Benchmarking of a multi-biomarker low-volume panel for Alzheimer’s Disease and related dementia research

Laura Ibanez1,2,3*, Menghan Liu1,3*, Aleksandra Beric1,3*, Jigyasha Timsina1,3, Pat Kholfeld1,3, Kristy Bergmann1,3, Joey Lowery1,3, Nick Sykora1,3, Brenda Sanchez-Montejo1,3, Will Brock1,3, John P. Budde1,3, Randall J. Bateman2,4,5,6, Nicolas Barthelemy2,5, Suzanne E. Schindler2,4,6, David M Holtzman2,4,6, Tammie L. S. Benzinger2,4,6, Chengjie Xiong8, Rawan Tarawneh9, Krista Moulder2,6, John C. Morris2,6, Yun Ju Sung1,3, Carlos Cruchaga1,2,3,4,6,10*

1. Department of Psychiatry, Washington University School of Medicine
2. Department of Neurology, Washington University School of Medicine
3. NeuroGenomics and Informatics Center, Washington University School of Medicine
4. Hope Center for Neurologic Diseases, Washington University School of Medicine
5. The Tracy Family SILQ Center, Washington University School of Medicine
6. Knight Alzheimer Disease Research Center, Washington University School of Medicine
7. Department of Radiology, Washington University School of Medicine
8. Division of Biostatistics, Washington University School of Medicine
9. Department of Neurology, University of New Mexico School of Medicine
10. Department of Genetics, Washington University School of Medicine

*Co-corresponding authors
Laura Ibanez, PhD
ibanezl@wustl.edu
Carlos Cruchaga, PhD
cruchagac@wustl.edu

#Co-first authors

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract
Alzheimer’s Disease (AD) biomarker measurement is key to aid in the diagnosis and prognosis of the disease. In the research setting, participant recruitment and retention and optimization of sample use, is one of the main challenges that observational studies face. Thus, obtaining accurate established biomarker measurements for stratification and maximizing use of the precious samples is key. Accurate technologies are currently available for established biomarkers, mainly immunoassays and immunoprecipitation liquid chromatography-mass spectrometry (IP-MS), and some of them are already being used in clinical settings. Although some immunoassays- and IP-MS based platforms provide multiplexing for several different coding proteins there is not a current platform that can measure all the established and emerging biomarkers in one run. The NUcleic acid Linked Immuno-Sandwich Assay (NULISA™) is a mid-throughput platform with antibody-based measurements with a sequencing output that requires 15µL of sample volume to measure more than 100 analytes, including those typically assayed for AD. Here we benchmarked and compared the AD-relevant biomarkers including in the NULISA against validated assays, in both CSF and plasma. Overall, we have found that CSF measures of Aβ42/40, NfL, GFAP, and p-tau217 are highly correlated and have similar predictive performance when measured by immunoassay, mass-spectrometry or NULISA. In plasma, p-tau217 shows a performance similar to that reported with other technologies when predicting amyloidosis. Other established and exploratory biomarkers (total tau, p-tau181, NRGN, YKL40, sTREM2, VILIP1 among other) show a wide range of correlation values depending on the fluid and the platform. Our results indicate that the multiplexed immunoassay platform produces reliable results for established biomarkers in CSF that are useful in research settings, with the advantage of measuring additional novel biomarkers using minimal sample volume.
Background

Alzheimer’s Disease (AD) is characterized by the pathological accumulation of misfolded amyloid beta (Aβ) plaques, and hyperphosphorylated tau tangles [1, 2]. Thus, most efforts have been focused on developing accurate measurements of amyloid and tau species to assist in the diagnosis and pathological monitoring of the disease. The amount and distribution of amyloid and tau pathology can be quantified in using Positron Emission Tomography (PET) imaging following injection of specific radiotracers, which is commonly used as reference standard for AD pathology in living individuals [3, 4]. However, the use of fluid biomarkers is more convenient, less invasive, and cost-effective. Levels of some CSF analytes are highly correlated with amyloid and tau pathology while the performance in plasma assays vary widely, with only some demonstrating high correlation with amyloid pathology [3, 5-8].

In both CSF and plasma, immunoprecipitation liquid chromatography-mass spectrometry (IP-MS) is the technology that has provided the highest correlation values with amyloid and tau PET and outperformed other assays in classification of pathologically positive vs. negative patients [9-11]. Especially, p-tau measurements that not only provide good correlation with cognitive decline but also outperform other measurements [9, 10, 12, 13], even when measured in plasma. Despite the good performance and the capability to measure many species of immunopurified protein, IP-MS multiplexing (measuring multiple single gene-coding proteins in one run) remains a challenge. Additionally, IP-MS requires relatively high volumes per analyte, specialized facilities, uses specialized machinery, and highly trained personnel which makes it difficult to use in a research setting. The need of bridging samples across measuring rounds is a limitation for both IP-MS and immunoassays, it increases the use of biofluid and the harmonization of values when performing longitudinal analyses its complex.

The use of immunoassays, despite having slightly poorer performance compared to IP-MS, is widespread. CSF assays (Aβ40, Aβ42, tau phosphorylated at position 181 (p-tau181) and more recently 217 (p-tau217)) show good correlation with Aβ brain deposits (as measured with amyloid PET) and acceptable discriminatory power to discern between amyloid positive and negative individuals [14, 15], especially when combined with APOE genotype and/or age [16-19]. IP-MS assays, especially for p-tau217 have also show very high predictive power to identify amyloid positive individuals based on imaging [8]. However, as the levels of some of the proteins are lower in plasma, immunoassays have presented with some performance challenges.

Besides Aβ40, Aβ42 and tau species, there are other biomarkers that are of interest including neurofilament light chain protein (NFL), neurogranin (NGRN), visinin like protein 1 (VILIP-1), soluble triggering receptor expressed on myeloid cells 2 (sTREM2), Chitinase 3 Like 1 (YKL40), and synaptosomal-associated protein 25 (SNAP25). NFL, and NGRN have been proposed as markers of neurodegeneration, a hallmark for AD and other neurodegenerative diseases [12]. Decreased levels of sTREM2, a marker of microglial activation and inflammatory changes, have been associated with increased rate of AD progression and is reported to be an endophenotype for AD [13–16]. Similarly, SNAP25, a member of the soluble N-ethylmaleimide-sensitive factor attachment protein receptors (SNARE) complex, has been proposed to be associated with decline in cognitive function resulting from synapse degradation, whereas variants and reduced levels of VILIP-1 have also been implicated in AD pathology [17–20]. Previous works have shown variable correlation values between measurements of these analytes assessed with high-throughput platforms such as SomaScan as compared with more classic immuno-assays [20, 21]. However,
neither SomaScan nor Olink includes all the current relevant AD biomarkers. High and mid-throughput technologies have been used to not only assess AD biomarkers, but also to perform discovery of new biomarkers and build predictive models [22-24].

Participant recruitment and retention along with sample acquisition, especially CSF via lumbar puncture, is one of the main challenges that observational studies face. Thus, maximizing use of the precious samples is key, especially in research settings. Volume is a minor limitation in a clinical setting. Input volume is highly dependent on the technology, but in general high and mid-throughput technologies require low volumes (55-100ul) to obtain thousands of measurements, which make them the most efficient for both volume and cost. There are several high throughput platforms available, including the relatively widely used Olink and SomaScan. Olink is an antibody-based technology that leverages proximity extension assays whereas SomaScan is aptamer based. However, in addition to the lack of robust measurements of established AD biomarkers, they present two additional disadvantages: (i) there is variable correlation between high throughput measures and the IP-MS or immunoassays values; [25] differences in correlation may be related to differences in the epitope measured, and differences in the binding efficiencies of the antibodies/aptamers; and (ii) most of the assays are not able to provide absolute quantifications values, which are typically used for medical interpretation. NUCleic acid Linked Immuno-Sandwich Assay (NULISA™) [26] is a mid-throughput platform that uses antibody-based measurements with a sequencing output. The main advantages are the low volume required (10 µL) combined with a high multiplexing capacity (more than 100 analytes) [26]. The central nervous system (CNS) panel includes traditional AD biomarkers, including Aβ40, Aβ42, p-tau181, and p-tau217 along with emerging and other AD biomarkers such as Aβ38, NFL, SNAP25, S100B, p-tau231, and p-tau217 among others.

Herein, we have benchmarked a new technology performing a head-to-head comparison of measures from the novel NULISA assay, IP-MS assays, and traditional immunoassays for Aβ40, Aβ42, Aβ42/40, total tau, p-tau181, and p-tau217 to assess utility in a research setting. For emerging biomarkers (NFL, NRGN, sTREM2, YKL40, SNAP25, GFAP, and VILIP1) we have compared NULISA to traditional immunoassays and high-throughput measurements. Finally, for other non-AD specific biomarkers we have compared NULISA to SomaScan and Olink.

**Material and Methods**

**Participants**

All research was performed in accordance with the protocols approved by the Institutional Review Board (IRB) at Washington University in Saint Louis.

We included 44 fasted CSF samples from 44 individuals and 44 unfasted plasma samples from 44 individuals from the Memory and Aging Project (MAP) at the Charles F. and Joanne Knight Alzheimer Disease Research Center (Knight-ADRC) at Washington University in Saint Louis with biomarker data measured available in CSF and/or plasma using reference and high-throughput assessments. Due to the benchmarking characteristics of this study, the inclusion criterion was data availability, no clinical criteria were used. Despite the presence of some individuals being present in both CSF and plasma samples (N=69), CSF and plasma samples/visits were not matched, and therefore represent different cohorts (Table 1) for the purpose of this study. As part of the regular study protocol, specimen collection, and imaging assessments are obtained at least every two years. CSF samples are collected close to the clinical visit at the Knight-ADRC MAP clinic after an overnight fast along with fasted and non-fasted plasma. CSF and plasma are immediately...
processed and stored at -80°C until used. Amyloid scans (obtained within six months before or after the biofluid collection) obtained and processed as previously described [27] were available for 43 of the 44 CSF samples and 39 out of 44 plasma samples.

For those participants with PET scans available (N=43 for CSF and N=39 for plasma samples), cortical Aβ depositions were quantified using 11C Pittsburgh Compound-B (PiB) PET and 18F-AV-1451 (florbetapir) PET. Individuals were classified as amyloid positive (N=14 for CSF samples, and N=20 for plasma samples) or negative (N=29 for CSF samples, and N=19 for plasma samples) based on imaging quantification. Given the use of two compounds, centiloids were standardized using Z-scores and positivity was assessed using a gaussian mixture model (GMM) data driven approach as previously described[21]. Additionally, all participants were stratified as AD cases (N=6 for CSF, and N=9 for plasma) or controls (N=35 for CSF and N=31 for plasma) based on clinical dementia rating (CDR®) and clinical diagnosis of AD. We were unable to classify three CSF samples and four plasma samples using this criterion.

**NUcleic acid Linked Immuno-Sandwich Assay (NULISA™) Measurements**

NULISA [26] is a novel technology based on a mechanism of dual capture and release. This approach minimizes assay background and increases sensitivity at the attomole level. Briefly, DNA conjugated antibodies form immunocomplexes with the target protein in solution, the immunocomplex is purified sequentially, first with an oligo-dT beads binding to the polyA sequence attached to the capture antibody, and then with streptavidin beads binding to the biotin on the detection antibody. Then, the DNA reporter is generated by proximity ligation and amplified by PCR to generate next-generation sequencing libraries. The DNA oligonucleotide contains the barcode to de-multiplex targeted protein and sample identification. This technology has the capability of measuring more than 200 proteins in the same 10µL of sample. To measure the samples included in this study we have utilized the NULISAseq CNS Disease Panel 120. Concentration is reported in NULISA Protein Quantification (NPQ) units. NPQ units are the result of normalizing the sequence quantification counts by intraplate and intensity variability and log2-transformed to approximate normality.

**Immunoassay and Mass Spectrometry Measurements**

All measurement were obtained from the Knight-ADRC Biomarker Core. Briefly, CSF Aβ40, Aβ42, total tau, and p-tau181 measurements were performed using a Lumipulse G1200 Chemiluminescent Enzyme Immunoassay Assay platform as previously described [3, 20, 21]. In contrast, plasma Aβ40 and Aβ42 levels were measured using mass spectrometry [5, 8] along with CSF p-tau217 [12, 28]. Only IP-MS concentration values have been retained for analyses. Other emerging AD biomarkers were quantified using a variety of immunoassays. CSF NiL was measured using commercially available immunoassays kits (Uman Diagnostics AB, Umea, Sweden) as previously described [29-31], whereas plasma levels were assessed using Quanterix[32]. CSF NRGN, SNAP25, and VILIP1 were performed using Single Molecule Counting (SMC) technology (EMD Millipore, Burlington, MA) as previously described [33-36]. YKL40 CSF concentration were measured using the MicroVue YKL40 EIA plate-based enzyme immunoassay (Quidel, San Diego, CA) for serum and plasma kit. We followed a previously described modified version of the protocol that allows quantification in CSF fluid. Briefly, samples were diluted threefold prior to assay, and washes included five cycles instead of four. Remaining steps were followed as indicated by the manufacturer [37]. Finally, sTREM2 was measured using
an in-house immunoassay [38-40] and plasma GFAP was measured using the ultra-sensitive single-molecule array (Simoa, Quanterix) [41].

High-Throughput Measurements

High-throughput aptamer-based protein measurements using SomaScan 7K were available for all plasma and CSF samples included in this comparison [42, 43]. In short, modified single-stranded DNA aptamers will specifically bind the targeted protein. After washing and processing, the aptamers are quantified by a DNA microarray. Quantification of protein concentrations is done using relative fluorescent units (RFU). Prior to analysis relative quantifications undergo stringent quality control, described elsewhere [23, 24]. Even though the full panel includes around seven thousand aptamers, we only included in here the ones that overlap with the NULISAseq CNS Panel and pass quality control in both platforms.

Similarly, 43 CSF samples were also measured using Olink Explore HT1. No Plasma samples were measured using Olink. This technology allows to measure thousands of proteins in a multiplex manner using Proximity Extension Assays (PEA) and Illumina Next Generation Sequencing using minimal sample volumes. Relative protein quantification is done via Normalized Protein eXpression (NPX) values that are already adjusted for intra and inter assay variability and log2 transformed to approximate normal distribution.

Statistical Analyses

All analyses were performed in R Project for Statistical Computing, including all the available data for each analyte, individual, and fluid. All the raw values corresponding to protein concentration and relative fluorescent units (RFU) from SomaScan were log10 transformed to approximate the normal distribution. NPQ values from NULISA are already log2 transformed and no transformation was used. Then, head-to-head comparisons were performed using Spearman correlations.

For differential abundance analyses we used case control classification along with quantitative imaging amyloid burden. The centiloid values were log10 transformed to approximate normality and used as phenotype in a regression. For both imaging and case/control status in plasma and CSF, the lm function from the stats R package was used to determine association. All analysis were adjusted by age and sex. P-values were FDR corrected to account for multiple testing.

Predictive performance comparison across measurement technologies were performed by using receiver operating characteristic (ROC) curves and calculating the area under the ROC curves (AUC). Then, we used multivariate logistic regression using normalized protein values to predict amyloid positivity. Additionally, we also included age and sex as covariates to account for baseline risk to evaluate if there was improvement in the prediction. Due to the limited sample size, no statistical test was used to formally compare performance.

Results

Established and emerging CSF AD Biomarkers

We compared the biomarker concentration of all the established AD biomarkers (Aβ40, Aβ42, Aβ42/40 ratio, total tau, and p-tau181) measured using NULISA (NPQ values) and Lumipulse (pg/mL; Figure 1A, Table S1). We observed that Aβ40 measured with Lumipulse and NULISA was highly correlated (ρ=0.73 [0.53-0.84], p=2.38×10^{-08}; Figure S1A) along with Aβ42 (ρ=0.84 [0.70-0.92], p=7.18×10^{-13}; Figure S1B). The Aβ42/40 ratio correlation was higher that the
Aβ species individually, with ρ=0.91 ([0.81-0.95], p=1.04×10⁻¹⁷; Figure S1C). We also tested the correlation between Aβ42/40 ratio and amyloid PET centiloid. The ratio obtained from Lumipulse showed a correlation of ρ=0.78 [0.59-0.89], (p=7.51×10⁻¹⁰, Figure S2A), similar to the one obtained from NULISA ρ=0.74 [0.53-0.85] (p=1.81×10⁻⁰⁸, Figure S2B). Total tau Lumipulse concentrations were correlated with NULISA (ρ=0.83 [0.63-0.95], p=2.96×10⁻¹²; Figure S1D). Similarly, p-tau181 Lumipulse values also correlated with NULISA (NPQ values ρ=0.98 [0.95-0.99], p=3.07×10⁻³³; Figure S1E). IP-MS was used to measure CSF p-tau217 concentration (pg/mL), that also showed high correlation with NULISA quantification (ρ=0.96 [0.92-0.98], p=1.60×10⁻²⁵; Figure S1F). Additionally, CSF p-tau217 concentration was correlated with centiloid values when measured with IP-MS (ρ=0.70 [0.49-0.83], p=2.00×10⁻⁰⁷; Figure S2C), and NULISA (ρ=0.75 [0.56-0.85], p=8.55×10⁻⁰⁹; Figure S2D).

Then, we analyzed markers of neurodegeneration, NFL and NRGN (Figure 1B, Table S1). NFL NULISA strongly correlated with immunoassay (ρ=0.96 [0.91-0.98]; p=3.45×10⁻²⁵; Figure S3A), SomaScan-based measures (ρ=0.93 [0.85-0.96]; p=6.50×10⁻²⁰; Figure S3), and Olink (ρ=0.93 [0.84-0.97]; p=1.71×10⁻¹⁹; Figure S3C). High correlations were also found for NULISA NRGN when comparing with SomaScan (ρ=0.94 [0.88-0.96]; p=1.26×10⁻²¹; Figure S3D), as well as SMC (ρ=0.91 [0.84-0.95]; p=2.42×10⁻¹⁷; Figure S3E), but lower, yet significant with Olink (ρ=0.64 [0.38-0.80]; p=3.03×10⁻⁰⁵; Figures S3F).

CSF sTREM2 levels, a marker of microglia activation and inflammation, were available in the four platforms (Figure 1B, Table S1), with an ρ=0.79 ([0.64-0.86], p=2.75×10⁻¹⁰) when comparing immunoassay and NULISA (Figure S4A) and ρ=0.87 ([0.76-0.92], p=2.31×10⁻¹⁴) when comparing SomaScan and NULISA quantifications (Figure S4B). Olink measurements were also significantly correlated with NULISA quantification (p=0.76 [0.60-0.85], p=2.06×10⁻⁰⁹; Figure S4C). YKL40, another biomarker of neuroinflammation, measurements using the immunoassay also showed significant correlation with NULISA values (ρ=0.78 [0.51-0.91]; p=1.38×10⁻⁰⁷; Figures 1B, 4D, Table S1), but not with those from SomaScan (ρ=0.33 [0.03-0.58], p=0.03; Figure 1S4E). However, Olink measurements were highly correlated with those from NULISA (ρ=0.79 [0.59-0.90], p=3.15×10⁻¹⁰; Figure S4F).

The biomarker for neuronal injury VILIP1 NULISA values showed high correlation across all three platforms (Figure 1B, Table S1). SMC quantification was highly correlated with NULISA relative quantification (ρ=0.90 [0.78-0.95], p=9.26×10⁻¹⁶; Figure S5A), with SomaScan (ρ=0.88 [0.78-0.93], p=2.04×10⁻¹⁵; Figure S5B), and Olink (ρ=0.79 [0.60-0.90], p=5.87×10⁻¹⁰; Figure S5C). SNAP25, a biomarker of synapse loss, also showed correlation between SMC and NULISA (ρ=0.81 [0.65-0.90], p=4.08×10⁻¹¹; Figures 1B, 5D, Table S1), Olink and NULISA (ρ=0.71 [0.47-0.85], p=2.73×10⁻⁰⁷; Figure S5E) but not between NULISA and SomaScan (ρ=-0.02 [-0.35-0.27], p=0.88; FigureS5F) as reported before [20]. CSF GFAP measurement, a marker of astroglia, was available in NULISA and SomaScan (Figure 1B, Table S1), with a moderate but significant correlation (ρ=0.64 [0.39-0.82], p=2.69×10⁻⁰⁶; Figure S5G), and Olink (ρ=0.92 [0.85-0.94], p=3.12×10⁻¹⁸; Figure S5H).

Plasma AD traditional and emerging biomarkers.
To assess the performance of NULISA assessments in plasma, we performed similar analyses as those described for CSF (Figure 1A, Table S1). IP-MS Aβ40 concentration was significantly correlated with NULISA relative quantification (\( \rho=0.74 \) [0.54-0.86], \( p=8.11 \times 10^{-69} \); Figure S6A), along with that of Aβ42 (\( \rho=0.72 \) [0.53-0.83], \( p=4.70 \times 10^{-68} \); Figure S6B). However, the Aβ42/40 ratio showed poor correlation between IP-MS and NPQ quantifications (\( \rho=0.25 \) [-0.07-0.51], \( p=0.10 \); Figure S6C). Along the same lines, the correlation of Aβ42/40 obtained from IP-MS measurements showed higher correlation (\( \rho=0.64 \) [0.40-0.82], \( p=1.20 \times 10^{-55} \); Figure S2F) than the one from NULISA assessments (\( \rho=0.37 \) [0.04-0.63], \( p=0.02 \); Figure S2G). Plasma p-tau181 and p-tau217 were not available for head-to-head comparison. We did however observe a highly significant correlation between p-tau217 and amyloid PET centroid (\( \rho=0.72 \) [0.55-0.85], \( p=1.87 \times 10^{-07} \); Figure S2H). NfL plasma measurements measured using Quanterix were highly correlated when compared to those assessed with NULISA (\( \rho=0.89 \) [0.76-0.94], \( p=5.74 \times 10^{-11} \); Figures 1C, S6D, Table S1) but not with SomaScan (\( \rho=0.29 \) [-0.01-0.55], \( p=0.06 \); Figure S6E). Plasma sTREM2 measurements with SomaScan were highly correlated with NULISA (\( \rho=0.94 \) [0.86-0.97], \( p=7.24 \times 10^{-20} \); Figures 1C, S6F, Table S1), whereas NRGN SomaScan values showed a moderate but significant correlation with NULISA (\( \rho=0.75 \) [0.55-0.88], \( p=7.76 \times 10^{-49} \); Figures 1C, S6G, Table S1). The YKL40 correlation between SomaScan plasma quantification and NULISA was significant (\( \rho=0.65 \) [0.44-0.79], \( p=2.15 \times 10^{-06} \); Figures 1C, S6H, Table S1). No correlation was observed between SomaScan and NULISA measurements for SNAP25 (\( \rho=0.05 \) [-0.24-0.34], \( p=0.74 \); Figure 1C, Table S1) or VILIP1 (\( \rho=0.14 \) [-0.20-0.45], \( p=0.41 \); Figure 1C, Table S1). Finally, plasma GFAP showed high correlation between Simoa and NULISA (\( \rho=0.83 \) [0.62-0.93], \( p=6.48 \times 10^{-49} \); Figures 1C, S6I, Table S1), but not between SomaScan and NULISA (\( \rho=0.19 \) [-0.08-0.45], \( p=0.22 \); Figure 1C, Table S1).

Comparison of NULISA with High Throughput measurements for other proteins

Similar to traditional biomarkers, we assessed if the measurements of the aptamer-multiplex platform SomaScan and the PEA-based platform Olink, were comparable to those quantified using the antibody-multiplexed platform quantifications NULISA (Table S2). We observed a wide range of correlation values for CSF (Figures S7A and S7B). SomaScan had 102 analytes in common with NULISA. The best correlations were observed for proteins such as BASP1 (Brain Abundant Membrane Attached Signal Protein 1, related to apoptotic mechanisms), NPTX2 (Neuronal Pentraxin II), or CXCL8 (Interleukin 8) with \( \rho \) values close to one, whereas others like IL10 (Interleukin 10), CD40L (CD40 ligand), or TNF (Tumor Necrosis Factor) showed the worst correlations (Table S2). There were 97 CSF proteins measured using Olink and NULISA. Overall, we observed a lower correlation with NULISA than SomaScan (Table S2). Top correlations between NULISA and Olink were observed for emerging biomarkers NFL, YKL40, and GFAP, with \( \rho > 0.90 \) and discussed in the previous section. None of the other quantifications show \( \rho > 0.90 \), MDH1 (Malate Dehydrogenase 1) showed the highest correlation (\( \rho=0.87 \) [0.73-0.93], \( p=2.62 \times 10^{-13} \)), followed by NPTX2 (\( \rho=0.86 \) [0.74-0.93], \( p=7.08 \times 10^{-14} \)). Non correlated measurements included BDNF (Brain-Derived Neurotrophic Factor), NPTXR (Neuronal Pentraxin Receptor) or CD40L among others. Differences in the targeted epitopes or in the sensitivity between assays might explain the wide range of correlation values observed, as previously described [26, 44].
In plasma, 95 protein measurements were available in NULISA and SomaScan. No measurement with Olink were available. Similar to CSF, the correlation ranges were very wide (Table S2, Figure S7C). Top correlated measurements include alpha synuclein, sTREM2, BDNF, and CRP (C-Reactive Protein), while the less correlated included established AD biomarkers such as p-tau217, p-tau231, and p-tau181.

**NULISA biomarker levels association against amyloid imaging and clinical status**

Next, we performed a differential abundance analysis using logistic regression to confirm the association with amyloid PET and clinical status for the most relevant proteins included in the NULISA CNS panel.

PET amyloid imaging was available for the 43 individuals with CSF measurements. We found the levels of ten CSF proteins significantly associated with the harmonized centiloid amyloid imaging measurements [21] (Figure S8A, Table S3). Of those, p-tau-217 (beta= 0.91 [0.68,1.14], p=2.45×10^{-09}), p-taur231 (beta= 0.73 [0.52,0.94], p=5.30×10^{-08}), p-tau-181 (beta=0.46 [0.28,0.64], p=1.44×10^{-05}), and total tau (Microtubule Associated Protein Tau (MAPT)-beta=0.35 [0.17,0.54], p=6.34×10^{-04}) showed the strongest association and remained significant after multiple test correction and adjusting by sex and age. Several other proteins were nominally associated with amyloid PET quantification (Aβ42, VILIP1, IL6R, YWHAG and TIMP3.

We also performed similar analyses in plasma (N=39) and found 18 proteins significantly associated with amyloid accumulation measured with PET (Figure S8B, Table S3). The levels of p-tau217 (beta=0.31 [0.23,0.39], p=6.22×10^{-09}) showed the strongest association with amyloid accumulation and remained significant after FDR correction followed by p-tau231 (beta=0.25 [0.10,0.40], p=2.33×10^{-03}), GFAP (beta=0.25 [0.09,0.41], p=0.004) and Aβ42 (beta=-0.17 [-0.31,-0.03], p=0.02). The remaining associations were related to the inflammatory pathway: CCL13; CCL26, CXCL8 IGFBP7, TIMP3, IL7, IL1B, SQSTM, BDNF, MDH1, FGF2, MME and HBA1.

When using clinical case-control classification for the differential abundance analyses, the results changed drastically, with a clear reduction in statistical power with six and nine proteins nominally significant in CSF and plasma respectively (Table S3). Of note, the comparison included six clinical AD cases for CSF, and nine in plasma. In CSF, IL1β was the protein with the strongest association (beta=-0.12 [-0.19,-0.05], p=1.76×10^{-03}). We were also able to detect Aβ42 (beta=-1.20 [-2.20,-0.20], p=0.02; Figure S8C) even with the unbalanced population studied. In plasma the strongest associations observed for NfL (beta=0.90 [0.45,1.35], p=3.99×10^{-04}) and p-tau217 (beta=0.49 [0.12,0.87], p=0.01; Figure S8D).

**Amyloid positivity predictive performance across platforms and tissues.**

We assessed the performance of each platform and protein to predict amyloid positivity assessed by PET for both CSF and plasma measurements as standalone measurement and adding age and sex at birth as covariates (Table S4). In CSF, the Aβ42/40 ratio showed an AUC of 0.91 [0.82-1.00] for immunoassay and 0.94 [0.88-1.00] for NULISA-based measurements when including sex and age (Figure 2A). Analogously, no performance differences were found for p-tau217 between IP-MS (AUC=0.92 [0.84-1.00]) and NULISA (AUC=0.93 [0.86-1.00]; Figure 2A). similarly, p-tau181 showed an AUC of 0.89 [0.79-0.99] and 0.90 [0.73-0.96] when measured using immunoassay techniques and NULISA respectively (Figure 2A). The AUC values for Total tau were very close when measured with NULISA (AUC=0.85 [0.73-0.96]) or with immunoassay (AUC=0.83 [0.71-0.95]; Figure 2A). NfL showed limited discriminatory performance regardless
of the technology (Table 3). Finally, NULISA p-tau231 showed similar classification performance (AUC=0.92 [0.84-1.00]; Figure 2A) than that previously reported in the literature [45].

In plasma, Aβ42/40 ratio show lower performance in NULISA (AUC=0.73 [0.57-0.89]; Figure 2B) compared to IP-MS (AUC=0.87 [0.76-0.99]; Figure 2B). Despite no comparison being available, p-tau217 measured by NULISA had similar predictive power (AUC=0.96 [0.90-1.00]; Figure 2B) than that reported previously [46]. The AUC for NfL was similar for both immunoassay and NULISA, regardless of the addition of covariates (AUC=0.60 [0.37-0.83] and AUC=0.60 [0.41-0.78] for immunoassay and NULISA respectively; Figure 2B), whereas for GFAP the predictive power of the immunoassay (AUC=0.83 [0.65-1.00]; Figure 2B) was higher than with NULISA (AUC=0.75 [0.59-0.91]; Figure 2B), in both cases when combined with age and sex.

Discussion

As the research community describes novel and better biomarkers, technology also advances to enable multiplex accurate measurements that use low sample volume. Using multiplex assays streamlines the integration of measurements from the same visit and reduce sample swap errors due to the processing of one sample per patient to obtain several quantifications, instead of processing several samples to obtain individual measurements than need to be matched. Low volume assays facilitate bridging and extends the use of each biological fluid. Overall, multiplex measurements are less prone to error, and more cost-efficient from the hands on-time and sample usage perspective. NULISA is a novel technology lacking a head-to-head comparison in the context of neurodegeneration-related biomarkers. Here, we have compared the NULISA quantifications (NPQ values) with those obtained using the gold-standard method (immunoassay or mass spectrometry-based) and high-throughput (SomaScan and Olink) quantifications to assess the accuracy and predictive power of this novel measurements in both CSF and plasma.

Overall, we have found that in CSF the NULISA values for the established biomarkers using well-validated gold standard assays (Lumipulse Aß42, Aß40, Aß42/40 ratio, total tau, p-tau181, and IP-MS p-tau217) are highly correlated and show similar performance in classifying amyloid PET status. In plasma, p-tau217 shows a similar AUC that the one reported in the literature [10, 46], as well as ptau231 [45], but the predictive power for Aß40/42 is significantly lower than established assays. These differences between CSF and plasma could be explained by the protocol optimization of IP-MS and immunoassay based on fluid. NULISA uses the same panel of antibodies and protocol regardless of tissue. Thus, it is plausible to think that the epitopes captured in CSF and plasma are optimized for immunoassay, but not for NULISA. Optimization of the NULISA CNS panel for plasma might improve the results reported in here.

When focusing on the emerging biomarkers, NULISA quantifications showed significant correlations for all the AD-relevant biomarkers such as sTREM2, GFAP, YKL40 when those were quantified using reference standard technologies. With high-throughput technologies the results are more variable, especially for those assays known to be non-optimal in the high-throughput platform such as YKL40 or SNAP25 in SomaScan. These findings suggest that NULISA can aid in the assessment of all AD-related biomarkers in multiplex manner for the first time.

Due to the inclusion of more than 100 analytes in the CNS panel, and the aim to assess the panel in a research setting, we performed differential abundance analysis to test if we were able to identify previously associated AD proteins, and potentially novel ones. The top protein in both CSF and plasma was p-tau217, already described to be an AD biomarker [28], closely followed by p-tau231. While in CSF we identified p-tau181, MAPT, Aß42 and NGRN, already associated
and known biomarkers for AD, only p-tau231 and GFAP were identified in plasma. All findings were in the correct direction of effect, supporting that correctly powered samples assessed using NULISA technology have the power to evaluate established AD biomarkers, measure AD-related processes such as neuronal injury or inflammation, and identify novel differentially accumulated proteins in a research setting.

This study has several limitations. The main one is the limited sample size used for the benchmarking, which could lead to wide confidence intervals for the correlations and the AUC calculations. For the same reason, we have not applied formal statistical comparison when calculating AUCs. However, the aim of this study was to assess the utility of the NULISA platform in research settings and assess the accuracy of its measures compared to gold standard and established assays measurements. Each assay performs within a range of concentrations, thus formal assay comparison in larger sample sizes with a wide range of concentrations could help assess if the lack of correlation, of special interest in plasma Aβ species quantification is due to lack of optimization or limit of detection. Overall, and despite the limitations, this work supports the use of NULISA quantification for research purposes, and evidence that the NPQ values have biological meaning and can be used to ask scientific questions. They provide accurate measurements and optimizes the use of precious biofluids. Similar analyses in larger samples sizes, multi-site studies, and diverse cohorts are needed to further validate its use in research or clinical settings.

In conclusion, this study reports for the first time the correlation values between established AD biomarker measurements assessed using the gold standard and a novel multiplexed assay. Given the high correlation in CSF and plasma, the minimal sample requirement, and the multiplexing capacity of NULISA, these assays have the potential of revolutionizing the scientific field by providing highly efficient measurements, saving precious sample and thus providing the means for more discoveries.

Acknowledgements

We thank all the participants and their families, as well as the many involved institutions and their staff.

This work was supported by grants from the National Institutes of Health (R01AG044546 (CC), P01AG003991(CC, JCM), RF1AG053303 (CC), RF1AG058501 (CC), U01AG058922 (CC), K99/R00-AG062723 (LI), RF1AG074007 (YJS), R01AG070941 (SES)), the Chan Zuckerberg Initiative (CZI), the Michael J. Fox Foundation (LI, CC), the Department of Defense (LI-W81XWH2010849), the Alzheimer’s Association Zenith Fellows Award (ZEN-22-848604, awarded to CC), the Alzheimer's Drug Discovery Foundation (LI), Bright Focus Foundation (LI), and an Anonymous foundation.

The recruitment and clinical characterization of research participants at Washington University were supported by NIH P30AG066444 (JCM), P01AG03991(JCM), and P01AG026276(JCM).

This work was supported by access to equipment made possible by the Hope Center for Neurological Disorders, the Neurogenomics and Informatics Center (NGI: https://neurogenomics.wustl.edu/) and the Departments of Neurology and Psychiatry at Washington University School of Medicine.

Conflict of interest
RB has received funding from NIH, Alzheimer’s Association, Biogen, AbbVie, Bristol Meyer Squibbs, Novartis, and Eisai. RB has equity and is on the scientific advisory board of C2N Diagnostics. SES has served on advisory boards and consulted on biomarker testing for Eisai, and she has received speaker fees from Eli Lilly. DMH has equity and is on the scientific advisory board of C2N Diagnostics. DMH is on the scientific advisory board of Denali, Genentech, and Cajal Neuroscience and consults for Asteroid. TLSB received research funding from Siemens and is an investigator on clinical trials and studies of AD with partial support from Eisai, Lilly, Roche, Janssen, and Biogen. TLSB is a consultant to Eisai. CC has received research support from: GSK and Eisai. CC is a member of the advisory board of Circular Genomics and owns stocks. CC is a member of the advisory board of ADmit.

The other authors report no conflict of interest.

The funders of the study had no role in the collection, analysis, or interpretation of data; in the writing of the report; or in the decision to submit the paper for publication.
**Figures and Tables**

**Table 1.** Demographic characteristics of the Knight-ADRC participants included in each of the two cohorts assessed in the present study.

<table>
<thead>
<tr>
<th></th>
<th>CSF Cohort</th>
<th>Plasma Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size (N)</td>
<td>44</td>
<td>44</td>
</tr>
<tr>
<td>Sex (N, % Females)</td>
<td>31 (70.45)</td>
<td>19 (43.18)</td>
</tr>
<tr>
<td>Age at Draw*</td>
<td>65 (50.91)</td>
<td>71 (50.91)</td>
</tr>
<tr>
<td>Non-Hispanic White (N, %)</td>
<td>42 (95.45)</td>
<td>42 (95.45)</td>
</tr>
<tr>
<td>APOe4 carriers*</td>
<td>16 (36.36)</td>
<td>24 (54.55)</td>
</tr>
<tr>
<td>PET Availability (N, %)</td>
<td>43 (97.73)</td>
<td>39 (88.64)</td>
</tr>
<tr>
<td>Positive Amyloidosis&amp;</td>
<td>14 (31.82)</td>
<td>20 (45.45)</td>
</tr>
<tr>
<td>Clinical AD Cases (N, %)</td>
<td>6 (13.64)</td>
<td>9 (20.45)</td>
</tr>
</tbody>
</table>

*mean (maximum, minimum); *Carriers of at least one APOe4 allele; *Positivity calculated using the normalized amyloid PET centroid and the GMM method; CSF=Cerebrospinal Fluid; PET=Positron emission tomography; AD=Alzheimer’s Disease.
Figure 1. Correlations between NULISA and other quantification methods for A. established AD biomarkers in CSF and plasma using gold-standard assessments, B. emerging biomarkers in CSF, and C. emerging biomarkers in Plasma. In green are indicated the comparison between NULISA and gold-standard measurements, in blue the comparisons to SomaScan, and in grey the comparisons to Olink. Circles indicate correlations in CSF whereas diamonds indicate correlations in plasma.

Figure 2. Performance of established and emerging biomarkers to classify brain amyloid positivity (based on Pittsburgh Compound-B (PiB) Positron emission tomography (PET) quantifications) expressed as Area Under the Receiver Operating Characteristic Curve (AUC) in A. CSF (circle), and B. Plasma (diamond). Red indicates AUC for NULISA, whereas green indicates the AUC for the gold standard method.
References


### Table A

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Assay</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aβ42/40 ratio</td>
<td>Lumipulse</td>
<td>43</td>
</tr>
<tr>
<td>Aβ42/40 ratio</td>
<td>NULISA</td>
<td>43</td>
</tr>
<tr>
<td>NFL</td>
<td>Uman ELISA</td>
<td>43</td>
</tr>
<tr>
<td>NFL</td>
<td>NULISA</td>
<td>43</td>
</tr>
<tr>
<td>p-τ181</td>
<td>Lumipulse</td>
<td>43</td>
</tr>
<tr>
<td>p-τ181</td>
<td>NULISA</td>
<td>43</td>
</tr>
<tr>
<td>p-τ217</td>
<td>IP-MS</td>
<td>43</td>
</tr>
<tr>
<td>p-τ217</td>
<td>NULISA</td>
<td>43</td>
</tr>
<tr>
<td>p-τ231</td>
<td>Literature</td>
<td>-</td>
</tr>
<tr>
<td>p-τ231</td>
<td>NULISA</td>
<td>43</td>
</tr>
<tr>
<td>Total Tau</td>
<td>Lumipulse</td>
<td>43</td>
</tr>
<tr>
<td>Total Tau</td>
<td>NULISA</td>
<td>43</td>
</tr>
</tbody>
</table>

### Table B

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Assay</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aβ42/40 ratio</td>
<td>IP-MS</td>
<td>39</td>
</tr>
<tr>
<td>Aβ42/40 ratio</td>
<td>NULISA</td>
<td>39</td>
</tr>
<tr>
<td>GFAP</td>
<td>Simoa</td>
<td>31</td>
</tr>
<tr>
<td>GFAP</td>
<td>NULISA</td>
<td>39</td>
</tr>
<tr>
<td>NFL</td>
<td>Quanterix</td>
<td>30</td>
</tr>
<tr>
<td>NFL</td>
<td>NULISA</td>
<td>39</td>
</tr>
<tr>
<td>p-τ217</td>
<td>Literature</td>
<td>-</td>
</tr>
<tr>
<td>p-τ217</td>
<td>NULISA</td>
<td>39</td>
</tr>
</tbody>
</table>

### Diagram

AUC

- **Assay**: NULISA, Gold Standard Tissue
- **Tissue**: CSF, Plasma

---

This analysis is licensed under CC-BY-ND 4.0 International.