Received: 24 June 2024 | Revised: 21 December 2024 | Accepted: 23 December 2024

DOI: 10.1002/alz.14549

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

RESEARCH ARTICLE

The Dementia SomaSignal Test (dSST): A plasma proteomic predictor of 20-year dementia risk

Michael R. Duggan[1] | Clare Paterson[2] | Yifei Lu[3] | Hannah Biegel[2] | Heather E. Dark[1] | Jenifer Cordon[1] | Murat Bilgel[1] | Naoto Kaneko[4] | Masaki Shibayama[4] | Shintaro Kato[4,5] | Makio Furuichi[4,5] | Iwao Waga[4,5,6] | Keita Hiraga[7] | Masahisa Katsuno[7,8] | Yukiko Nishita[9] | Rei Otsuka[9] | Christos Davatzikos[10] | Guray Erus[10] | Kelsey Loupy[2] | Melissa Simpson[2] | Alexandria Lewis[11] | Abhay Moghekar[11] | Priya Palta[12] | Rebecca F. Gottesman[13] | Susan M. Resnick[1] | Josef Coresh[14] | Stephen A. Williams[2] | Keenan A. Walker[1]

[1]Laboratory of Behavioral Neuroscience, National Institute on Aging, National Institutes of Health, Baltimore, Maryland, USA

[2]Department of Clinical and Research Development, Standard BioTools, Boulder, Colorado, USA

[3]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

[4]Innovation Laboratory, NEC Solution Innovators Limited, Tokyo, Koto-ku, Japan

[5]FonesLife Proteomics Laboratory, FonesLife Corporation, Chuo City, Tokyo, Japan

[6]Well-being Design Institute for Health, Tohoku University, Aoba-ku, Sendai, Japan

[7]Department of Neurology, Nagoya University Graduate School of Medicine, Nagoya, Aichi, Japan

[8]Department of Clinical Research Education, Nagoya University Graduate School of Medicine, Nagoya, Aichi, Japan

[9]Department of Epidemiology of Aging, National Center for Geriatrics and Gerontology, Obu, Aichi, Japan

[10]Artificial Intelligence in Biomedical Imaging Laboratory, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[11]Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

[12]Department of Neurology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

[13]Stroke Branch, National Institute of Neurological Disorders and Stroke, Bethesda, Maryland, USA

[14]Departments of Population Health and Medicine, New York University Grossman School of Medicine, New York, New York, USA

Correspondence
Keenan A. Walker, Laboratory of Behavioral Neuroscience, National Institute on Aging, National Institutes of Health, Full Postal Address: NIH BRC BG RM 04B311, 251 Bayview Blvd, Baltimore, MD 21224, USA.
Email: keenan.walker@nih.gov

Funding information
Intramural Research Program (IRP) of the National Institute on Aging (NIA); NHLBI, NIA, NINDS, NIDCD, Grant/Award Numbers: 75N92022D00001, U01HL096917,

Abstract

INTRODUCTION: There is an unmet need for tools to quantify dementia risk during its multi-decade preclinical/prodromal phase, given that current biomarkers predict risk over shorter follow-up periods and are specific to Alzheimer's disease.

METHODS: Using high-throughput proteomic assays and machine learning techniques in the Atherosclerosis Risk in Communities study (n = 11,277), we developed the Dementia SomaSignal Test (dSST).

RESULTS: In addition to outperforming existing plasma biomarkers, the dSST predicted mid-life dementia risk over a 20-year follow-up across two independent cohorts

75N92022D00002, 75N92022D00003, 75N92022D00004, 75N92022D00005, U01HL096812, U01HL096814, U01HL096899, U01HL096902; NEC Solution Innovators Limited; National Center for Geriatrics and Gerontology; Nagoya University

with different ethnic backgrounds (areas under the curve [AUCs]: dSST 0.68–0.70, dSST+age 0.75–0.81). In a separate cohort, the dSST was associated with longitudinal declines across multiple cognitive domains, accelerated brain atrophy, and elevated measures of neuropathology (as evidenced by positron emission tomography and plasma biomarkers).

DISCUSSION: The dSST is a cost-effective, scalable, and minimally invasive protein-based prognostic aid that can quantify risk up to two decades before dementia onset.

KEYWORDS
dementia, machine learning, prognosis, proteomics

Highlights

- The Dementia SomaSignal Test (dSST) predicts 20-year dementia risk across two independent cohorts.
- dSST outperforms existing plasma biomarkers in predicting multi-decade dementia risk.
- dSST predicts cognitive decline and accelerated brain atrophy in a third cohort.
- dSST is a prognostic aid that can predict dementia risk over two decades.

1 | BACKGROUND

There is an unmet need for dementia risk prediction, particularly in the era of disease-modifying therapies when early detection may benefit patients most.[1] Positron emission tomography (PET), cerebrospinal fluid (CSF), and plasma measures of amyloid beta (Aβ), and phosphorylated tau (p-tau) have improved diagnostic accuracy for Alzheimer's disease (AD); however, these markers can sometimes fail to predict cognitive impairment,[2–4] and as specific indicators of AD, these measures do not capture the effects of non-AD neuropathologies that are especially prevalent among non-White individuals.[5–7] Advances in blood-based assays provide opportunities for cost-effective, scalable, and minimally invasive tools for dementia prediction. However, existing assays have been developed over limited follow-up times to specifically detect near-term (≤8 years) or prevalent AD.[8–10] Given the multi-decade preclinical and prodromal phase that precedes dementia onset and the growing consensus that early intervention will be most therapeutically effective, improved dementia prediction is critically needed.[1,11,12]

In the current study, we used a high-throughput proteomic assay to develop and validate the Dementia SomaSignal Test (dSST)—a machine learning–derived protein-based predictor of 20-year dementia risk—in 11,277 participants enrolled in the atherosclerosis risk in communities (ARIC) study (Figure 1). After demonstrating this measure's predictive accuracy in a sample of middle-aged adults and its capacity to outperform existing Alzheimer's disease and related dementias (ADRD) plasma biomarkers, we showed that it can perform even better as a predictor of 5-year dementia risk among older adults. After validating the dSST in an external cohort, the national institute for longevity sciences-longitudinal study of aging (NILS-LSA) in Japan, we also leveraged data from a third cohort, the baltimore longitudinal study of aging (BLSA), to show that the dSST predicted longitudinal declines across multiple cognitive domains, accelerated regional brain atrophy, and elevated measures of neuropathology (as evidenced by PET and plasma biomarkers of AD, neuronal injury, and astrogliosis).

2 | METHODS

All participants provided written informed consent, and procedures were approved by the institutional review boards (IRBs) affiliated with each cohort. Participants in these complete-case analyses were free of dementia diagnosis at the time of blood draw.

2.1 | ARIC study

2.1.1 | Study sample

The ARIC study is a community-based, prospective cohort study conducted in four U.S. communities: Forsyth County, NC; Jackson, MS; the northwest suburbs of Minneapolis, MN; and Washington County, MD. The ARIC study was originally designed to investigate the etiology and natural history of atherosclerosis and cardiovascular disease but has since expanded to facilitate understanding of other disease areas, including neurological health. ARIC enrolled 15,792 mostly White and Black participants, 45–64 years of age, between 1987 and 1989.[13] After initial enrollment, participants had four additional in-person

visits: Visit 2 (1990–1992), Visit 3 (1993–1995), Visit 4 (1996–1999), and Visit 5 (2011–2013). Participants were then invited back for Visit 6 (2016–2017) after 5 years. The present analyses included follow-up through the end of Visit 6. Blood was drawn for proteomic analysis at Visits 3 and 5. Twenty-year dementia risk was assessed between Visits 3 and 5, and 5-year dementia risk was assessed between Visits 5 and 6. Participants were eligible for inclusion if they had available SomaScan Assay (Standard BioTools) and dementia status data and were excluded based on baseline dementia diagnosis. Study protocols were approved by IRBs at each participating center: University of North Carolina at Chapel Hill, Chapel Hill, NC; Wake Forest University, Winston-Salem, NC; Johns Hopkins University, Baltimore, MD; University of Minnesota, Minneapolis, MN; and University of Mississippi Medical Center, Jackson, MS. All ARIC participants provided written informed consent at each study visit; proxies provided consent for participants who were judged to lack capacity.

2.1.2 | Protein measurement

Proteins were measured in ARIC using the SomaScan Assay, which uses DNA-based binding reagents (modified aptamers) to assess plasma levels of ≈5000 proteins (v4.0) with high specificity and limits of detection largely comparable to antibody-based assays.[14] Plasma was collected with standardized protocols and frozen at −80°C until analysis. The proteomic model for 20-year dementia-risk prediction was trained and validated using samples from ARIC Visit 3 (median age 60) and included all participants without known dementia at the time of blood draw. Subsequent assessment of the proteomic model for 5-year dementia risk prediction was performed in dementia-free individuals from ARIC Visit 5 (median age 75). Log_{10}-transformed protein measurements were centered and scaled using means and SDs from the training data sets. Samples from participants that did not pass SomaScan Assay quality control (QC) criteria were excluded ($n = 78$ [Visit 3], $n = 31$ [Visit 5]). At Visit 3, using 426 blind duplicates, the median coefficient of variation (CV) was 5.1% for all aptamers and 5.8% for the aptamers used to calculate the dSST. At Visit 5, using 204 blind duplicates, the median CV was 5.4% for all aptamers and 5.7% for the aptamers used to calculate the dSST. dSST scores at Visits 3 and 5 were standardized using means and SDs derived from the training data sets at Visits 3 and 5, respectively. $A\beta_{40}$, $A\beta_{42}$, glial fibrillary acidic protein (GFAP), neurofilament light (NfL), and p-tau181 concentrations were measured using the Single Molecule Array (Simoa), Neurology 4-Plex E (N4PE), and p-tau181 (V2) assays on the Simoa HD-X instrument (Quanterix).[15] Using 90 blind duplicates, CVs were 7.3%, 8.5%, 2.3%, 3.7%, and 5.6% for $A\beta_{40}$, $A\beta_{42}$, GFAP, NfL, and p-tau181, respectively. Values for GFAP, NfL, and p-tau181 were log_2 transformed to correct for skewness. $A\beta_{42/40}$ ratio and standardized values were used in analyses. Biomarker values that were 5 interquartile ranges (IQRs) above and below the third and first quartiles, respectively, were excluded.

RESEARCH IN CONTEXT

1. **Systematic review**: As indicated by available sources of evidence (e.g., PubMed), current biomarkers can predict dementia risk over short follow-up periods and are specific to Alzheimer's disease (AD), but there is an unmet need for prognostic tools to quantify dementia risk during its multi-decade prodromal/preclinical phase and independent of a specific neuropathology.
2. **Interpretation**: Using large cohort studies with 20-year follow-up periods, machine learning approaches, and state-of-the-art proteomics, we developed the Dementia SomaSignal Test (dSST) as a cost-effective, scalable, and minimally invasive prognostic aid that can quantify dementia risk up to two decades before symptom onset.
3. **Future directions**: Although additional studies are needed to validate the potential utility of dSST in clinical practice, our results address an unmet medical need in the form of a blood-based prognostic aid that predicts dementia risk during its multi-decade preclinical/prodromal phase and independent of a specific neuropathology.

2.1.3 | Dementia assessment

Dementia diagnosis was adjudicated in ARIC through surveillance of cognitive assessment tests, telephone screenings, informant ratings, hospital records, and death record reviews, as described previously.[16] At ARIC Visits 2–4, a three-instrument cognitive assessment was applied (delayed word recall task, digit symbol substitution from the Wechsler Adult Intelligence Scale–Revised [WAIS-R], and a letter fluency task). At ARIC Visits 5 and 6, participants received a comprehensive cognitive exam and a functional assessment that included the Clinical Dementia Rating (CDR) scale and Functional Activities Questionnaire (FAQ). Using these data, dementia was classified based on National Institute on Aging and Alzheimer's Association (NIA-AA) and Diagnostic and Statistical Manual of Mental Disorder, Fifth Edition (DSM-5) criteria. In the time between Visit 5 and Visit 6, participants were contacted annually via phone and administered the Six-Item Screener (SIS), a brief cognitive assessment. If participants received a low score on the SIS, or if they were unable to participate in the screening via phone, the Ascertain Dementia 8 (AD8) was administered to the participant's informant. For participants who attended Visit 6, SIS and AD8 were used to define the date of dementia onset. For participants who did not attend Visit 6, the SIS, AD8, hospital discharge, and death certificate codes were used to define dementia diagnosis and date of dementia onset.

2.1.4 | *APOE* genotyping

Apolipoprotein E (*APOE*) ε4 carrier status was available for a subset of individuals in the ARIC study and was determined with the TaqMan Assay (Applied Biosystems). Depending on the combination of alleles, an individual could possess one of six *APOE* genotypes (ε2ε2, ε2ε3, ε3ε3, ε2ε4, ε3ε4, or ε4ε4). Analyses were conducted using *APOE* genotype as a dichotomous variable based on the presence of an ε4 allele (i.e., non-carriers [ε2ε2, ε2ε3, or ε3ε3] vs carriers [ε2ε4, ε3ε4, or ε4ε4]).

2.2 | The National Institute for Longevity Sciences-Longitudinal Study of Aging (NILS-LSA)

2.2.1 | Study sample

NILS-LSA is a population-based, prospective cohort study designed to understand age-related changes among middle-aged and older community-dwellers in Aichi Prefecture, Japan.[17] Enrollment occurred from communities surrounding the NILS-LSA area, including Obu and Hagashiura. The first wave of enrollment was initiated in 1997 through 2000 and included 2267 randomly selected men and women 40–79 years of age. After the baseline visit, follow-up clinical evaluations occurred every 2 years, which included body composition and anthropometry measurements, physical function assessments, nutritional analysis, and neuropsychological assessments. The present analyses included follow-up through 2022. Blood was drawn for proteomic measurement from dementia-free individuals at the second study visit (Wave 2; 2000–2002). Participants were eligible for inclusion if they had available SomaScan Assay and dementia status data and were excluded based on baseline dementia diagnosis. Blood samples for the current analyses were selected using a case–cohort approach from the whole NILS-LSA cohort. Random selection (total $n = 340$, event/dementia cases $n = 170$, non-event/non-dementia cases $n = 170$) initially resulted in significantly different mean age (Wilcox rank sum test $p = 4.349e\text{-}13$) between dementia (mean = 71.3) and non-dementia (mean = 59.4) cases. To ensure that analyses were comparable to ARIC (i.e., where the mean age of dementia cases was 62.4 and the mean age of non-dementia cases was 59.5), random selection was reapplied after excluding older dementia cases (>78.0) and younger non-dementia cases (<55.0), resulting in age differences between dementia (mean = 68.6) and non-dementia (mean = 64.9) cases that were more similar to those observed in ARIC. The NILS-LSA protocols were approved by the IRB of the National Center for Geriatrics and Gerontology (22TB5 and 1665-3), and all participants provided written informed consent prior to participation.

2.2.2 | Protein measurement

Proteins were measured in NILS-LSA with the SomaScan Assay, which measures plasma levels of ≈7000 proteins (v4.1). Plasma was collected with standardized protocols and frozen at −80°C until analysis. All samples passed SomaScan Assay QC criteria. As no blind duplicates were included in NILS-LSA analyses, assay reliability was assessed using the variance in plasma QC samples, which was derived from 15 participants and run in triplicate on each of the five assay plates. The median CV was 4.1% for all aptamers, and 4.6% for the aptamers used to calculate the dSST. dSST scores were standardized using means and SDs derived from the whole NILS-LSA data set.

2.2.3 | Dementia assessment

Dementia diagnosis was adjudicated in NILS-LSA through Long Term Care Insurance, a mandatory form of national social insurance to assist activities of daily living in the disabled elderly in Japan.[18] The dementia criteria use a scale-based adjudication, requiring information from questionnaires (developed by the Japanese Ministry of Health, Labor, and Welfare) to assess the degree of functional disability, and official documentation based on ascertainment of an attending physician (the Doctor's Opinion Paper). According to the Levels of Independence Degree in Daily Living for Elderly with Dementia reported by the attending physician, the applicant's independence is classified into six ranks (0, I–IV, and M). Consistent with prior studies, individuals with a dementia scale degree of ≥IIa were classified as dementia cases.[19] Such dementia criteria display 73% and 96% sensitivity and specificity, respectively (compared to neuropsychiatrist-based clinical diagnoses), and correlate well with Mini-Mental State Examination (MMSE) scores.[20,21]

2.2.4 | *APOE* genotyping

APOE ε4 carrier status was determined with the Type IIP enzyme Hhai (Toyobo). Depending on the combination of alleles, an individual could possess one of six common *APOE* genotypes (ε2ε2, ε2ε3, ε3ε3, ε2ε4, ε3ε4, or ε4ε4). Analyses were conducted using *APOE* genotype as a dichotomous variable based on the presence of an ε4 allele (i.e., non-carriers [ε2ε2, ε2ε3, or ε3ε3] vs carriers [ε2ε4, ε3ε4, or ε4ε4]).

2.3 | The Baltimore Longitudinal Study of Aging (BLSA)

2.3.1 | Study sample

The BLSA is a community-based, prospective cohort study based in Baltimore, MD, USA, which is designed to assess physical and cognitive measures in a cohort of community-dwelling volunteers.[22] Participants free of major chronic diseases as well as cognitive and functional impairment at the time of enrollment received comprehensive health and functional screening evaluations at subsequent study visits, which were completed by licensed health care professionals (e.g., nurse practitioner, medical doctor). Study visits occurred biennially until 2005, and then every 1–4 years depending on age (age <60 years, every 4

years; age 60–79 years, every 2 years; age ≥80 years, every year). For a group of participants enrolled in the BLSA neuroimaging substudy, visits occurred annually beginning in 1994. Cognitive tasks used in the current analyses were initiated in the BLSA between 1984 and 1993. PET scans for ^{11}C-Pittsburgh compound B (PiB) and ^{18}F-flortaucipir (FTP) were implemented in 2005 and 2016, respectively. Beginning in 2009–2010, serial 3T magnetic resonance imaging (MRI) scans were collected. Blood was drawn for proteomic and ADRD biomarker measurement at baseline MRI or PET scan. For participants with multiple blood draws, the earliest sample was used in the current analyses. Due to BLSA's continuous enrollment, participants entered the study at different times and thus varied with respect to follow-up times. The present analyses included follow-up through 2020. The BLSA protocol was approved by the IRB of the National Institute of Environmental Health Science, National Institutes of Health (NIH; 03AG0325), all participants provided written informed consent prior to participation and de-identified data were used for analyses. Participants were eligible for inclusion if they had available SomaScan Assay and cognition, MRI, PET, or plasma biomarker data, and were excluded based on neurologic conditions that could affect brain structure or function (e.g., strokes, seizures) and baseline cognitive impairment (i.e., dementia, mild cognitive impairment [MCI], impaired but not MCI; procedures for determining cognitive status are detailed elsewhere[23]).

2.3.2 | Protein measurement

Proteins were measured with the SomaScan Assay, which measures plasma levels of ≈7000 proteins (v4.1).[24] Plasma was collected with standardized protocols and frozen at −80°C until analysis. Samples from participants that did not pass SomaScan QC criteria were excluded ($n = 7$). The median CV was 4.5% for all aptamers, and 4.0% for the aptamers used to calculate the dSST (102 blind duplicates). The CV for the dSST was 5.4%. The median inter-plate (128 blind duplicates) and inter-batch (27 blind duplicates) CVs for the aptamers used to calculate the dSST were 4% and 7%, respectively.[25] dSST scores were standardized using means and SDs derived from the whole BLSA data set. ADRD biomarkers ($A\beta_{40}$, $A\beta_{42}$, GFAP, NfL, and p-tau181) were measured using Simoa, N4PE, and p-tau181 (V2) assays on the Simoa HD-X instrument (Quanterix). Assays were run in duplicate and values averaged. Intra-assay CVs were 2.8%, 1.9%, 5.0%, 5.1%, and 4.4% for $A\beta_{40}$, $A\beta_{42}$, GFAP, NfL, and p-tau181, respectively. Values for GFAP, NfL, and p-tau181 were $\log_2$ transformed to correct for skewness. $A\beta_{42/40}$ ratio and standardized biomarker values were used in the analyses.

2.3.3 | Cognitive performance

Composite scores across five cognitive domains (visuospatial ability, verbal memory, verbal fluency, executive functioning, attention) were calculated from standardized (converted to a *z*-score using the baseline mean and SD) and averaged individual task components.[26,27] Because certain cognitive tasks were initiated in the BLSA at different periods according to protocol changes, composite scores for participants at each visit were computed from those tasks available at the time of assessment. Visuospatial ability was assessed using a modified version of the Educational Testing Service Card Rotations Test and two Clock Drawing Tests (CDTs), which asked participants to draw the hands and face of clocks indicating 3:25 and 11:10; here, a composite score was calculated using the mean of the standardized *z*-scores from the Card Rotations Test and the mean of the CDTs. Verbal memory was measured using immediate (sum of five learning trials) and long-delay free recall from the California Verbal Learning Test. Verbal Fluency-Letters (F, A, and S) and Verbal Fluency-Categories (fruits, animals, and vegetables) were calculated to determine the verbal fluency composite score. Executive function was assessed using the Trail Making Test Part B and the Digit Span Backward subset of the WAIS-R, whereas attention was evaluated using Trail Making Test Part A and the Digit Span Forward subset of the WAIS-R. Scores of Trail Making Test Parts A and B were first natural log transformed, *z* scored, and then signs inverted so that higher scores reflect higher performance, consistent with the direction of performance across other cognitive tasks. MMSE standardized scores (total correct) were used to assess general cognitive functioning.

2.3.4 | Magnetic resonance imaging

Brain scans were collected using 3T MRI.[28] A validated, Multi-atlas Region Segmentation Utilizing Ensembles (MUSE) anatomic labeling method, which is specifically designed to achieve a consistent parcellation of brain anatomy in longitudinal MRI studies using T1-weighted sequences,[29] was applied to T1-weighted magnetization-prepared rapid gradient echo (MPRAGE) scans acquired on a 3T Philips Achieva (repetition time [TR] = 6.8 ms, echo time [TE] = 3.2 ms, flip angle = 8°, image matrix = 256 × 256, 170 slices, pixel size = 1 × 1 mm, slice thickness = 1.2 mm). Voxel-wise tissue density maps for different brain tissue types were also calculated using the RAVENS methodology (Regional Analysis of Volumes Examined in Normalized Space).[30] To achieve optimal consistency between regional (volumetric) and voxel-wise analyses, MUSE region of interest (ROI) labels were used to segment the brain into gray matter (GM), white matter (WM), ventricles (VN), and cerebrospinal fluid (CSF) for RAVENS map calculations. RAVENS map intensity values quantify the regional tissue volumes for a subject at each voxel on a common template space, with one RAVENS map for each tissue type. RAVENS maps allow for calculating voxel-wise statistical maps from a group of subjects without using prior regional definitions; this approach has been used extensively and validated in large-scale neuroimaging studies.[31,32] For voxel-based morphometry (VBM) analyses, images were smoothed at a 2 mm Gaussian filter, and the Montreal Neuroimaging Institute (MNI) template space was employed. Using RAVENS estimates, we calculated a machine learning–based score known as the SPARE-AD (Spatial Pattern of Atrophy for Recognition of Alzheimer's disease), which captures multi-variate changes in brain structure that accurately discriminate cognitively normal subjects from other neurodegenerative

phenotypes, particularly AD.[33–36] In brief, it is computed by training a support vector machine (SVM) classification model to distinguish cognitively normal from clinically diagnosed dementia populations using structural brain features, and has been shown to discriminate between normal cognition and MCI as well as conversion for such individuals to MCI and AD dementia, respectively.[37–40] Because of scanner changes over time (3T MRIs were not initiated until 2009–2010 but PiB was initiated in 2005), the ROI masks for PiB participants before 2008 were generated using MPRAGE scans collected on a 1.5T scanner (Philips Intera, repetition time = 6.8 ms, echo time = 3.3 ms, flip angle = 8°, image matrix = 256 × 256, 124 slices, pixel size = 0.94 × 0.94 mm, slice thickness = 1.5 mm) or spoiled gradient-recalled (SPGR) scans on a 1.5T scanner (GE Signa, repetition time = 35 ms, echo time = 5 ms, flip angle = 45°, image matrix = 256 × 256, 124 slices, pixel size = 0.94 × 0.94 mm, slice thickness = 1.5 mm).

2.3.5 | PET

PiB distribution volume ratios (DVRs) and FTP standardized uptake value ratios (SUVRs) were measured using positron PET. PiB scans (70 min) were acquired on a GE Advance or Siemens High Resolution Research Tomograph (HRRT) scanner following an intravenous bolus injection of ≈555 MBq of the radiotracer. DVRs were computed with a spatially constrained simplified reference tissue model using cerebellar GM as a reference. Mean cortical Aβ reflected the average DVR values across the cingulate, frontal, parietal (including precuneus), lateral temporal, and lateral occipital regions, excluding the pre- and post-central gyri. Mean cortical DVR values were harmonized between the two scanners by leveraging longitudinal data available on both scanners for 79 participants. Aβ PET status (±) was defined based on a Gaussian mixture model threshold of 1.064 mean cortical DVR.[41] Participants who converted from Aβ− to Aβ+ (n = 7) had ≥3 scans; for these participants, observations at which a participant's Aβ PET status reflected the participant's status at ≥50% of scans were preserved. One participant who converted from PiB+ to PiB− across their only two scans was excluded. FTP scans (30 min) were acquired on a Siemens HRRT scanner 75 min after an i.v. bolus injection of ≈370 MBq of the radiotracer. Using PET images partially volume corrected with a region-based voxel-wise method, SUVRs were computed with the inferior cerebellar GM as a reference. Averaged bilateral SUVRs for the entorhinal cortex (ERC) and the inferior temporal gyrus (ITG) were examined. Full details of PiB and FTP acquisition and processing have been described elsewhere.[41]

2.3.6 | *APOE* genotyping

APOE ε4 carrier status was available for a subset of individuals in the BLSA study and was defined via polymerase chain reaction (PCR) with restriction isotyping using the Type IIP enzyme Hhai or the Taqman method. Depending on the combination of alleles, an individual could possess one of six *APOE* genotypes (ε2ε2, ε2ε3, ε3ε3, ε2ε4, ε3ε4, or ε4ε4). Analyses were conducted using *APOE* genotype as a dichotomous variable based on the presence of an ε4 allele (i.e., non-carriers [ε2ε2, ε2ε3, or ε3ε3], carriers [ε2ε4, ε3ε4, or ε4ε4], or missing).

2.4 | Protein characterization

A systematic, manual literature search utilizing PubMed was used to identify literature-based evidence that reported relationships of individual proteins included in the dSST calculation with dementia or dementia-related neuropathology. Articles written in English and accepted (or *in press*) following peer review were considered. For comparative gene expression across cell types, we used consensus transcript expression levels (normalized Transcripts per Million [nTPM]) from the Human Protein Atlas (https://www.proteinatlas.org/about/). If a central nervous system (CNS) cell (i.e., astrocytes, excitatory/inhibitory neurons, microglia, etc.) was among the top five cell types with the highest expression of a given gene, it was considered highly expressed in a CNS cell. Plasma and CSF SomaScan Assay v4.1 proteomic measurements derived from 18 control and 18 AD participants made available by the Emory Goizueta Alzheimer's Disease Research Center[42] were used to assess the associations between specific proteins measured in plasma and CSF. Supplemental information relevant to AD, including expression levels (RNA, protein) in post-mortem brain tissue, was obtained from the AD Knowledge Portal (https://adknowledgeportal.synapse.org), a platform for accessing data, analyses, and tools generated by the Accelerating Medicines Partnership Program for AD (AMP-AD) and other NIA-supported programs. Expression levels in neurovasculature cell types were obtained from the Human BBB (https://twc-stanford.shinyapps.io/human_bbb/), a transcriptomic dataset generated using VINE (Vessel Isolation and Nuclei Extraction)-seq[43]; Welch's independent samples t-test was used to determine if cognate genes encoding proteins were differentially expressed in neurovasculature cell types in AD brain tissue. The Open Targets Platform (https://platform.opentargets.org) was used to identify medications that target specific proteins.

2.5 | Diversity, equity, and inclusion

We are committed to the promotion of diversity, equity, and inclusion. We believe greater diversity in race, ethnicity, gender, and other factors (e.g., gender orientation, socioeconomic status, and so on) is needed to enhance equitable outcomes. Although self-reported race and sex can serve as a proxy for a number of different social, economic, and health factors that may disproportionately affect minority individuals, we intentionally prioritized dementia-risk prediction without the use of demographic information or clinical characteristics to maximize clinical utility and scalability across populations. We recognize NIH definitions of sex (biological differences between females and males) and gender (socially constructed roles and behaviors); self-reported sex (rather than gender) was used in analyses based on available demographic information in the current cohorts.

2.6 | Statistical analyses

To maximize the clinical utility of the dSST, we prioritized dementia risk prediction independent of demographic characteristics and comorbidities. Statistical significance was defined at two-sided $p < 0.05$. Analyses were performed in R version 4.2.

2.6.1 | ARIC

The dSST was developed using data from ARIC Visit 3, which were randomly split into training (70%), tuning (15%), and validation (15%) data sets. All protein levels in this study were first $\log_{10}$ transformed and then centered and scaled using means and SD of the ARIC Visit 3 training data set. First, unadjusted Cox proportional hazards (PH) regression models were fit in the training data set to identify proteins linked to incident dementia over 20 years. Second, based on *p*-values of beta coefficients in these univariate models, the 50 proteins that showed the strongest associations with dementia risk were included as features in a single Cox least absolulte shrinkage and selection operator (LASSO) iteration, which reduced the feature set to 25 proteins. Fitting a Cox elastic net model using 10-fold cross-validation supported no further reduction of the feature set. An accelerated failure time (AFT) model with Weibull distribution resulted in a final 25-protein model fit, which generated continuous scores (i.e., reflecting the absolute likelihood of developing dementia over 20 years) that we called the dSST (e.g., a dSST score of 1 reflected a 1% probability of developing dementia within 20 years of blood draw). Robustness of the model was assessed using the tuning data set, allowing for the adjustment of the modeling approach to account for any overfitting. After the model was deemed robust, no further adjustments to the modeling approach were made, and the validation data set was used to test model performance.[44,45] The average risk of developing dementia (i.e., the predicted probability of developing dementia based on the mean of the 25 proteins selected for the dSST) within 20 years was 15% (or a dSST score of 15) in the training data set. Four non-overlapping risk strata were classified based on relative risk to the training data set average: more than 50% reduced risk (*low*; dSST score $\leq$7), between 50% reduced risk and average risk (*medium-low*; dSST score >7 and $\leq$15), between average risk and 50% elevated risk (*medium-high*; dSST score >15 and $\leq$22), and greater than 50% elevated risk (*high*; dSST score >22).

Using beta estimates obtained from the AFT Weibull model at ARIC Visit 3, which were then applied to the same proteins measured at ARIC Visit 5 (i.e., $\log_{10}$ transformed protein measurements centered and scaled using means and SDs of the ARIC Visit 3 training data set), dSST scores were calculated to assess dSST performance for 5-year risk prediction among older adults. ARIC Visit 5 data were randomly split into training (70%), tuning (15%), and validation (15%) data sets, and preliminary analyses in the training data set indicated that the dSST developed to predict 20-year dementia risk in mid-life was underpredicting 5-year dementia risk in late life due to differences in event rates. Using the Visit 5 training data set, the dSST (which was calculated with beta estimates obtained from the AFT Weibull model at ARIC Visit 3) was recalibrated considering linear regression, isotonic regression, and Platt scaling, resulting in a piecewise linear recalibration of the original model. Robustness of the linear-transformed, recalibrated model was assessed using the tuning data set, which showed that no further adjustments to the modeling approach were necessary. The validation data set was then used to test model performance. The average risk of developing dementia within 5 years was 6.5% (or a dSST score of 6.5) in the training data set. Four non-overlapping risk strata were classified based on relative risk to the training data set average: more than 50% reduced risk (*low*; dSST score $\leq$ 3.1), between 50% reduced risk and average risk (*medium-low*; dSST score > 3.1 and $\leq$ 6.5), between average risk and 50% elevated risk (*medium-high*; dSST score > 6.5 and $\leq$ 9.7), and greater than 50% elevated risk (*high*; dSST score > 9.7).

Discriminative performance of continuous dSST scores was primarily evaluated using area under the receiver-operating characteristic (ROC) curve (AUC). Sensitivity and specificity were assessed using a dSST cutoff score that predicted dementia status with a maximized Youden Index (ARIC Visit 3: 0.160; ARIC Visit 5: 0.073). Sensitivity and specificity at ARIC Visit 5 were also assessed using the same dSST cutoff score that predicted dementia status with a maximized Youden Index at ARIC Visit 3 (i.e., 0.160); here, the cutoff score was subjected to the same linear transformation applied to ARIC Visit 5 data set to account for differences in dementia event rates between mid-life and late life, resulting in a cutoff score of 0.190. Performance of continuous dSST scores was also evaluated using Cox PH regression models. Discriminative performance of dSST risk categories was evaluated primarily using the dynamic range, defined as the ratio of Kaplan–Meier (KM) event rates between the *low*- and *high*-risk categories. With the *low*-risk category serving as the comparator, performance of dSST risk categories was also evaluated using Cox PH regression models. Event-free probabilities and observed event rates were based on KM estimates, whereas predicted event rates were based on continuous dSST scores; confidence intervals (CIs) for dynamic range were calculated using bootstrapping, and bootstrap iterations where there were no diagnoses within the timeframe were not included in the CI estimate. Model calibration was assessed with the Hosmer–Lemeshow Goodness-of-Fit using the Greenwood–Nam-D'Agostino method and deciles. The Cox proportionality assumption was tested by computing and plotting Schoenfeld residuals. Among participants with *APOE* genotype information available, discriminative performance of *APOE* genotype was also evaluated using AUCs and Cox PH regression models. Discriminative performance of continuous dSST scores and standardized ADRD plasma biomarker levels ($A\beta_{42/40}$, GFAP, NfL, and p-tau181) was evaluated using Cox PH regression models. Cox PH regression models were also used examined the associations of dSST scores with late-life dementia risk among participants whose dSST risk categories did and did not shift across mid-life and late-life (e.g., among participants in the *low*-risk category in mid-life, what is the association between dSST scores and late-life dementia risk among participants who retained their *low*-risk category in late life compared those participants who shifted to the *medium-low*, *medium-high*, or *high*-risk categories). Age and age+dSST comparator models were made by

fitting AFT Weibull models to age and dSST scores from the training data set at each ARIC Visit. Prior to fitting these models, age was standardized, whereas dSST scores were transformed from risk probability space to linear space and standardized. Baseline risk for each of these comparator models was calculated using the average values (age and dSST) from the training data sets, and risk bins were defined based on the previously defined relative risk strata. Continuous dSST scores evaluated using Cox PH regression models were standardized, with scores from ARIC Visit 3 standardized using means and SDs derived from training data set at Visit 3, and scores from ARIC Visit 5 standardized using means and SDs derived from training data set at Visit 5.

2.6.2 | NILS-LSA

Using beta estimates obtained from the AFT Weibull model at ARIC Visit 3, which were then applied to the same proteins measured in NILS-LSA (i.e., $\log_{10}$ transformed protein measurements centered and scaled using means and SDs of the ARIC Visit 3 training data set), dSST scores in this separate cohort were calculated. Because of potential differences in overall protein abundance across different populations and cohort studies, NILS-LSA data were additionally z-scored to improve interpretability. Four non-overlapping risk strata were classified based on relative risk to the training data set average observed at ARIC Visit 3. However, due to limited sample size in the *low*-risk category, three risk categories were used in NILS-LSA analyses: *low/medium-low*, *medium-high*, and *high*. Model calibration in NILS-LSA was not assessed due to the case–cohort design of sample selection. Discriminative performance of continuous dSST scores was primarily evaluated using AUCs. Sensitivity and specificity were assessed using a dSST cutoff score that predicted dementia status with a maximized Youden Index (0.139). Using dSST scores centered and scaled using means and SDs of the ARIC Visit 3 training data set (i.e., without additional *z*-scoring to account for potential differences in overall protein abundance), sensitivity and specificity were also assessed using the same dSST cutoff score that predicted dementia status with a maximized Youden Index at ARIC Visit 3 (i.e., 0.160). Performance of continuous dSST scores was also evaluated using Cox PH regression models. Discriminative performance of dSST risk categories was primarily evaluated using the dynamic range, defined as the ratio of KM event rates between the *low/medium-low* and *high*-risk categories. With the *low/medium-low* risk category serving as the comparator, performance of dSST risk categories was also evaluated using Cox PH regression models. Event-free probabilities and observed event rates were based on KM estimates, whereas predicted event rates were based on continuous dSST scores. The Cox proportionality assumption was tested by computing and plotting Schoenfeld residuals. Discriminative performance of *APOE* genotype was also evaluated using AUCs and Cox PH regression models. Due to the case–cohort selection of NILS-LSA samples used in the current analyses, Cox PH regression models were adjusted for a sampling weight that reflected the 20-year incident dementia rate of the larger NILS-LSA cohort. For age and age+dSST comparator models, four non-overlapping risk strata were classified based on relative risk to the training data set average observed at ARIC Visit 3. However, due to limited sample size in the *low*-risk category, three risk categories were used in NILS-LSA analyses: *low/medium-low*, *medium-high*, and *high*. Continuous dSST scores evaluated using Cox PH regression models were standardized, with scores standardized using means and SDs derived from the whole NILS-LSA data set.

2.6.3 | BLSA

Using beta estimates obtained from the AFT Weibull model at ARIC Visit 3, which were then applied to the same proteins measured in BLSA (i.e., $\log_{10}$ transformed protein measurements centered and scaled using means and SDs of the ARIC Visit 3 training data set), dSST scores in this separate cohort were calculated. Because of potential differences in overall protein abundance across different populations and cohort studies, BLSA data were additionally z-scored to improve interpretability. Four non-overlapping risk strata were classified based on relative risk to the training data set average observed at ARIC Visit 3. Linear mixed-effects regression models adjusted for time and dSST*time were used to examine associations of continuous dSST scores and its risk categories with cross-sectional and longitudinal cognitive performance and brain volumes. From these models, the fixed effect of the dSST variable estimated the cross-sectional association between the dSST variable and baseline cognitive domain score/brain volume, the dSST variable*time estimated the effect of dSST variable on cognitive domain score/brain volume rates of change, and time estimated the average rate of cognitive domain score/brain volume change when all the predictors that interact with time are equal to zero. Random effects of intercept and time with unstructured covariance were included to account for the within-subject correlation of the repeated measurements. For VBM analyses, a threshold of 50 voxels with an uncorrected $p < 0.001$ was used to define significant clusters. Logistic regression was used to examine the cross-sectional associations of continuous dSST scores with Aβ PET status (±). Performance of continuous dSST scores and standardized ADRD plasma biomarker levels (A$\beta_{42/40}$, GFAP, NfL, and p-tau181) for discriminating Aβ PET status (±) were evaluated using logistic regression. Linear regression models were used to examine cross-sectional associations of continuous dSST scores with continuous regional tau PET measures (ERC and ITG). Due to limited sample size, the relationship between dSST risk categories and PET measures was not assessed. Linear regression was used to examine associations of continuous dSST scores and its risk categories with cross-sectional standardized ADRD plasma biomarker levels (A$\beta_{42/40}$, GFAP, NfL, and p-tau181). The *low*-risk category served as the comparator in analyses examining risk categories. Brain MRI analyses adjusted for intracranial volume (defined at age 70) and tau PET analyses adjusted for Aβ PET status (±). Continuous dSST scores evaluated using linear mixed-effects, logistic and linear regression models were standardized, with scores standardized using means and SDs derived from the whole BLSA data set.

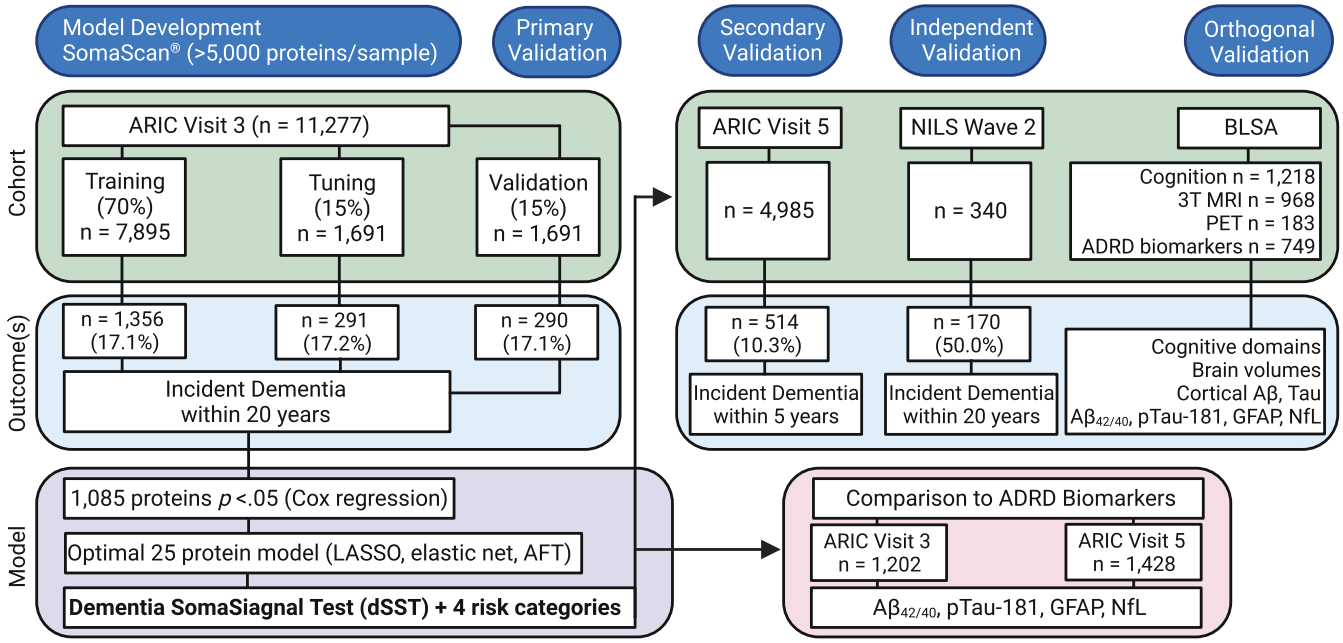

FIGURE 1 Study design. Using high-throughput plasma proteomic data from the ARIC study, a machine-learning based predictor of mid-life (20-year) incident dementia risk called the dSST was developed and validated. After testing the dSST's capacity for predicting both mid- and late-life (5-year) dementia risk and comparing its performance to established ADRD plasma biomarkers in ARIC, the NILS-LSA in Japan was used to test the dSST's predictive validity for 20-year dementia risk. The BLSA was also used to assess the dSST's relationships with domain-specific cognitive decline, changes in 3T MRI-derived brain volumes, and differences in neuropathology, as evidenced by PET and ADRD plasma biomarkers. ADRD, Alzheimer's disease and related dementias; AFT, accelerated failure time (model); ARIC, Atherosclerosis Risk in Communities; BLSA, Baltimore Longitudinal Study of Aging; dSST, Dementia SomaSignal Test; LASSO, least absolute shrinkage and selection operator; MRI, magnetic resonance imaging; NILS-LSA, National Institute for Longevity Sciences-Longitudinal Study of Aging; PET, positron emission tomography.

3 | RESULTS

3.1 | dSST predicts 20-year and 5-year dementia risk in ARIC

The dSST was generated using plasma proteomic data in a cohort who were middle-aged at time of blood draw and followed for 20 years (total n = 11,277), namely ARIC participants at Visit 3 (1993–1995) followed until Visit 5 (2011–2013). The dSST was also tested in an older cohort from a later wave of the study (total n = 4,985), namely ARIC participants at Visit 5 followed until Visit 6 (2016–2017). Proteins were measured with the SomaScan Assay v4.0 (≈5,000 proteins), although all proteins used to develop the dSST are also available on 4.1 (≈7000 proteins; Table S1). In brief, proteins that showed the strongest associations with dementia risk in Cox PH models were included in a single Cox LASSO iteration followed by 10-fold cross-validation, and an AFT model was applied to the reduced feature set of 25 proteins to generate continuous relative risk (i.e., dSST) scores. Along with the continuous dSST, four non-overlapping risk strata were defined based on relative risk in the training data set: more than 50% reduced risk (*low*), between 50% reduced risk and average risk (*medium-low*), between average risk and 50% elevated risk (*medium-high*), and greater than 50% elevated risk (*high*).

We identified an optimal, 25-protein-only survival model prognostic of mid-life dementia risk over a 20-year follow-up (training n = 7,895; tuning n = 1,691; validation n = 1,691; median age 60; Figure 1; Table S2-4). dSST scores were elevated among males, *APOE* ε4 carriers, with older age and higher body mass index (BMI); levels of most proteins used in the dSST calculation were also significantly elevated with older age (Figure S1A, B; Table S4, 5; Figure S1A, B). The dSST displayed AUCs of 0.732 and 0.700 in the training and validation data sets, respectively (Figure 2A; Table S6). Although age (AUC: 0.776) significantly outperformed the dSST, a model that included age+dSST (AUC: 0.811) performed significantly better than age alone, and demonstrated superior discrimination compared to *APOE* alone (AUC: 0.600). In sex-stratified analyses, model performance was similar and did not differ from performance when female and male data were combined. The combination of the dSST, age, and sex yielded the same performance compared to the age+dSST model (AUC: 0.811).

dSST risk categories reliably stratified clinically relevant differences in dementia risk, with 20-year observed event rates of 3.8% in *low*, 10.9% in *medium-low*, 18.4% in *medium-high*, and 35.6% in *high* (Figure 2B). The observed and predicted event rates according to risk categories and deciles showed the dSST was well calibrated (Figure 2C; Figure S1C). The relative difference in risk between the *low*- and *high*-risk categories was 9.39-fold (10.98-fold after adding age to the model), compared to a 7.78-fold and 1.99-fold difference associated with age and *APOE*, respectively. Although the dSST outperformed AD-specific ($A\beta_{42/40}$, p-tau181) plasma biomarkers in discriminating 20-year dementia risk, this difference in performance was not statistically

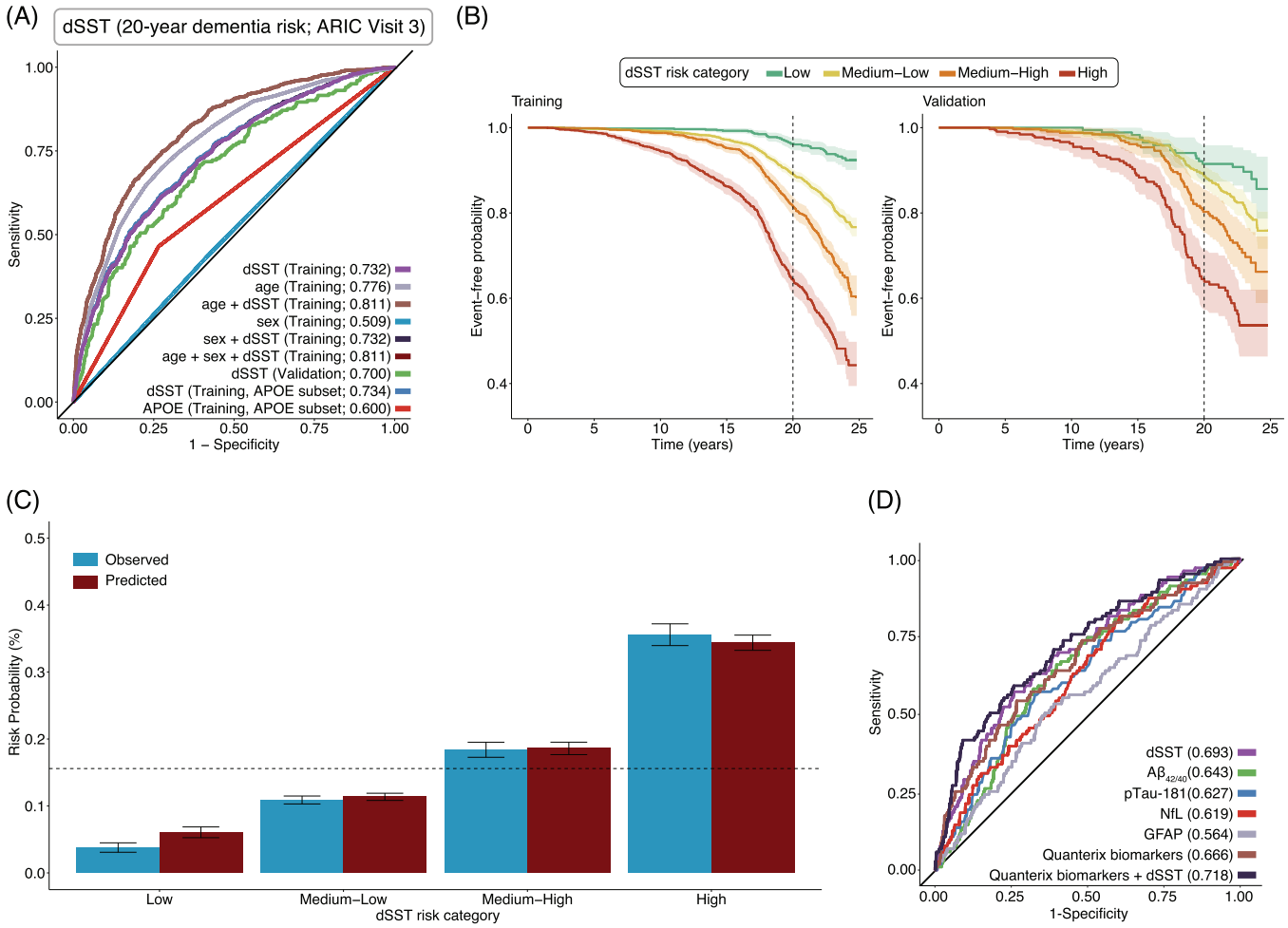

FIGURE 2 The dSST predicts 20-year dementia risk in the ARIC study. (A) Using blood drawn from Visit 3 (1993–1995) and dementia follow-up through Visit 5 (2011–2013), ROC curves show the AUCs of dSST scores in discriminating 20-year dementia risk across training and validation data sets, as well as performance of age and *APOE* genotype. (B) Kaplan–Meier plots show the observed 20-year dementia event-free probabilities grouped according to dSST risk categories (*low*, *medium-low*, *medium-high*, and *high*) across training and validation data sets using proteomic data from Visit 3. Event-free probabilities were calculated across the observed times for each risk group. (C) Bar charts show the observed and predicted 20-year dementia event rate probabilities according to dSST risk categories (*low*, *medium-low*, *medium-high*, and *high*) in the training data set. Predicted event rates were calculated as the mean dSST score for each risk group; the dotted line reflects average risk. (D) ROC curves show the AUC of dSST scores and ADRD plasma biomarker levels ($A\beta_{42/40}$, GFAP, NfL, and p-tau181) in discriminating 20-year dementia risk. Results derived from Cox proportional hazard regression models. Aβ, amyloid beta; *APOE*, apolipoprotein E; ARIC, Atherosclerosis Risk in Communities; AUC, area under the curve; dSST, Dementia SomaSignal Test; GFAP, glial fibrillary acidic protein; NfL, neurofilament light; p-tau, phosphorylated tau; ROC, receiver-operating characteristic.

significant (Table S7). However, the dSST offered significantly improved performance compared to non-specific neurodegenerative (GFAP, NfL) plasma biomarkers, and significantly improved predictive accuracy when used in combination (Quanterix biomarkers AUC: 0.666 vs Quanterix biomarkers + dSST AUC: 0.718) (Figure 2D; Figure S1D; Table S7).

The ability of the dSST to predict late-life dementia risk over a 5-year follow-up was also assessed (training n = 3,488; tuning n = 748; validation n = 749; median age 75; Tables S2, 5; Figure S2A, B). The dSST displayed AUCs of 0.778 and 0.776 in the training and validation data sets, respectively, which were significantly better compared to performance in mid-life (Figure 3A; Table S6). The dSST significantly outperformed *APOE* (AUC: 0.570). Although the dSST also offered improved prediction compared to age (AUC: 0.755), this difference was not statistically different. However, the combination of age+dSST (AUC: 0.807) performed significantly better than age alone. In sex-stratified analyses, model performance was similar and did not differ from performance when female and male data were combined. The combination of the dSST, age, and sex yielded the same performance compared to the age+dSST model (AUC: 0.807). Using a dSST cutoff score that predicted dementia status with a maximized Youden Index at ARIC Visit 3 instead of ARIC Visit 5 impacted sensitivity and specificity, suggesting that cutoff scores applied to prediction of 20-year dementia risk among middle-aged adults should not be applied to 5-year dementia-risk prediction in older adults.

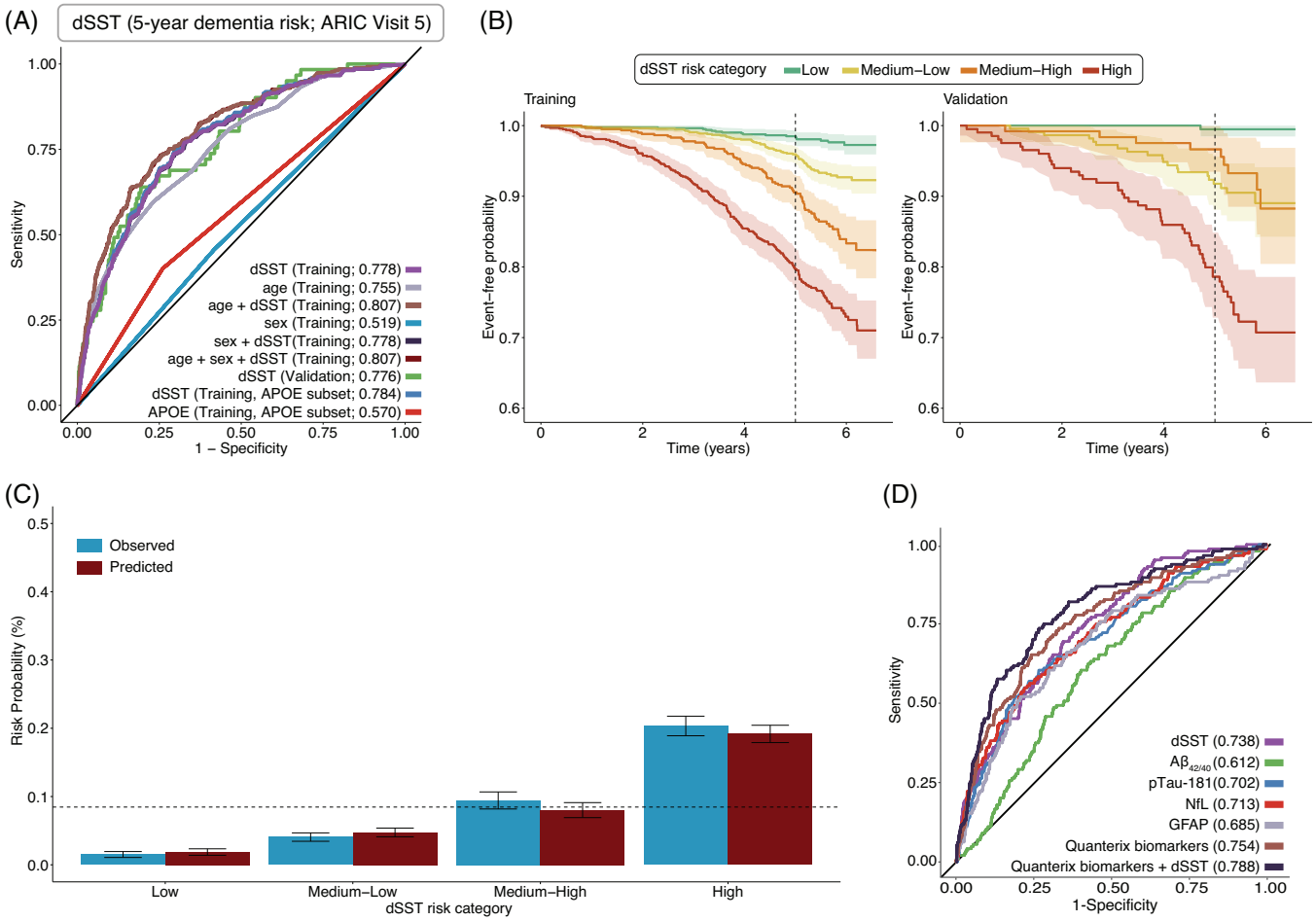

FIGURE 3 The dSST predicts 5-year dementia risk in the ARIC study. (A) Using blood drawn from Visit 5 (2011–2013) and dementia follow-up through Visit 6 (2016–2017), ROC curves show the AUCs of dSST scores in discriminating 5-year dementia risk across training and validation data sets, as well as performance of age and *APOE* genotype. (B) Kaplan–Meier plots show the observed 5-year dementia event-free probabilities grouped according to dSST risk categories (*low*, *medium-low*, *medium-high*, and *high*) across training and validation data sets using proteomic data from Visit 5. Event-free probabilities were calculated across the observed times for each risk group. (C) Bar charts show the observed and predicted 5-year dementia event rate probabilities according to dSST risk categories (*low*, *medium-low*, *medium-high*, and *high*) in the training data set. Predicted event rates were calculated as the mean dSST score for each risk group; the dotted line reflects average risk. (D) ROC curves show the AUC of dSST scores and ADRD plasma biomarker levels ($A\beta_{42/40}$, GFAP, NfL, p-tau-181) in discriminating 5-year dementia risk. Results derived from Cox proportional hazard regression models. Aβ, amyloid beta; *APOE*, apolipoprotein E; ARIC, Atherosclerosis Risk in Communities; AUC, area under the curve; dSST, Dementia SomaSignal Test; GFAP, glial fibrillary acidic protein; NfL, neurofilament light; pTau, phosphorylated tau; ROC, receiver-operating characteristic.

dSST risk categories reliably stratified clinically relevant differences in dementia risk, with 5-year observed event rates of 1.5% in *low*, 4.1% in *medium-low*, 9.4% in *medium-high*, and 20.3% in *high* (Figure 3B). The dSST was well calibrated (Figure 3C; Figure S2C). The relative difference in risk between the *low*- and *high*-risk categories was 13.47-fold (21.17-fold after adding age to the model), compared to a 16.50-fold and 1.78-fold difference associated with age and *APOE*, respectively. The dSST outperformed each plasma biomarker in discriminating 5-year dementia risk, but this difference in performance was only statistically significant for $A\beta_{42/40}$. Similar to ARIC Visit 3, the dSST significantly improved predictive accuracy when used in combination with plasma biomarkers (Quanterix biomarkers AUC: 0.754 vs Quanterix biomarkers + dSST AUC: 0.788) (Figure 3D; Figure S2D; Table S7).

For ARIC participants with blood samples collected at Visits 3 and 5 (n = 4,480; median age 57 and 75 at Visits 3 and 5, respectively), 44.4% retained the same dSST risk category over 20 years between mid-life and late life, with over 65% of participants in the *low*- and *high*-risk groups maintaining the same risk category (Figure 4; Table S8). Using 102 blind duplicates to assess the stability of risk classification in the BLSA, we observed no participant switch from *low*- to *high*-risk groups, or from *high*- to *low*-risk groups. ARIC participants who shifted to higher risk categories between mid-life and late life showed greater dementia risk over the subsequent 5-year follow-up, whereas participants who shifted to lower risk categories showed attenuated subsequent dementia risk during this time. For descriptive purposes, baseline (ARIC Visit 3, ARIC Visit 5) levels of individual proteins used in dSST calculations are shown in boxplots across participants who did

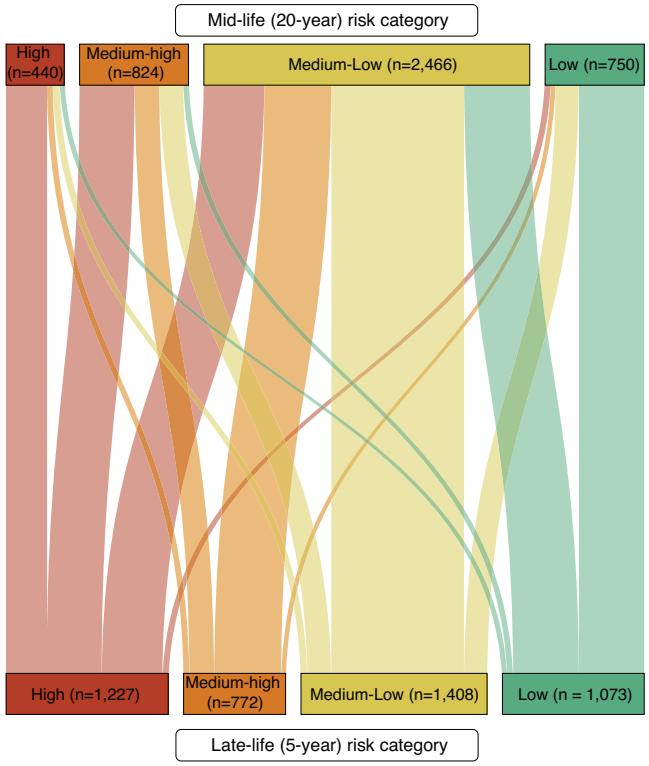

FIGURE 4 Changes in dSST risk categories over time in the ARIC study. A Sankey diagram shows the distribution of ARIC participants who retained the same dSST risk category across mid- and late-life, and for those ARIC participants who did shift between categories, the distribution of categories to which such participants transitioned. Mid-life dementia risk was estimated using blood draws from Visit 3 (1993–1995) and dementia follow-up through Visit 5 (2011–2013), and late-life dementia risk was estimated using blood draws from Visit 5 (2011–2013) and dementia follow-up through Visit 6 (2016–2017). ARIC, the Atherosclerosis Risk in Communities; dSST, Dementia SomaSignal Test.

and did not develop dementia over the subsequent follow up periods (Figures S3, 4).

3.2 | dSST predicts 20-year dementia risk in NILS-LSA

To determine the generalizability of the dSST, its performance was assessed in an external cohort based in Japan ($n = 340$; median age 67; Table S9). The dSST in the NILS-LSA cohort yielded an AUC of 0.677 for predicting 20-year dementia risk (Figure 5A;Table S6). Although age (AUC: 0.741) outperformed the dSST, this difference in performance was not statistically significant. The combination of the dSST and age achieved a significantly better AUC (0.751) than the dSST alone. dSST+age displayed a higher AUC than age alone; however, this difference was not statistically significant. The dSST demonstrated significantly superior discrimination compared to *APOE* alone (AUC: 0.550). In sex-stratified analyses, model performance was similar and did not differ from performance when female and male data were combined. The combination of the dSST, age, and sex yielded the same performance compared to the age+dSST model (AUC: 0.751). Using a dSST cutoff score that predicted dementia status with a maximized Youden Index at ARIC Visit 3 instead of NILS-LSA did not substantially alter sensitivity and specificity, suggesting that cutoff scores applied to prediction of 20-year dementia risk among middle-aged adults are comparable across these two cohorts. Due to limited sample size in the *low*-risk category, three risk categories were used in NILS-LSA analyses: *low/medium-low*, *medium-high*, and *high*. dSST risk categories stratified clinically relevant differences in dementia risk, with 20-year observed event rates of 29.0% in *low/medium-low*, 48.7% in *medium-high*, and 63.4% in *high* (Figure 5B). The relative difference in risk between the *low/medium-low* and *high*-risk categories was 2.20-fold (4.77-fold after adding age to the model), compared to a 2.86-fold and 1.41-fold difference associated with age and *APOE*, respectively.

3.3 | dSST is related to greater domain-specific cognitive decline in BLSA

The association between the dSST and cognitive decline was examined in the BLSA using linear mixed-effects regression models ($n = 1,214$; median age 71; Table S10). The average follow-up time was 6.7 (SD = 3.5) years with an average of 4.5 (SD = 2.5) assessments per participant (range: 2 to 22). At baseline, higher dSST scores and risk categories were associated with significantly lower performance across each cognitive domain and the MMSE (Figure S5A,B; Table S11). In longitudinal analyses, higher dSST scores were also associated with significantly greater decline in each cognitive domain and the MMSE (Figure S5C). A 1 SD higher dSST score was associated with 0.272 lower verbal memory scores (*z*-score) at baseline and an additional decrease of 0.020 per year, whereas a 1 SD higher age was associated with 0.024 lower verbal memory scores at baseline and an additional decrease of 0.002 per year. The *high* dSST risk category predicted accelerated declines in all cognitive domains (Figure S5D).

3.4 | dSST is linked to accelerated brain atrophy in BLSA

We then asked how the dSST related to brain volume loss in the BLSA using similar models ($n = 968$; median age 68; Table S10). The average follow-up time was 5.4 (SD = 2.3) years with an average of 3.5 (SD = 1.5) scans per participant (range: 2 to 11). Higher dSST scores and risk categories (i.e., using the *low* dSST risk category as the reference) were associated with lower brain volumes at baseline and greater rates of atrophy in total brain, GM, and WM volumes, as well as in each lobar region. With higher dSST risk categories, we observed step-wise effects on brain structure at baseline, but not on longitudinal rates of change (Figure 6A; Table S12). A 1 SD higher dSST score was associated with 46.29 cm^3 lower total brain volume at baseline and an additional decrease of 0.37 cm^3 per year, whereas a 1 SD higher age was associated with 3.28 cm^3 lower total brain volume at baseline

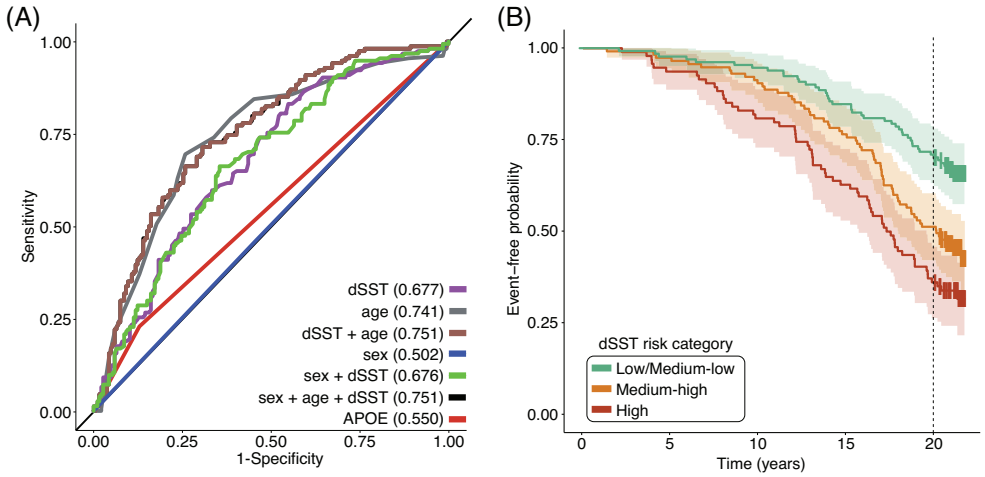

FIGURE 5 The dSST predicts 20-year dementia risk in the NILS-LSA. (A) Using blood drawn in 2000 and dementia follow-up through 2022, ROC curves show the AUC of dSST scores in discriminating 20-year dementia risk, as well as performance of age and *APOE* genotype. (B) Kaplan–Meier plots show the observed 20-year dementia event-free probabilities grouped according to dSST risk categories (*low/medium-low*, *medium-high*, and *high*). Event-free probabilities were calculated across the observed times for each risk group. *APOE*, apolipoprotein E; AUC, area under the curve; dSST, Dementia SomaSignal Test; NILS-LSA, National Institute for Longevity Sciences-Longitudinal Study of Aging; ROC, receiver-operating characteristic.

and an additional decrease of 0.05 cm^3 per year. Voxel-based analyses showed that the dSST was most strongly associated with atrophy in medial-temporal regions (Figure 6B; Figure S6;Table S13). Higher dSST scores and risk categories (i.e., using the *low* dSST risk category as the reference) were also associated with higher baseline SPARE-AD scores and accelerated SPARE-AD increases, reflecting accelerated brain volume loss in brain regions vulnerable atrophy in AD.[28,33] With higher dSST risk categories, we observed step-wise effects on SPARE-AD scores at baseline and longitudinal rates of change (Figure 6C; Table S12).

3.5 | dSST is associated with PET and plasma ADRD biomarkers in BLSA

Next, we examined the dSST's associations with PET measures of Aβ (n = 183; median age 76) and tau (n = 60; median age 76), as well as plasma markers of $A\beta_{42/40}$, GFAP, and NfL (n = 749; median age 69) and p-tau-181 (n = 664; median age 67; Table S10) using logistic or linear regression models. A 1 SD higher dSST score elevated risk of Aβ-positive PET status by nearly 2-fold (odds ratio = 1.99; Figure 7A; Table S14). The dSST (AUC: 0.724) enhanced discrimination of Aβ PET status compared to age (AUC: 0.605), *APOE* (AUC: 0.527), and plasma biomarkers alone (Quanterix biomarkers AUC: 0.877 vs Quanterix biomarkers + dSST AUC: 0.895). As expected, AD-specific biomarkers as well as GFAP (which has been closely tied to cortical amyloid[46]) outperformed the dSST for prediction of Aβ-positive status (Figure 7B). Non-significant, positive associations were observed between dSST scores and tau-PET levels in the ERC and ITG. Among the four plasma ADRD biomarkers, the dSST showed the strongest association with NfL (adjusted R^2 = 0.247), a non-specific marker of neuronal injury (Figure 7C;Table S14); this association may, in part, be influenced by SomaScan NfL's inclusion in dSST calculations (Table S4).

3.6 | dSST protein characterization

For the 25 proteins used to develop the dSST, literature-based evidence indicated that 20 were implicated previously in dementia or dementia-related neuropathology, and revealed their contributions to neuroinflammation, blood–brain barrier (BBB) dysfunction, and altered Aβ metabolism (Table S15). The levels of 9 proteins were significantly correlated across plasma and CSF (Table S16), and 11 cognate genes encoding proteins were highly expressed in at least one CNS cell type (Table S17), suggesting that many proteins used to develop the dSST are sensitive markers of neurologic health. Results from post-mortem tissue analyses showed that every cognate gene with available transcriptomic data was differentially expressed at the RNA level in AD brains, with another three also differentially expressed at the protein level (Table S18). In neurovascular cell types, patterns of AD-related expression were most evident in endothelial cells (capillary and arterial), oligodendrocytes, and extracellular matrix-regulating pericytes (Table S19).

4 | DISCUSSION

Using data from multiple large cohorts, the current study applied high-throughput proteomic assays and machine learning techniques to develop and validate a blood-based prognostic aid for dementia risk. The dSST predicted mid-life dementia risk over a 20-year follow-up across two independent cohorts with different ethnic backgrounds

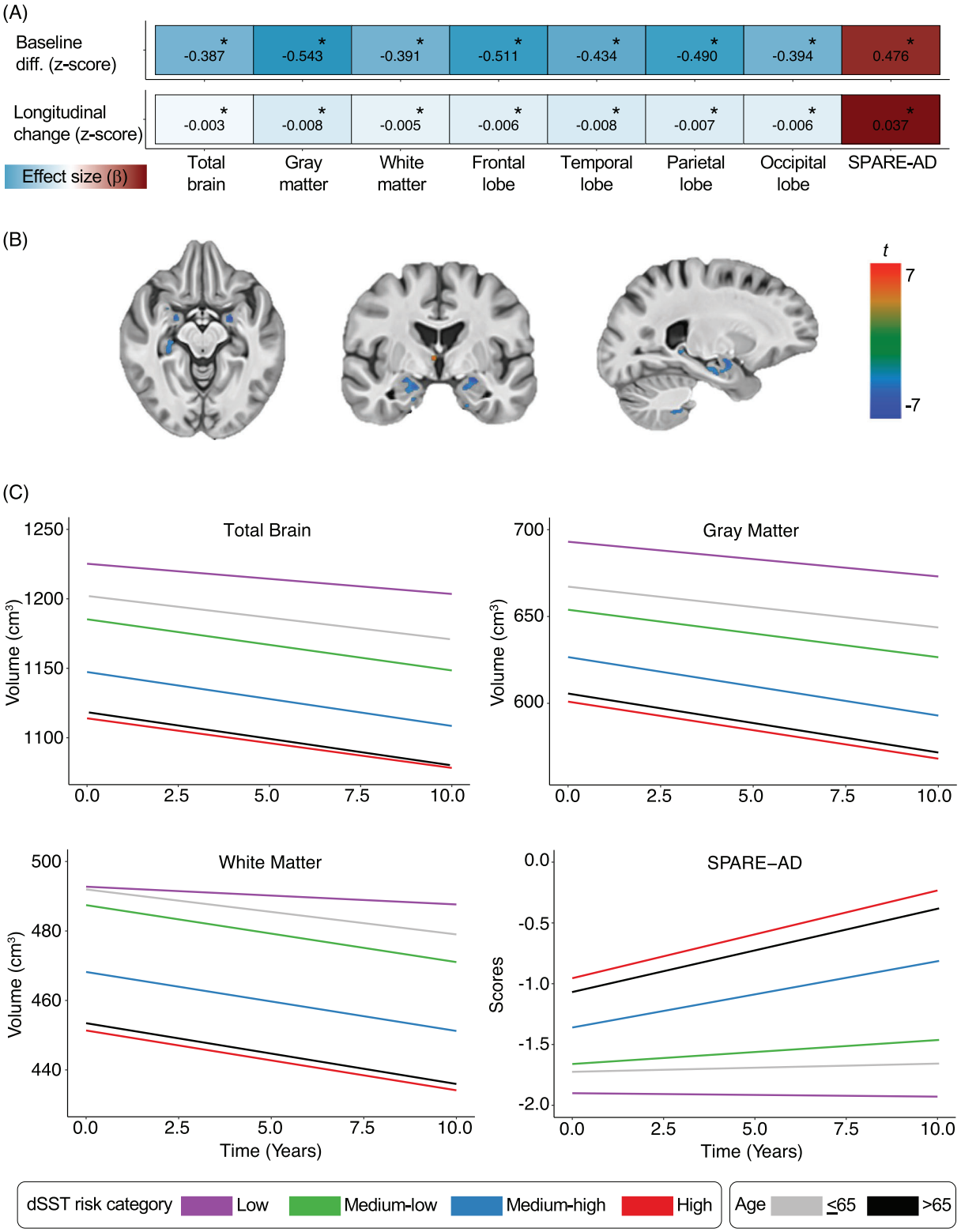

FIGURE 6 The dSST relates to differences in baseline and longitudinal brain atrophy in the BLSA. (A) Heatmaps show the associations of dSST scores with cross-sectional differences and longitudinal rates of change in standardized regional brain volumes and SPARE-AD scores. *Statistically significant ($p < 0.05$). (B) Axial, coronal, and sagittal images show the associations of dSST scores with voxel-wise longitudinal differences in gray matter volumes. A threshold of 50 voxels with an uncorrected $p < 0.001$ was used to define significant clusters. (C) Line graphs show the associations of dSST risk categories (relative to the low-risk category) with cross-sectional differences and longitudinal rates of change in total brain volume (cm^3), total gray matter volume (cm^3), total white matter volume (cm^3), and SPARE-AD scores. Differences in baseline and annual rates of change in brain volumes, voxels, and SPARE-AD scores associated with dSST scores, and its risk categories were derived from linear mixed-effects regression models. dSST, Dementia SomaSignal Test; BLSA, Baltimore Longitudinal Study of Aging; SPARE-AD, Spatial Pattern of Atrophy for Recognition of Alzheimer's Disease.

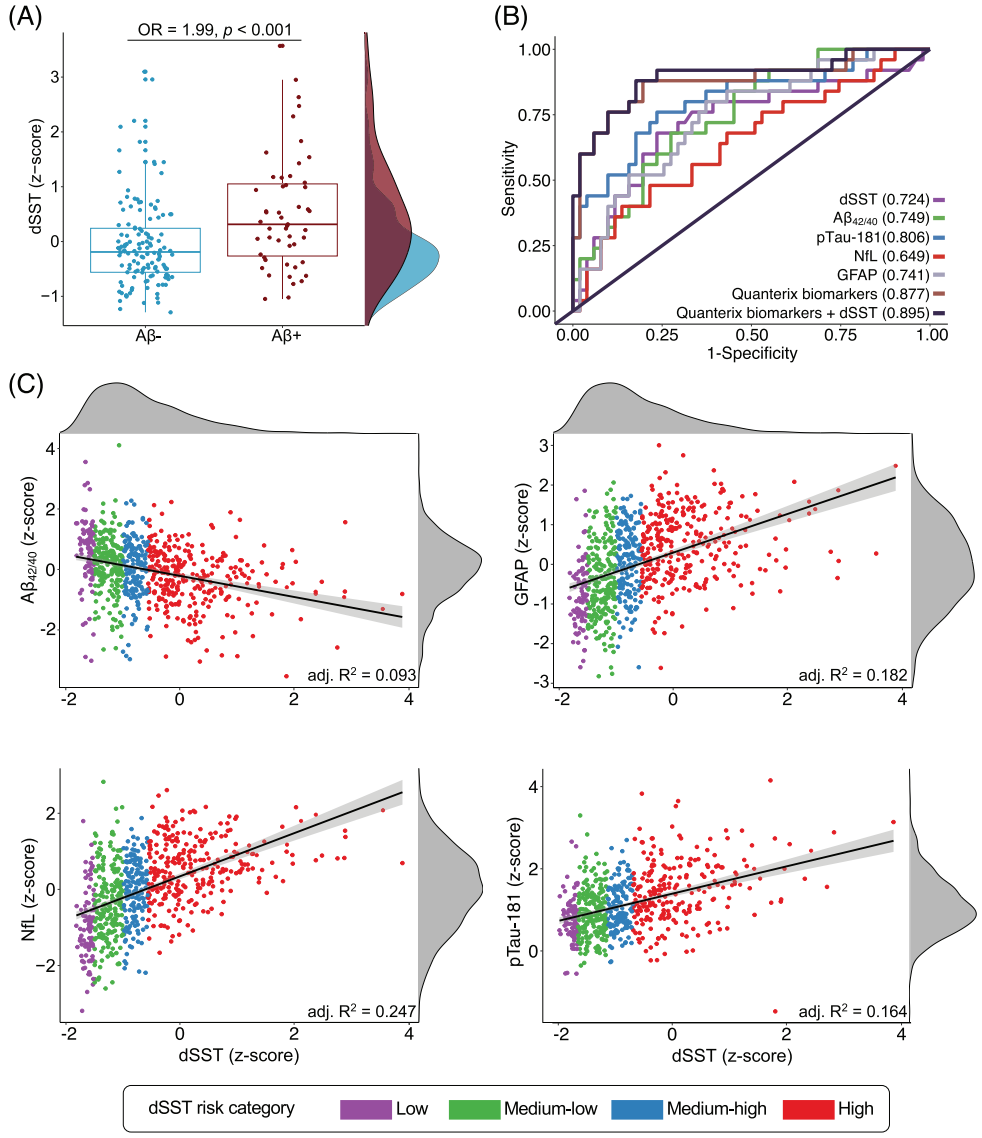

FIGURE 7 The dSST relates to differences in biomarkers in the BLSA. (A) Boxplot (and corresponding density plots along *y*-axis) shows the distribution of dSST scores across amyloid-negative (Aβ−) and amyloid-positive (Aβ+) PET participants. Results derived from logistic regression models. (B) ROC curves show the AUC of dSST scores and ADRD plasma biomarker levels (A$\beta_{42/40}$, GFAP, NfL, and p-tau181) in discriminating Aβ PET status (±). Results derived from logistic regression models. (C) Scatterplots and lines of best fit show differences in A$\beta_{42/40}$, GFAP, NfL, and p-tau181 levels associated with dSST scores. Density plots along the y-axes display the distribution of ADRD plasma biomarker levels. Density plots along the *x*-axis display the distribution of dSST scores. Results derived from linear regression models. Aβ, amyloid beta; ADRD, Alzheimer's disease and related dementias; AUC, area under the curve; BLSA, Baltimore Longitudinal Study of Aging; dSST, Dementia SomaSignal Test; GFAP, glial fibrillary acidic protein; NfL, neurofilament light; OR, odds ratio; PET, position emission tomography; p-tau, phosphorylated tau; ROC, receiver-operating characteristic.

(AUCs: dSST 0.68–0.70, dSST+age 0.75–0.81), and displayed even greater performance for late-life dementia risk over a 5-year follow-up (AUCs: dSST 0.78, dSST+age 0.81). The combination of the dSST with age offered improved prediction compared to age or the dSST alone. Along with higher AUCs compared to each ADRD plasma biomarker (A$\beta_{42/40}$, p-tau181, GFAP, and NfL) in discriminating 20- and 5-year dementia risk, the dSST improved predictive accuracy when used in combination with such biomarkers. In a separate cohort, the dSST also predicted decline across multiple cognitive domains, accelerated brain atrophy, and elevated measures of neuropathology. Our findings suggest that the dSST is a minimally invasive prognostic aid that reliably quantifies risk up to two decades before dementia onset.

In contrast to other biomarkers, which estimate AD-specific dementia over shorter follow-up times, the dSST was developed to estimate dementia risk over multiple decades preceding the typical age of dementia onset in a manner that is agnostic to etiology. Like the dSST, the cardiovascular risk factors, aging, and incience of dementia (CAIDE) risk score was developed to predict 20-year dementia risk with AUCs ranging from 0.65 to 0.77.[47,48] Unlike the dSST, this measure requires 15 different clinical variables (e.g., low-density

lipoprotein [LDL]/high-density lipoprotein [HDL] levels) that may not be available outside the context of a comprehensive medical workup. A recent UK Biobank study used the Olink proteomic panel (≈1,400 proteins) in a subset of participants (n = 683) to select 10 plasma proteins (including NfL and GFAP) that together predict 10-year AD dementia risk with an AUC of 0.80,[49] but it is unclear whether this model's predictive accuracy generalizes to other (more diverse) cohorts, longer follow-up windows, or to non-AD dementias. Blood NfL levels are elevated in several neurodegenerative conditions and can predict progression from MCI to all-cause and AD dementia over the near-term (<6 years) with good accuracy (AUCs >0.78).[9] The dSST outperformed NfL (as well as all other ADRD biomarkers) in predicting dementia risk across both 20-year and 5-year timescales. The dSST's higher AUCs compared to Quanterix's $A\beta_{42/40}$ and p-tau181 might be expected given that these biomarkers were developed specifically to detect AD, whereas our community-based sample of all-cause dementia cases used for dSST development likely contained participants with more heterogenous neuropathologies.[50]

Although the dSST showed similar performance for discriminating 20-year dementia risk in two independent cohorts, we did observe large differences in the dynamic range of *low* versus *high*-risk bins across ARIC and NILS-LSA. This discrepancy may be attributed to differences in study design or participant characteristics. With respect to chronological age, the dSST either showed superior prediction or its addition accounted for unique variance (and therefore clinical utility) not otherwise captured by this leading dementia risk factor. We also showed that individuals who lowered their dSST scores across mid-life to late life displayed a corresponding attenuation in dementia risk, indicating that the dSST may be leveraged as a surrogate marker to track response to interventions. Because of such potential to be an independent, modifiable, surrogate marker, we reported both the dSST's stand-alone performance as well as its performance in combination with basic demographic factors (i.e., age, sex), which may enhance prediction, but are not readily modifiable. Although the dSST quantified 20-year dementia risk across two large cohorts, we also demonstrated its capacity to predict a range of neurocognitive outcomes in a third cohort. The dSST's sensitivity to declines in cognitive performance across a variety of domains indicates its potential utility for detecting typical and atypical (e.g., non-amnestic) dementia variants.[51] Similarly, the dSST's robust associations with accelerated atrophy across brain regions and tissue types (GM/WM) suggests its utility for detecting dementia subtypes that are characterized by different patterns of neurodegeneration (e.g., limbic-predominant AD[52]).

Our study has several strengths, including the use of independent, large cohorts with multi-decade follow-up periods, state-of-the-art proteomics, and longitudinal neuroimaging. However, our study also has several limitations. First, we prioritized dementia-risk prediction without the use of demographic information or clinical characteristics. Although improved prediction would almost certainly be achievable by adding these features, we intentionally focused on developing a protein-only model to maximize clinical utility and scalability, while minimizing costs. Second, because the dSST was trained on long-term incident all-cause dementia, its ability to detect other cognitive phenotypes over shorter follow-up times (e.g., mild cognitive impairment) and to distinguish different dementia subtypes (e.g., vascular dementia vs AD dementia) may be limited. Third, because our study relied predominantly on data from participants of European and Asian ancestry, future investigations will be needed to determine the generalizability of our results to more diverse cohorts. Despite these considerations, our results address an unmet medical need in the form of a blood-based prognostic aid that predicts dementia-risk during its multi-decade preclinical/prodromal phase and independent of a specific etiology.

ACKNOWLEDGMENTS

We thank the staff and participants of the Atherosclerosis Risk in Communities (ARIC) study for their important contributions, as well as the National Institue for Longevity Sciences-Longitudinal Study of Aging (NILS-LSA) and Baltimore Longitudinal Study of Aging (BLSA) participants and staff for their participation and continued dedication. This research was supported by the Intramural Research Program (IRP) of the National Institutes of Health (NIH), National Institute on Aging (NIA). ARIC is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute (NHLBI) contracts (75N92022D00001, 75N92022D00002, 75N92022D00003, 75N92022D00004, 75N92022D00005). The ARIC Neurocognitive Study is additionally supported by U01HL096812, U01HL096814, U01HL096899, U01HL096902, and U01HL096917 from the NIH (NHLBI, NIA, NINDS (National Institute of Neurological Disorders and Stroke) and NIDCD (National Institute on Deafness and Other Communication Disorders). NILS-LSA was supported by contracts between NEC Solution Innovators Limited and the National Center for Geriatrics and Gerontology (21-18) and Nagoya University (2020-0628).

CONFLICT OF INTEREST STATEMENT

C.P., K.L., M.S., H.B., and S.A.W. are current or former employees of SomaLogic Operating Co., Inc, and/or Standard BioTools. N.K., M.S., S.K., M.F., and I.W. are current employees of NEC Solution Innovators Limited and/or FonesLife Corporation. The remaining authors declare no conflicts of interest. Author disclosures are available in the Supporting Information.

DATA AVAILABILITY STATEMENT

All data generated in the current study are included in this article (or its Supplementary Materials), available upon reasonable request, or available in an online public repository. Atherosclerosis risk in communities (ARIC) study proteomic data is available through the National Heart Lung, and Blood Institute (NHLBI) Biologic Specimen and Data Repository Information Coordinating Center (https://biolincc.nhlbi.nih.gov/studies/aric/). Additional requests for clinical or proteomic data may be submitted to the ARIC Steering Committees and will be reviewed to ensure that data can be shared without compromising participant confidentiality or breaching intellectual property restrictions. Participant-level demographic, clinical, and proteomic data may be partially restricted based on prior participant consent, and data sharing restrictions may also be applied to ensure consistency with confidentiality or privacy laws and considerations

(https://sites.cscc.unc.edu/aric/). Anonymized baltimore longitudinal study of aging (BLSA) data not published within this article may be shared upon request from qualified investigators. Researchers who wish to use BLSA data are encouraged to develop a pre-analysis plan that can be submitted for approval (https://blsa.nia.nih.gov/how-apply). National Institute for Longevity Sciences-Longitudinal Study of Aging (NILS-LSA) data are available from the authors (iwa-waga@nec.com) upon reasonable request and with permission of NILS-LSA investigators (https://www.ncgg.go.jp/research/lab/cgss/department/ep/index.html); as the data are under license for the current study, some restrictions to data sharing may apply.

CONSENT STATEMENT

All participants provided written informed consent.

ORCID

Michael R. Duggan https://orcid.org/0000-0002-1029-4423

REFERENCES

1. Sims JR, Zimmer JA, Evans CD, et al. Donanemab in early symptomatic Alzheimer disease: the TRAILBLAZER-ALZ 2 randomized clinical trial. *JAMA*. 2023;330(6):512-527. doi:10.1001/jama.2023.13239
2. Dubois B, Epelbaum S, Nyasse F, et al. Cognitive and neuroimaging features and brain β-amyloidosis in individuals at risk of Alzheimer's disease (INSIGHT-preAD): a longitudinal observational study. *Lancet Neurol*. 2018;17(4):335-346. doi:10.1016/S1474-4422(18)30029-2
3. Sturchio A, Dwivedi AK, Young CB, et al. High cerebrospinal amyloid-β 42 is associated with normal cognition in individuals with brain amyloidosis. *EClinicalMedicine*. 2021;38. doi:10.1016/j.eclinm.2021.100988
4. Smith R, Hägerström D, Pawlik D, et al. Clinical utility of tau positron emission tomography in the diagnostic workup of patients with cognitive symptoms. *JAMA Neurol*. 2023. doi:10.1001/jamaneurol.2023.1323
5. Boyle PA, Wang T, Yu L, et al. To what degree is late life cognitive decline driven by age-related neuropathologies?. *Brain*. 2021;144(7):2166-2175. doi:10.1093/brain/awab092
6. Wilkins CH, Windon CC, Dilworth-Anderson P, et al. Racial and ethnic differences in amyloid PET positivity in individuals with mild cognitive impairment or dementia: a secondary analysis of the Imaging Dementia-Evidence for Amyloid Scanning (IDEAS) cohort study. *JAMA Neurol*. 2022;79(11):1139-1147. doi:10.1001/jamaneurol.2022.3157
7. Barnes LL, Leurgans S, Aggarwal NT, et al. Mixed pathology is more likely in black than white decedents with Alzheimer dementia. *Neurology*. 2015;85(6):528-534. doi:10.1212/wnl.0000000000001834
8. Palmqvist S, Tideman P, Cullen N, et al. Prediction of future Alzheimer's disease dementia using plasma phospho-tau combined with other accessible measures. *Nat Med*. 2021;27(6):1034-1042. doi:10.1038/s41591-021-01348-z
9. Silva-Spínola A, Lima M, Leitão MJ, et al. Blood biomarkers in mild cognitive impairment patients: relationship between analytes and progression to Alzheimer disease dementia. *Eur J Neurol*. 2023;30(6):1565-1573. doi:10.1111/ene.15762
10. Ashton NJ, Brum WS, Di Molfetta G, et al. Diagnostic accuracy of a plasma phosphorylated tau 217 immunoassay for Alzheimer disease pathology. *JAMA Neurol*. 2024;81(3):255-263. doi:10.1001/jamaneurol.2023.5319
11. Bateman RJ, Xiong C, Benzinger TL, et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. *N Engl J Med*. 2012;367:795-804.
12. Walker KA, Chen J, Zhang J, et al. Large-scale plasma proteomic analysis identifies proteins and pathways associated with dementia risk. *Nat Aging*. 2021;1(5):473-489. doi:10.1038/s43587-021-00064-0
13. Wright JD, Folsom AR, Coresh J, et al. The ARIC (Atherosclerosis Risk In Communities) study: jACC focus seminar 3/8. *J Am Coll Cardiol*. 2021;77(23):2939-2959. doi:10.1016/j.jacc.2021.04.035
14. Tin A, Yu B, Ma J, et al. Reproducibility and variability of protein analytes measured using a multiplexed modified aptamer assay. *J Appl Lab Med*. 2019;4(1):30-39. doi:10.1373/jalm.2018.027086
15. Lu Y, Pike JR, Chen J, et al. Changes in Alzheimer disease blood biomarkers and associations with incident all-cause dementia. *Jama*. 2024;332(15):1258-1269. doi:10.1001/jama.2024.6619
16. Knopman DS, Gottesman RF, Sharrett AR, et al. Mild cognitive impairment and dementia prevalence: the Atherosclerosis Risk in Communities Neurocognitive study (ARIC-NCS). *Alzheimers Dement*. 2016;2:1-11. doi:10.1016/j.dadm.2015.12.002
17. Shimokata H, Ando F, Niino N. A new comprehensive study on aging–the National Institute for Longevity Sciences, Longitudinal Study of Aging (NILS-LSA). *J Epidemiol*. 2000;10(1):S1-S9. doi:10.2188/jea.10.1sup_1
18. Ikegami N. Public long-term care insurance in Japan. *JAMA*. 1997;278(16):1310-1314. doi:10.1001/jama.1997.03550160030017
19. Otsuka R, Zhang S, Ihira H, et al. Dietary diversity and risk of late-life disabling dementia in middle-aged and older adults. *Clin Nutr*. 2023;42(4):541-549. doi:10.1016/j.clnu.2023.02.002
20. Matsui Y, Tanizaki Y, Arima H, et al. Incidence and survival of dementia in a general population of Japanese elderly: the Hisayama study. *J Neurol Neurosurg Psychiatry*. 2009;80(4):366-370. doi:10.1136/jnnp.2008.155481
21. Noda H, Yamagishi K, Ikeda A, Asada T, Iso H. Identification of dementia using standard clinical assessments by primary care physicians in Japan. *Geriatr Gerontol Int*. 2018;18(5):738-744. doi:10.1111/ggi.13243
22. Shock NW, Greulich RC, Aremberg D, Costa PT, Lakatta EG, Tobin JD. *Normal Human Aging: The Baltimore Longitudinal Study of Aging*. National Institutes of Health; 1984.
23. Kawas C, Gray S, Brookmeyer R, Fozard J, Zonderman A. Age-specific incidence rates of Alzheimer's disease. *Neurology*. 2000;54:2072-2077. doi:10.1212/wnl.54.11.2072<bib>
24. Candia J, Daya GN, Tanaka T, Ferrucci L, Walker KA. Assessment of variability in the plasma 7k SomaScan proteomics assay. *Sci Rep*. 2022;12(1):17147. doi:10.1038/s41598-022-22116-0
25. Candia J, Fantoni G, Delgado-Peraza F, et al. Variability of 7 and 11K SomaScan Plasma Proteomics Assays. *J Proteome Res*. 2024;23(12):5531-5539. doi:10.1021/acs.jproteome.4c00667
26. Varadaraj V, Munoz B, Deal JA, et al. Association of vision impairment with cognitive decline across multiple domains in older adults. *JAMA Netw Open*. 2021;4(7):e2117416-e2117416. doi:10.1001/jamanetworkopen.2021.17416
27. Duggan MR, Peng Z, An Y, et al. Herpes viruses in the Baltimore Longitudinal Study of Aging: associations with brain volumes, cognitive performance, and plasma biomarkers. *Neurology*. 2022;99(18):e2014-e2024. doi:10.1212/wnl.0000000000201036
28. Resnick SM, Pham DL, Kraut MA, Zonderman AB, Davatzikos C. Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. *J Neurosci*. 2003;23(8):3295-3301. doi:10.1523/jneurosci.23-08-03295.2003
29. Doshi J, Erus G, Ou Y, et al. MUSE: mUlti-atlas region segmentation utilizing ensembles of registration algorithms and parameters, and locally optimal atlas selection. *Neuroimage*. 2016;127:186-195. doi:10.1016/j.neuroimage.2015.11.073
30. Davatzikos C, Genc A, Xu D, Resnick SM. Voxel-based morphometry using the RAVENS maps: methods and validation using simulated

longitudinal atrophy. *Neuroimage*. 2001;14(6):1361-1369. doi:10.1006/nimg.2001.0937
31. Driscoll I, Beydoun MA, An Y, et al. Midlife obesity and trajectories of brain volume changes in older adults. *Hum Brain Mapp*. 2012;33(9):2204-2210. doi:10.1002/hbm.21353
32. Lee S, Zipunnikov V, Reich DS, Pham DL. Statistical image analysis of longitudinal RAVENS images. Methods. *Front Neurosci*. 2015;9:368. doi:10.3389/fnins.2015.00368
33. Davatzikos C, Xu F, An Y, Fan Y, Resnick SM. Longitudinal progression of Alzheimer's-like patterns of atrophy in normal older adults: the SPARE-AD index. *Brain*. 2009;132(pt 8):2026-2035. doi:10.1093/brain/awp091
34. Davatzikos C, Bhatt P, Shaw LM, Batmanghelich KN, Trojanowski JQ. Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. *Neurobiol Aging*. 2011;32(12):2322.e19-e27. doi:10.1016/j.neurobiolaging.2010.05.023
35. Fan Y, Shen D, Gur RC, Gur RE, Davatzikos C. COMPARE: classification of morphological patterns using adaptive regional elements. *IEEE Trans Med Imaging*. 2006;26(1):93-105.
36. Davatzikos C, Resnick SM, Wu X, Parmpi P, Clark CM. Individual patient diagnosis of AD and FTD via high-dimensional pattern classification of MRI. *Neuroimage*. 2008;41(4):1220-1227. doi:10.1016/j.neuroimage.2008.03.050
37. Toledo JB, Weiner MW, Wolk DA, et al.; Alzheimer's Disease Neuroimaging Initiative. Neuronal injury biomarkers and prognosis in ADNI subjects with normal cognition. *Acta Neuropathol Commun*. 2014;2(1):1-9.
38. Da X, Toledo JB, Zee J, et al.; Alzheimer's Neuroimaging Initiative. Integration and relative value of biomarkers for prediction of MCI to AD progression: spatial patterns of brain atrophy, cognitive scores, APOE genotype and CSF biomarkers. *NeuroImage Clin*. 2014;4:164-173.
39. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20(3):273-297. doi:10.1007/BF00994018
40. Vapnik V. *The Nature of Statistical Learning Theory*. Springer science & business media; 1999.
41. Bilgel M, Bannerjee A, Shafer A, An Y, Resnick SM. Vascular risk is not associated with PET measures of Alzheimer's disease neuropathology among cognitively normal older adults. *Neuroimage Rep*. 2021;1(4). doi:10.1016/j.ynirp.2021.100068
42. Dammer EB, Ping L, Duong DM, et al. Multi-platform proteomic analysis of Alzheimer's disease cerebrospinal fluid and plasma reveals network biomarkers associated with proteostasis and the matrisome. *Alzheimers Res Ther*. 2022;14(1):174. doi:10.1186/s13195-022-01113-5
43. Yang AC, Vest RT, Kern F, et al. A human brain vascular atlas reveals diverse mediators of Alzheimer's risk. *Nature*. 2022;603(7903):885-892. doi:10.1038/s41586-021-04369-3
44. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Stat Soc Ser B Stat Methodol*. 2005;67(2):301-320. doi:10.1111/j.1467-9868.2005.00503.x
45. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization paths for Cox's Proportional Hazards Model via coordinate descent. *J Stat Softw*. 2011;39(5):1-13. doi:10.18637/jss.v039.i05
46. Benedet AL, Milà-Alomà M, Vrillon A, et al. Differences between plasma and cerebrospinal fluid glial fibrillary acidic protein levels across the Alzheimer Disease continuum. *JAMA Neurol*. 2021;78(12):1471-1483. doi:10.1001/jamaneurol.2021.3671
47. Kivipelto M, Ngandu T, Laatikainen T, Winblad B, Soininen H, Tuomilehto J. Risk score for the prediction of dementia risk in 20 years among middle aged people: a longitudinal, population-based study. *Lancet Neurol*. 2006;5(9):735-741. doi:10.1016/S1474-4422(06)70537-3
48. Chosy EJ, Edland SD, Gross N, et al. The CAIDE dementia risk score and the Honolulu-Asia Aging Study. *Dement Geriatr Cogn Disord*. 2019;48(3-4):164-171. doi:10.1159/000504801
49. Danni AG, Robert FH, Zhana K, et al. Blood protein levels predict leading incident diseases and mortality in UK Biobank. *medRxiv*. 2023. doi:10.1101/2023.05.01.23288879
50. Dark HE, Duggan MR, Walker KA. Plasma biomarkers for Alzheimer's and related dementias: a review and outlook for clinical neuropsychology. *Arch Clin Neuropsychol*. 2024;39(3):313-324. doi:10.1093/arclin/acae019
51. Graff-Radford J, Yong KXX, Apostolova LG, et al. New insights into atypical Alzheimer's disease in the era of biomarkers. *Lancet Neurol*. 2021;20(3):222-234. doi:10.1016/S1474-4422(20)30440-3
52. Daniel F, Agneta N, Eric W. Biological subtypes of Alzheimer disease. *Neurology*. 2020;94(10):436. doi:10.1212/WNL.0000000000009058

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Duggan MR, Paterson C, Lu Y, et al. The Dementia SomaSignal Test (dSST): A plasma proteomic predictor of 20-year dementia risk. *Alzheimer's Dement*. 2025;21:e14549. https://doi.org/10.1002/alz.14549