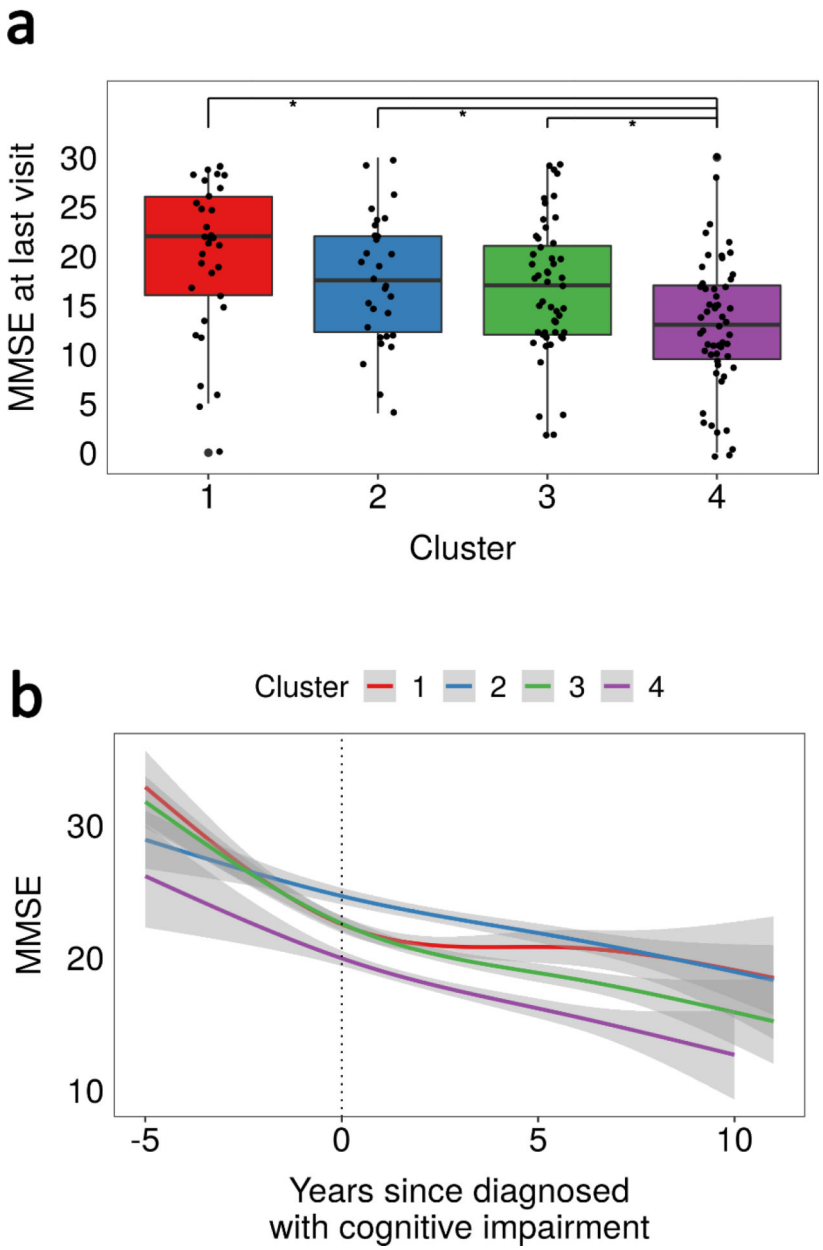


Fig. 3. Mini-Mental State Examination (MMSE) scores, stratifying by cluster status.

In panel (**a**), boxplots are shown for test scores at last visit by clusters. * indicates the significant pairwise comparison between cluster means based on the Bonferroni-Holm adjusted p-value of less than 0.05. In panel (**b**), longitudinal trajectories for MMSE test scores are shown. The x-axis indicates years since first diagnosed with cognitive impairment, that is, x = 0 represents the year when initially diagnosed as either impaired not MCI, MCI, or dementia based on the variable "NACCUDSD" in NACC UDS. Negative values of the x-axis represent the number of years before the first diagnosis of cognitive impairment.

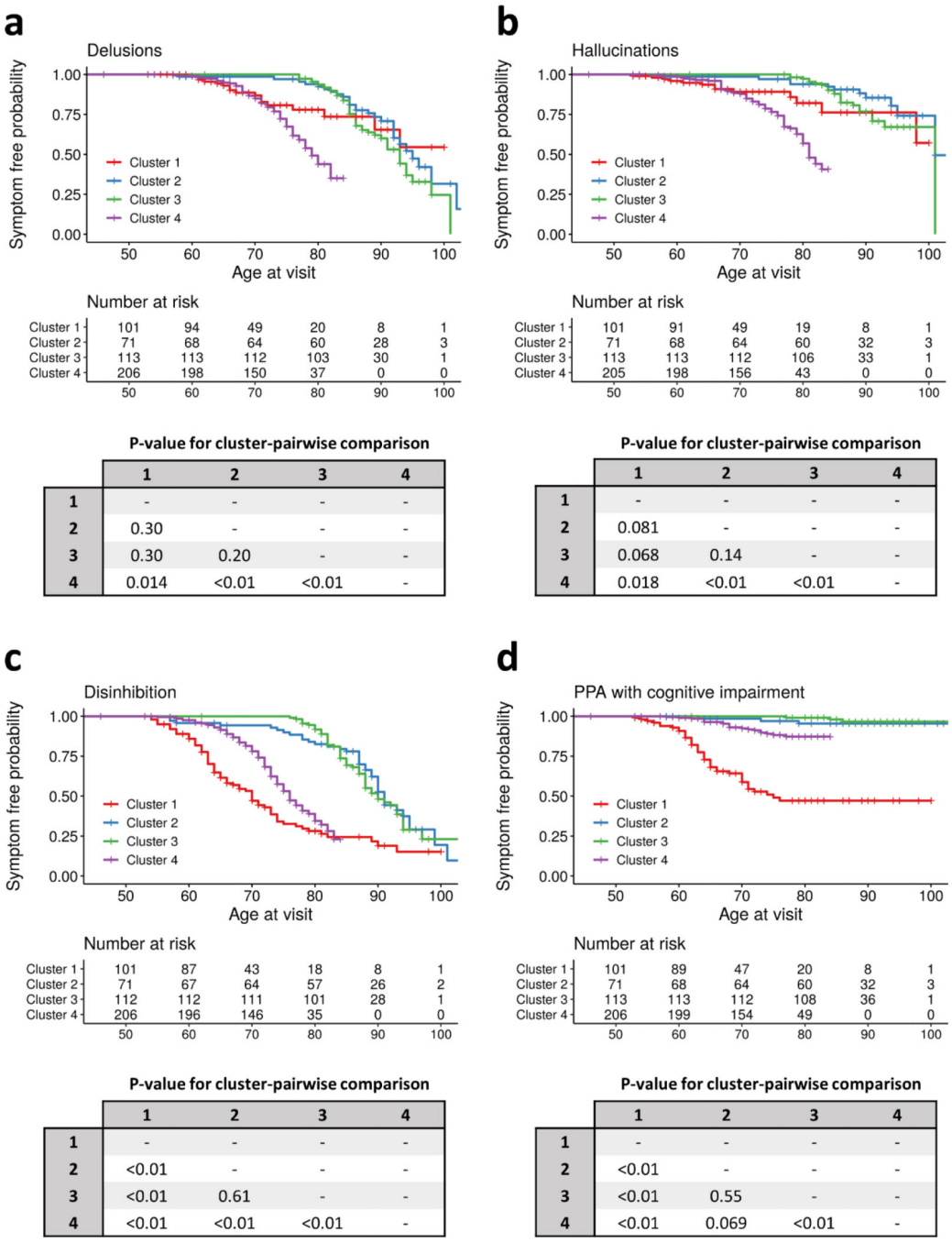


Fig. 4. Kaplan-Meier curves and Bonferroni-Holm adjusted p-values from pairwise log rank test for age at visit.

The y-axis shows the symptom free probability and the x-axis indicates age at visit.

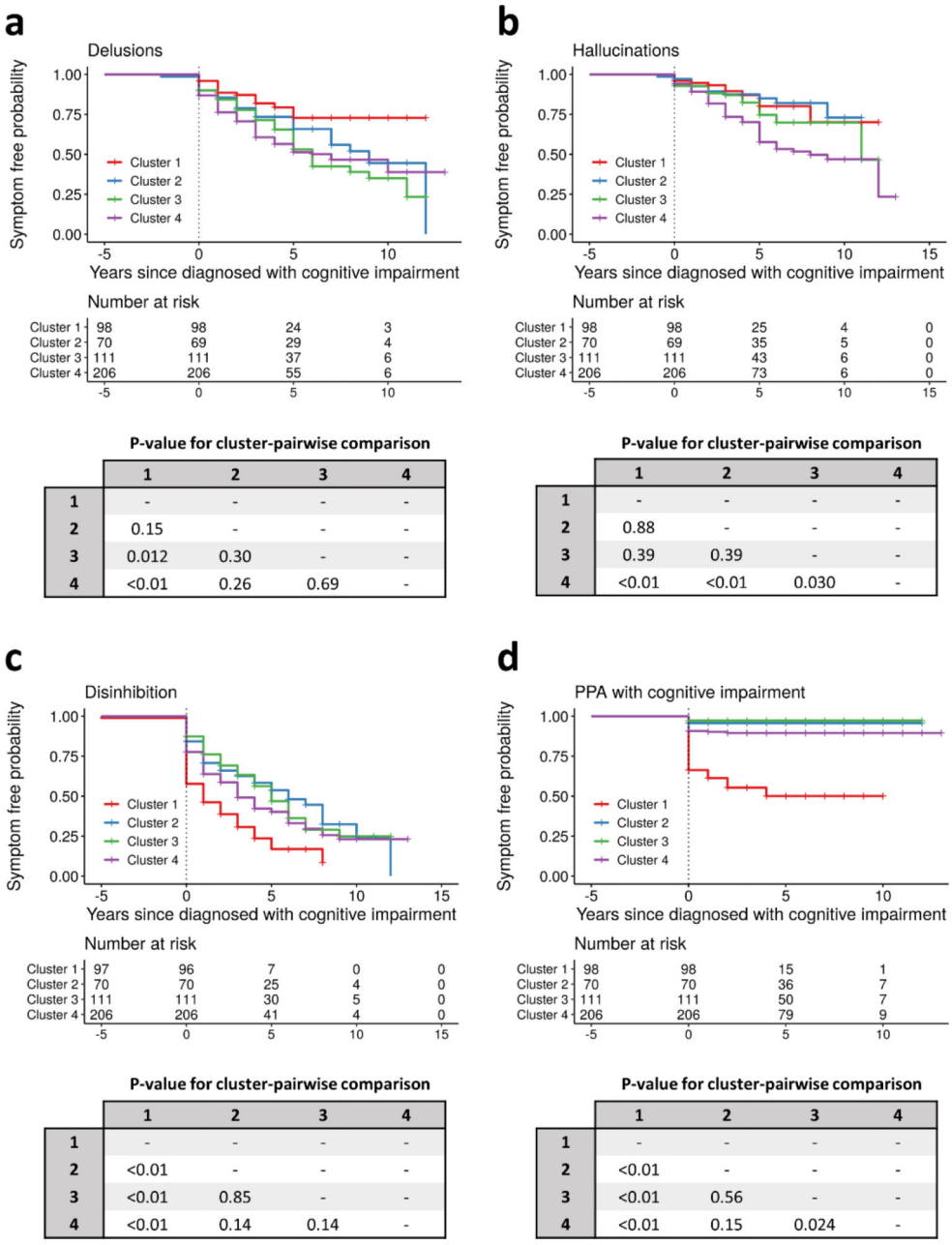


Fig. 5. Kaplan-Meier curves and Bonferroni-Holm adjusted p-value from pairwise log rank test for years since diagnosed as cognitive impairment.

The y-axis shows the symptom free probability and the x-axis indicates years since first diagnosed with cognitive impairment, that is, x = 0 represents the year when initially diagnosed as either impaired not MCI, MCI, or dementia based on the variable "NACCUDSD" in NACC UDS. Negative values of the x-axis represent the number of years before first diagnosis of cognitive impairment. MCI = mild cognitive impairment; NACC = National Alzheimer's Coordinating Center; UDS = Uniform Data Set; PPA = primary progressive aphasia; FTLD-TDP = frontotemporal lobar degeneration with TDP-43 proteinopathy

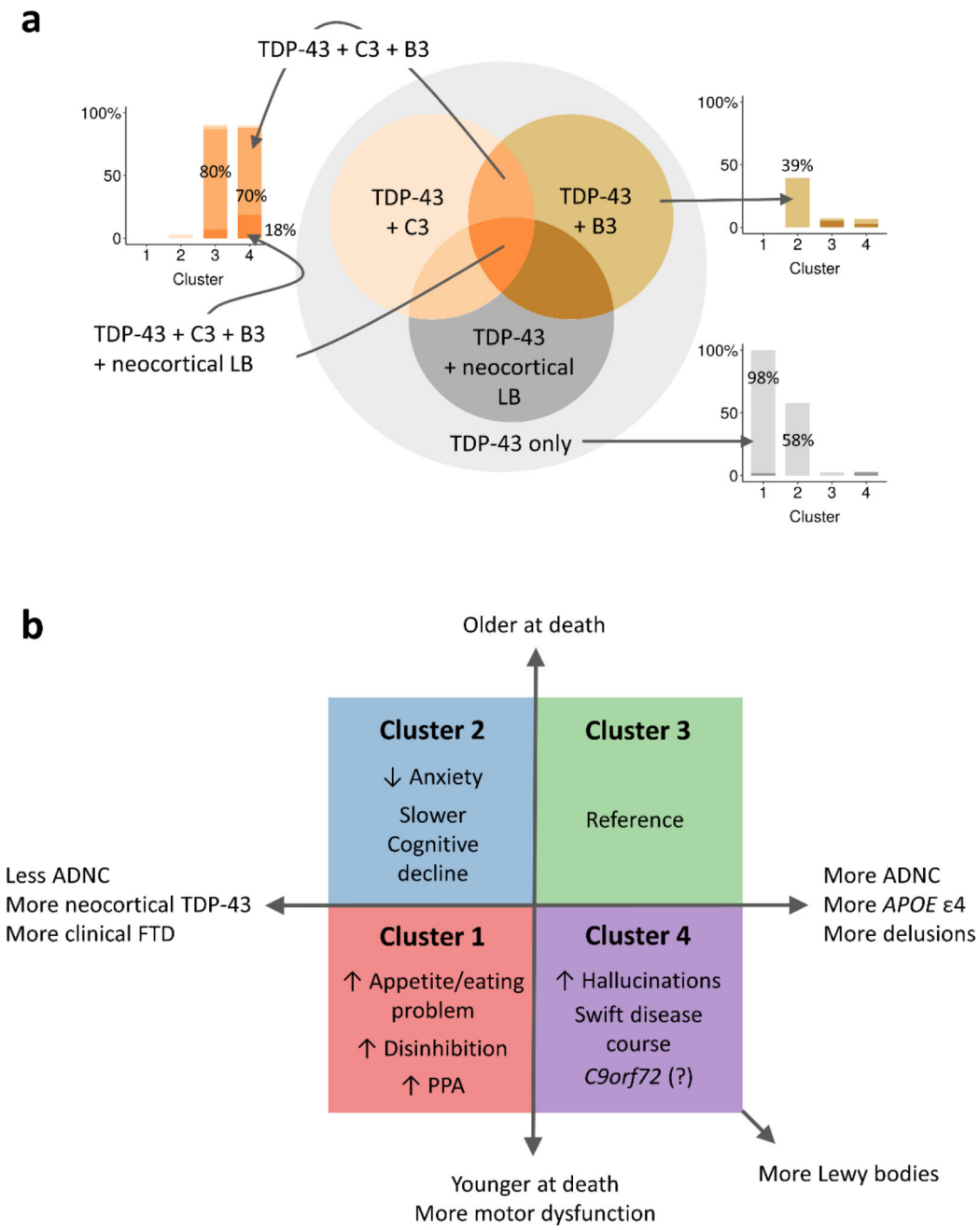


Fig. 6. Overall findings for each of four clusters.

Proportion of misfolded protein combinations in each of the clusters (**a**). Overall interpretation of the associated phenotypes for each of the clusters (**b**). C3 = frequent neocortical neuritic plaques; B3 = Braak NFT stage V or VI; LB = Lewy body; FTD = frontotemporal dementia clinical syndrome; ADNC = Alzheimer's disease neuropathologic change; PPA = primary progressive aphasia

Table 1.

Characteristics of the study subjects (n=495) on variables used in uniform manifold approximation and projection (UMAP) dimensionality reduction method by clusters, National Alzheimer's Coordinating Center (NACC) data, March 2020 data freeze

Variable	Cluster 1 (n = 103)	Cluster 2 (n = 71)	Cluster 3 (n = 114)	Cluster 4 (n = 207)
Age at death, n (%)				
< 65	18 (17.5)	4 (5.6)	1 (0.9)	12 (5.8)
≥ 65 and < 85	73 (70.9)	6 (8.5)	0 (0)	195 (94.2)
≥ 85	12 (11.6)	61 (85.9)	113 (99.1)	0 (0)
Neuritic plaques (C score), n (%) [a]				
C0	77 (74.8)	6 (8.5)	5 (4.4)	1 (0.5)
C1	14 (13.6)	19 (26.8)	1 (0.9)	2 (1.0)
C2	12 (11.7)	46 (64.7)	0 (0)	15 (7.2)
C3	0 (0)	0 (0)	108 (94.7)	189 (91.3)
Thal phase (A score), n (%) [b]				
A0	45 (43.7)	4 (5.6)	1 (0.9)	0 (0)
A1	48 (46.6)	7 (9.9)	0 (0)	0 (0)
A2	7 (6.8)	19 (26.8)	5 (4.4)	3 (1.4)
A3	3 (2.9)	41 (57.7)	108 (94.7)	204 (98.6)
Braak NFT stage (B score), n (%) [c]				
B0	26 (25.2)	3 (4.2)	1 (0.9)	0 (0)
B1	60 (58.3)	5 (7.0)	4 (3.5)	3 (1.4)
B2	17 (16.5)	33 (46.5)	7 (6.1)	11 (5.3)
B3	0 (0)	30 (42.3)	102 (89.5)	193 (93.2)
Lewy body pathology, n (%)				
No	88 (85.4)	55 (77.5)	58 (50.9)	73 (35.3)
Others	13 (12.6)	14 (19.7)	44 (38.6)	87 (42.0)
Neocortical	2 (1.9)	2 (2.8)	12 (10.5)	47 (22.7)
Cerebral amyloid angiopathy, n (%)				
None	89 (86.4)	31 (43.7)	17 (14.9)	32 (15.5)
Mild	9 (8.7)	21 (29.6)	43 (37.7)	89 (43.0)
Moderate	3 (2.9)	16 (22.5)	35 (30.7)	50 (24.2)
Severe	2 (1.9)	3 (4.2)	19 (16.7)	36 (17.4)
Infarct and lacunes, n (%)				
No	96 (93.2)	58 (81.7)	98 (86.0)	185 (89.4)
Yes	7 (6.8)	13 (18.3)	16 (14.0)	22 (10.6)
Microinfarcts, n (%)				
No	85 (82.5)	51 (71.8)	80 (70.2)	164 (79.2)
Yes	18 (17.5)	20 (28.2)	34 (29.8)	43 (20.8)
Arteriolosclerosis, n (%)				
None	18 (17.5)	14 (19.7)	12 (10.5)	29 (14.0)

Variable	Cluster 1 (n = 103)	Cluster 2 (n = 71)	Cluster 3 (n = 114)	Cluster 4 (n = 207)
Mild	49 (47.6)	11 (15.5)	26 (22.8)	70 (33.8)
Moderate	29 (28.2)	33 (46.5)	52 (45.6)	75 (36.2)
Severe	7 (6.8)	13 (18.3)	24 (21.1)	33 (15.9)

[a] C score: C0 = no, C1 = sparse, C2 = moderate, and C3 = frequent

[b] A score: A0 = phase 0, A1 = phase 1 or 2, A2 = phase 3, and A3 = phase 4 or 5

[c] B score: B0 = stage 0, B1 = stage I or II, B2 = stage III or IV, and B3 = stage V or VI

NFT = neurofibrillary tangle

Table 2.

Demographic and clinical characteristics of the study subjects on variables not used in uniform manifold approximation and projection (UMAP) dimensionality reduction method by clusters, National Alzheimer's Coordinating Center (NACC) data through the March 2020 data freeze (n = 495)

Variable	Cluster 1 (n = 103)	Cluster 2 (n = 71)	Cluster 3 (n = 114)	Cluster 4 (n = 207)
Age at death, mean ± SD	73.5 ± 10.4	88.9 ± 9.7	89.9 ± 5.0	76.6 ± 6.3
Gender, n (%)				
Men	47 (45.6)	35 (49.3)	44 (38.6)	108 (52.2)
Women	56 (54.4)	36 (50.7)	70 (61.4)	99 (47.8)
Years of education, mean ± SD	15.9 ± 2.6	16.4 ± 3.0	14.8 ± 3.4	15.6 ± 3.0
Years of follow-up, mean ± SD	6.2 ± 3.1	7.5 ± 3.3	7.9 ± 3.1	6.5 ± 2.9
APOE genotype, n (%)				
-/-	70 (77.8)	38 (60.3)	48 (45.7)	50 (27.3)
-/ε4	18 (20.0)	24 (38.1)	46 (43.8)	97 (53.0)
ε4/ ε4	2 (2.2)	1 (1.6)	11 (10.5)	36 (19.7)
Clinical status at last visit, n (%)				
Normal	5 (4.8)	1 (1.4)	4 (3.5)	1 (0.5)
Impaired-not-MCI	1 (1.0)	2 (2.8)	1 (0.9)	0 (0)
MCI	7 (6.8)	9 (12.7)	2 (1.7)	2 (1.0)
Dementia	90 (87.4)	59 (83.1)	107 (93.9)	204 (98.5)
Clinical diagnosis of AD at last visit, n (%)				
Normal	5 (4.8)	1 (1.4)	4 (3.5)	1 (0.5)
Yes	29 (28.2)	60 (84.5)	100 (87.7)	183 (88.4)
Cognitive impairment but not AD	69 (67.0)	10 (14.1)	10 (8.8)	23 (11.1)
Clinical diagnosis of FTD at last visit, n (%)				
Normal	13 (12.6)	12 (16.9)	7 (6.1)	3 (1.4)
Yes	39 (37.9)	6 (8.5)	1 (0.9)	12 (5.8)
Cognitive impairment but not FTD	51 (49.5)	53 (74.6)	106 (93.0)	192 (92.8)

SD = standard deviation; MCI = mild cognitive impairment; AD = Alzheimer's disease; FTD = frontotemporal dementia

Table 3.

Neuropathological characteristics of the study subjects on variables not used in uniform manifold approximation and projection (UMAP) dimensionality reduction method by clusters, National Alzheimer's Coordinating Center (NACC) data through the March 2020 data freeze (n = 495)

Variable	Cluster 1 (n = 103)	Cluster 2 (n = 71)	Cluster 3 (n = 114)	Cluster 4 (n = 207)
TDP-43 pathology, n (%)				
Amygdala				
No	13 (16.0)	7 (13.5)	8 (8.8)	14 (8.1)
Yes	68 (84.0)	45 (86.5)	83 (91.2)	158 (91.9)
Hippocampus				
No	8 (8.2)	5 (7.2)	15 (13.3)	48 (23.4)
Yes	90 (91.8)	64 (92.8)	98 (86.7)	157 (76.6)
Neocortex				
No	21 (23.3)	42 (70.0)	86 (80.4)	156 (83.9)
Yes	69 (76.7)	18 (30.0)	21 (19.6)	30 (16.1)
NIA-AA ADNC (ABC score)				
Not AD	45 (43.7)	4 (5.6)	1 (0.9)	0 (0)
Low ADNC	56 (54.4)	10 (14.1)	5 (4.4)	3 (1.4)
Intermediate ADNC	2 (1.9)	31 (43.7)	12 (10.5)	15 (7.2)
High ADNC	0 (0)	26 (36.6)	96 (84.2)	189 (91.3)
FTLD-TDP, n (%)				
No	24 (23.3)	59 (83.1)	109 (95.6)	196 (94.7)
Yes	79 (76.7)	12 (16.9)	5 (4.4)	11 (5.3)
Hippocampal sclerosis, n (%)				
No	68 (66.7)	41 (57.7)	69 (62.2)	139 (67.8)
Yes	34 (33.3)	30 (42.3)	42 (37.8)	66 (32.2)
Whole brain weight (g), mean ± SD	1079.6 ± 175.8	1123.2 ± 201.1	1132.5 ± 158.2	1130.5 ± 159.7
Cerebral cortex atrophy, n (%)				
None	12 (15.0)	10 (14.3)	14 (12.7)	18 (9.2)
Mild	16 (20.0)	28 (40.0)	39 (35.5)	57 (29.2)
Moderate	20 (25.0)	21 (30.0)	42 (38.2)	66 (33.8)
Severe	32 (40.0)	11 (15.7)	15 (13.6)	54 (27.7)
Lobar atrophy, n (%)				
None	25 (29.4)	50 (72.5)	86 (78.2)	150 (77.3)
Yes	60 (70.6)	19 (27.5)	24 (21.8)	44 (22.7)

SD = standard deviation; NIA-AA = National Institute on Aging - Alzheimer's Association; ADNC = Alzheimer disease neuropathologic change; FTLD-TDP = frontotemporal lobar degeneration with TDP-43 pathology

Table 4.

Feature of each of the clusters

Cluster	Feature
1	Enriched for FTLD-TDP and FTD clinically
2	Lower neuritic amyloid plaques, died at older age
3	Most with severe ADNC who died at older age
4	Most with severe ADNC who died at younger age

FTLD-TDP = frontotemporal lobar degeneration with TDP-43 proteinopathy; FTD = frontotemporal dementia; ADNC = Alzheimer's disease neuropathologic change