

*Review*

# Proteomic Discovery and Validation of Novel Fluid Biomarkers for Improved Patient Selection and Prediction of Clinical Outcomes in Alzheimer's Disease Patient Cohorts

Shivangi Awasthi<sup>1</sup>, Daniel S. Spellman<sup>1</sup>, Nathan G. Hatcher<sup>1\*</sup><sup>1</sup> Merck & Co., Inc., Kenilworth, NJ, USA

\* Correspondence: nathan\_hatcher@merck.com; Tel.: +1 (215) 652 6017

**Abstract:** Alzheimer's disease (AD) is an irreversible neurodegenerative disease characterized by progressive cognitive decline. The two cardinal neuropathological hallmarks of AD include buildup of cerebral  $\beta$  amyloid ( $A\beta$ ) plaques and neurofibrillary tangles of hyperphosphorylated tau. The current disease-modifying treatments are still not effective enough to lower the rate of cognitive decline. The paucity of early detection and disease progression biomarkers also seems to present a major obstacle to AD drug development. The current established readouts based on expression levels of amyloid beta, tau and phospho tau have shown many discrepancies in patient samples when linked to disease progression. There is an urgent need to identify diagnostic and disease progression biomarkers from blood, CSF or other biofluids that can facilitate early detection of the disease and provide pharmacodynamic readouts for new drugs being tested in clinical trials. Advances in proteomic approaches using state-of-the-art mass spectrometry, are now being increasingly applied to study AD disease mechanisms, identify drug targets and novel disease biomarkers. In this report, we describe applications of the quantitative proteomic approaches for understanding AD pathophysiology, summarize the current knowledge gained from proteomic investigations of AD and discuss development and validation of new predictive and diagnostic disease biomarkers.

**Keywords:** Neuroproteomics; Alzheimer's disease biomarker; neurodegeneration

---

## 1. Introduction

The absence of therapeutics to halt or even slow the progression of neurodegenerative diseases such Alzheimer's disease (AD), Parkinson's disease (PD) and frontotemporal lobar degeneration (FTLD) is a vast unmet medical need that continues to worsen as the population ages. A major challenge facing drug development for neurodegenerative diseases lies in the early detection of disease to select patient groups whose pathology have not yet progressed beyond the potential for intervention and to enable early prediction of drug efficacy in clinical trials testing novel therapies [1, 2]. There has been significant investment in the development of neurodegenerative disease biomarkers, with innovations in AD research as an example, that include fluid biomarkers measured in cerebral spinal fluid (CSF) and discovery of PET ligands that enable determination of both amyloid beta and tau protein pathologies. While clinical biomarkers are yet to be discovered for other major neurodegenerative diseases such as PD and FTLN, even in the example of AD, where Tau and beta amyloid clinical biomarkers of disease state have been validated, utility in predicting clinical outcomes is yet to be demonstrated. Thus, there remains a need for additional biomarker discovery for early detection of neurodegenerative disease pathology, progression rate, prediction of conversion to late-stage pathology and assessment of clinical efficacy in drug trials. Changes in protein content of patient fluids such as CSF can provide insights into neurodegenerative disease state and offer a rich potential source for discovery and validation of putative biomarkers in a readily accessible fluid specimen.

Proteomic analyses of patient CSF can provide an unbiased discovery strategy to identify disease related changes in protein content. Modern discovery proteomics approaches follow a variety of analyte identification and quantification strategies to differentiate proteins that are statistically up or down regulated in patient fluid specimens. These analyses typically utilize extensive analyte prefractionation and separation using liquid chromatography coupled to high mass accuracy mass spectrometers, and these approaches follow a variety of analyte identification and quantification strategies from sample preparation to bioinformatic and statistical modeling of data to differentiate putative protein fluid biomarkers for follow-up validation studies. In this review, we discuss fluid biomarker development from discovery to validation with examples taken from the published studies within AD research. We review the current quantitation workflows and bioinformatic tools for differential biomarker discovery, discuss experimental design considerations for effective comparisons using these unbiased approaches and close with an overview of follow-on targeted quantification of biomarker candidates required for final validation using the Alzheimer's Disease Neuroimaging Initiative (ADNI) Longitudinal Proteomic Changes in CSF studies as an example.

## **2. Clinically validated fluid biomarkers for AD**

### **2.1. AD Epidemiology**

Alzheimer's disease (AD) is recognized as a global public health priority by the World Health Organization. Current estimates in the United States show that a total 6 million Americans are living with AD, with future projections of incidences expected to rise to nearly 13 million by 2050 as advanced age is the primary contributing factor for AD and other dementias, and occurrence is predicted to increase with increasing global life expectancies.

AD is the most common type of dementia, accounting for 50-70% of all cases. AD is the sixth leading cause of death in the country, and while deaths from heart diseases have decreased by 7.3% between 2000 and 2019, AD mortalities have increased by 145%. In 2021, healthcare estimates for AD and other dementias are \$355 billion and by 2050, these costs could rise as high as \$1.1 trillion [3]. Beyond costs, AD patients suffer from long-term, progressive losses in memory, cognitive function, speech, reasoning, personality, and motor functions that can result in significant ancillary burden for relatives and health care providers.

### **2.2. AD pathology and diagnosis**

A neuropathological feature shared by most common neurodegenerative diseases include aggregation or deposition of certain proteins in neurons or extracellular spaces which is often termed as "proteopathies." The primary hallmarks of AD pathology are amyloid beta A $\beta$ -containing plaques and tau-containing neurofibrillary tangles [4]. Although roles of A $\beta$  and tau pathology in neurodegeneration have not been elucidated, deposition of extracellular plaques containing A $\beta$  peptides have been linked to synaptic and neuronal defects, a phenomenon termed as "amyloid cascade hypothesis" [5]. The transmembrane amyloid precursor protein (APP), upon cleavage by  $\beta$ -secretase (BACE1) and  $\alpha$ -secretase releases insoluble A $\beta$ 42 peptides that form aggregates which accumulate into larger fibrillar structures forming the defining plaque tissue pathology within the brain [6]. Tau protein acts to stabilize microtubules important for intracellular transport in axonal and dendritic processes and maintains neuronal cytoskeleton integrity. Aberrant phosphorylation and truncations of tau proteins form the major components of neurofibrillary tangles [7, 8]. Hyperphosphorylated tau inhibits assembly and disrupts microtubules and impairs the axoplasmic flow and causes a loss of neuronal connection [9]. It has also been shown that hyperphosphorylated tau can spread in a prion-like fashion and be taken up by surrounding neurons via endocytosis [10].

The criteria for AD diagnosis were first published by the National Institute on Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) that helped the clinicians differentiate between symptoms as "unlikely", "probable", "possible" or "definite" AD [11]. These clinical exclusion criteria required a neuropathological examination upon death to ascertain a "definite AD" diagnosis as no imaging or fluid biomarkers were yet available. In 2007, the international working group (IWG) published the first research criteria for diagnosis of prodromal AD, which allowed for the use of biomarkers like CSF A $\beta$ 42, tau, volumetric MRI, and amyloid PET [12]. The National Institute on Aging (NIA) and the Alzheimer's Association (NIA-AA) published guidelines updating the 1984 clinical diagnostic criteria [11]. These guidelines elaborated the AD dementia beyond memory loss by classifying developmental disease stages beginning with an early and long-lasting asymptomatic preclinical stage and followed by mild cognitive impairment (MCI) before progressing to a final stage defined as AD dementia. Based on the amyloid cascade hypothesis, AD pathology can begin to set in several years before onset of any clinical symptoms. MCI was introduced as a term to denote the transitional zone between cognitive decline seen in normal aging and cognitive dysfunction seen in AD [13]. Updates have been published for MCI due to AD [14] as well as AD dementia [2, 15]. In recent years, there has been a focus on describing a phase that precedes MCI i.e., pre-MCI phase [16, 17], although, no standard definition is available, efforts have been directed to analyze the symptomatic debut of AD. AD patients with early onset (age < 65 years, EOAD) and late onset (age > 65 years, LOAD) present different cognitive and brain atrophy profiles. EOAD patients present multimodal cognitive dysfunction, performing poorly in visuo-spatial, speech, logical, and attention tasks while memory impairment is not dominant. LOAD patients, on the other hand, exhibit episodic memory impairment as a central symptom even in the early MCI phase [18]. Longitudinal studies have demonstrated a more rapid decline in EOAD than LOAD [19].

The 2011 NIA-AA updates were made followed by considerable advancements in knowledge regarding the disease pathology as well as maturation of the clinical, imaging and research techniques. These new guidelines added the use of CSF and imaging biomarkers as secondary diagnostic techniques although these were limited primarily to the supportive tools in clinical research. The IWG criteria was updated in 2014, where the CSF biomarkers, low A $\beta$ 42 combined with high total tau or phospho tau had more of a primary role while volumetric MRI and PET were to be used as secondary tools to monitor neurodegeneration [20]. In 2018, NIA-AA defined a new "research framework", A/T/N which identified imaging and CSF biomarkers as valid diagnostic tools [21, 22]. This revision of relying on biomarkers as indicators of disease progression reduces the dependence on patient biopsies or autopsies to obtain a definite diagnosis. Under this A/T/N framework, a subject is diagnosed with AD by presentation of amyloid pathology (A+, by amyloid PET or in CSF), by presence of absence of tau fibrillation (T, by PET or phospho-tau in CSF) and extent of neurodegeneration (N, by structural MRI or total tau in CSF). The A/T/N framework allows for addition of new biomarkers as and when they become available to better support the overall characterization of AD. This new protocol distinguished the three CSF core biomarkers of AD as per the pathological mechanisms to which they contribute.

### 2.3. Clinical utility of CSF A $\beta$ 42, total Tau and phosphor Tau

AD patients exhibit lower CSF A $\beta$ 42 concentration when compared to controls [23-25]. A significant increase in CSF total tau and phospho tau (p-threonine-181) in AD dementia patients has also been shown and replicated by many studies [23, 26]. Since plaques and tangles are at the core of the disease biology, A $\beta$ 42, total tau and phospho tau (p-tau) are referred to as the core biomarkers. Compared to healthy controls, AD patients exhibit low CSF levels of A $\beta$ 42 whereas total A $\beta$  is unchanged. These values are often expressed as an A $\beta$ 42/ A $\beta$ 40 ratio that accounts for intra-individual variations. Additionally, AD patients

exhibit high CSF levels of total tau and phospho tau. This distribution of A $\beta$  and tau in CSF is common among individuals with AD and is generally referred to as the “AD signature” [27].

CSF A $\beta$ 42/ A $\beta$ 40 as well as total and p-tau signatures have shown promise for early detection and prognosis for risk of progression. In a three-year longitudinal study following early MCI and healthy age-matched controls, Hansson and coworkers monitored CSF levels of tau and A $\beta$ 1–42 and were able to demonstrate diagnostic sensitivity of 95% and a specificity of 83% discriminating AD in patients with MCI as well as increased risk of progression to AD in MCI individuals exhibiting abnormal total tau and A $\beta$ 1–42 at baseline measures.[28]

Recent data suggests a critical role of hyperphosphorylation of tau occurring at different stages of disease progression which follows distinct trajectories over time. A comprehensive investigation that measured phosphorylation occupancy at multiple tau residues in CSF showed that tau phosphorylation is dynamic and begins once the aggregated A $\beta$  pathology is seen, changes over two decades tracking the clinical progression of the disease and goes down significantly in a site-dependent manner near the onset of cognitive decline [29]. Specifically, CSF p-threonine-217 levels are closely related to amyloidosis at asymptomatic and symptomatic stages and has been proposed as a biomarker to monitor AD pathology [30]. A follow up study found a positive  $\neg$  correlation between p-threonine-217 levels in plasma and CSF in clinical samples [31]. Another longitudinal analysis of plasma p-threonine-217 levels showed clinically relevant changes that correlated with preclinical disease stage, MCI as well as MCI to Alzheimer’s disease converters [32]. A recent study also highlighted the localization of p-threonine-217 in neurofibrillary tangles (NFTs), neuropil threads and multi-vesicular bodies in the postmortem AD brain tissue strengthening the potential of p-threonine-217 to be used as a biomarker [33].

#### 2.4. The unmet medical need

A true disease biomarker should highlight the activity of the disease process, predict outcome, and provide a pharmacological readout to therapeutic intervention. For a biomarker to be translated to a validated clinical test, it should be easy to use, cost-efficient, reproducible and should provide high sensitivity and specificity [34]. These points were emphasized further by both Food and Drug Administration (FDA) and European Medicines Agency’s (EMA) Committee for Medicinal Products for Human Use (CHMP) in context of AD biomarker discovery [35][36].

Currently, AD biomarkers can be divided into several categories, those that are derived from neurogenetics [37, 38], neuroimaging and neurophysiology [39, 40], and biochemical fluid biomarkers (CSF [41] and blood-based [42]). However, the success of such a multimodal approach is yet to be determined and needs additional development in terms of specificity and selectivity [43].

The accuracy of AD diagnosis in the clinic remains low and experienced neurologists and clinicians face challenges diagnosing AD in subjects with dementia. The rates of accuracy of AD diagnosis have been reported to be generally low; with sensitivities ranging between 71% to 88% and specificities from 44% to 71% [44]. It was also reported that the rate of misdiagnosis of AD can be as high as 20% [45]. This may be due to the presence of other comorbidities as the 2015 ADNI study showed that over 50% of the AD diagnosed patients show other pathologies as well [46].

The effects of AD span beyond the neurological system as it is known that the AD patients show physical decline as well. These systemic manifestations are in part driven by gradual and progressive cognitive and behavioral failures linked to neurodegeneration. The high degree of heterogeneity observed in the biological genotypes as well as behavioral phenotypes is evident by a great variety of pathological lesions, kinds of clinical symptoms,

onset age, types of behavioral manifestations. This well-documented information along with the presence of a plethora of molecular mechanisms involved in AD pathophysiology emphasize the multifactorial nature of the disorder. Although, the clinical accuracy of CSF A $\beta$ 42, tau and phospho tau CSF levels for diagnosis of the “AD signature” is robust, there are issues pertaining to sensitivity and specificity. There are examples of individuals who lack the AD CSF signature but have developed dementia. Additionally, it has been reported that tau CSF levels can increase in normal aging and can fail to discriminate AD from other forms of dementias like vascular dementia (VaD), frontotemporal dementia (FTD) and Creutzfeldt–Jakob disease (CJD) [47, 48]. Altered levels of A $\beta$ 42, tau and phosphorylated tau in CSF have been accepted to varying degrees as a diagnostic tool for AD in different countries. Even though IWG and NIA-AA have advocated for use of these fluid biomarkers to add diagnostic value, they do not endorse the use of these biomarkers for routine diagnosis as there is a need for further research, qualification, and standardization of such tests for use as diagnostics [12, 15]. CSF levels of A $\beta$  and tau are currently considered supportive indicators within a broader clinical phenotype [2].

Alterations in CSF A $\beta$  and tau CSF have utility in discriminating early MCI and offer potential in relating disease progression, but these changes occur over a relatively short increment of disease and thus have inherent limitations for early detection, prognosis as well as for use as surrogate efficacy biomarkers in drug trials. For example, CSF A $\beta$ 42, tau and phospho tau show minimal changes when patients progress from MCI to AD dementia [30, 31] or when patients are in the clinical phase of AD dementia [32]. This indicates that the alteration of normal to pathological levels of these CSF biomarkers precedes clinical symptoms and occurs during the preclinical asymptomatic phase of the disease. Since these biomarkers are stable over MCI and AD dementia, they don't always allow for AD prognosis and accurate disease staging. Moreover, clinical trials currently rely on relatively late-stage indicators of disease where levels of CSF A $\beta$  and tau may be maximally altered and no longer sensitive to therapeutic intervention.

Given these limitations, there is an urgent need to discover additional CSF biomarkers that cover a greater portion of AD disease progression. To support advances in AD intervention, prevention, and treatment, discovery of new diagnostic biomarkers is needed at the earliest stage in the disease continuum where the intervention is expected to provide the most long-term benefits. Additionally, new fluid biomarkers would benefit selection of prodromal patients with greater risk for developing AD and therefore could aid appropriate patient selection for clinical trials. Within clinical trials, there is a need to determine target engagement and potential disease modification of therapeutic candidates. Thus, likely multiple additional biomarkers are needed that can be used as surrogates of clinical efficacy across the specific timepoints or stages of disease within selected patient populations.

### 3. Mass spectrometry-based discovery proteomics

Recent advances in proteomic approaches using state-of-the-art mass spectrometry, are now being increasingly applied to elucidate disease mechanisms, identify drug targets and to identify novel disease biomarkers [49]. Neuroproteomics has emerged as a subcategory of proteomics and has been utilized extensively in the study of neurodegenerative diseases to identify potential biomarker candidates with diagnostic, prognostic and therapy predictive utilities as demonstrated by the growth rate of publications in this area [50]. The Human Proteomics Brain Proteome Project (HUPO BPP, <http://www.hbpp.org/>) an international initiative under the global Human Proteome Organization (HUPO) aims to further consolidate proteomic efforts to facilitate scientific discussions and collaborations and translate such efforts to clinical testing in various neurodegenerative disorders, including AD. The following sections will review proteomic applications using quantita-

tive mass spectrometry-based approaches and the current knowledge gained from proteomic investigations of AD with a focus on the discovery and validation of CSF biomarkers relating AD progression.

Proteins control the synthetic, catalytic, and regulatory biochemical processes and often assume a higher structured order referred to as a proteome [51, 52]. The large-scale systematic measurement of proteomes to generate biological insights about a system is referred to as proteomics. Mass spectrometry (MS) based approaches have emerged for reliable and exhaustive investigation into the composition and function of the proteome as an integrated system [53-56]. Liquid chromatography coupled to tandem MS (LC MS/MS) based proteomics analysis may comprise of simply identifying the proteins, looking into the nature and location of the post-translational modifications (phosphorylations, glycosylations, etc.) [57, 58], measuring dynamic quantitative changes between conditions or studying protein conformations or interactions in the context of biological signaling pathways [59, 60].

Mass spectrometry has been successfully applied to the proteomics analysis owing to its inherent specificity of identification and its exceptional sensitivity. The field of mass spectrometry has seen tremendous growth in the past four decades with regards to kinds and configurations of ionization sources, mass analyzers [61, 62] and ion detectors which has significantly impacted the proteome analysis [63]. The most significant of these developments include implementation of “soft” ionization techniques electrospray ionization (ESI) [64, 65] and matrix-assisted laser desorption/ionization (MALDI) [66], although developed in 1980s, remain prevalent for biomolecular analysis today. These “soft” ionization techniques are well suited for proteins and peptide analysis as these are polar, nonvolatile, and unstable analytes that need to be ionized without extensive degradation. MS-based proteomics has led to successful identification and accurate quantification of many proteins in complex biological matrices, thereby advancing the understanding of the cellular signaling pathways, characterizing the dynamics of protein-protein interactions in varied cellular states and locations, shedding light on the complex disease mechanisms, and providing unique biological insights [55, 63, 67].

### 3.1. Bottom-up shotgun proteomics (DDA vs DIA)

In a typical bottom-up proteomics experiment, sequence specific protease (trypsin is most common) cleaves proteins into peptides, which are separated through reverse-phase liquid chromatographic (RP LC) system (at nanoliter/minute flow rates [68]) and thereafter subjected to electrospray ionization (or nanospray (NSI) [69]) followed by a full MS1 scan [53, 70]. The topmost abundant ions are selected from the MS1 scan, fragmented by collision with gases like nitrogen or argon (most commonly by collision induced dissociation (CID) or higher energy collisional dissociation [71]), followed by acquisition of a MS2 scan of the resulting fragments (referred as data dependent acquisition (DDA) or “topN”). A single proteomics experiment will often have thousands of peptides being eluted off the LC column while the mass spectrometer generates the tandem spectra for the same. This tandem mass spectra provides the necessary information for peptide sequencing [72].

Standard instruments for DDA proteomics experiments include high resolution and high mass accuracy analyzers [73-75]. Mass analyzers such as the Orbitrap and the QTOF provide detailed structural information for accurate peptide and protein identification at acquisition speeds that can enable identification of hundreds of proteins in single LC run, but comprehensive coverage of the proteome is far from complete in these analyses. The dynamic range of protein expression in a cellular system can range from seven to eight orders of magnitude [76-78]. One study found that in a typical shotgun LC MS/MS run, approximately 100,000 “peptide-like” features elute but the current instrumentation is limited in its sequencing speed, sensitivity, and precursor ion isolation capabilities [79].

Various fractionation strategies at the protein or peptide level are often applied to overcome instrument limitations and delve deeper into the proteome (see section 4.4).

Whereas DDA selects the most abundant peaks in MS1 for fragmentation, Data Independent Acquisition (DIA) workflows aim to acquire complete MS/MS sequencing data systematically throughout the chromatographic run, independent of the detected analytes in the MS1 scan. The isolation windows for precursor ions in MS1 are set to cover the entire m/z range of the eluting peptides and can therefore generate extensive MS/MS maps for all the observable analytes in a sample. Opening of the precursor isolation windows leads to chimeric MS2 spectra originating from co-fragmentation of multiple peptides. As a large-scale discovery method, DIA can potentially overcome the under-sampling restrictions of DDA and limitations with the number of analytes that can be analyzed by selected/multiple reaction monitoring (SRM/MRM), without compromising analytical and quantitative performance [80]. However, proper interpretation of DIA data can be cumbersome because of the complex MS2 data that contain data from multiple peptides and can therefore be difficult to analyze.

### 3.2. Quantitative proteomics for differential feature identification

A major advantage of mass spectrometry based global proteomics is its capability to quantify proteins in a biological system. These quantitative measurements of protein abundances help build a functional network and provide a snapshot of protein changes associated with different states. These altered levels of proteins may provide clues for potential drug targets or disease biomarkers. MS-based quantitative proteomics can be broadly divided into two groups, 1) relative quantitative proteomics, where protein abundances are compared between two or more samples/conditions. The advancements of mass spectrometry instrumentation and bioinformatics today, allow for quantification of protein expression changes on a few to tens of thousands of proteins under normal or perturbed conditions. 2) absolute quantitative proteomics deals with measuring absolute abundances of proteins in samples. A variety of approaches that are utilized to carry out quantitative proteomic experiments are discussed in the following sections.

#### 3.2.1. Labeling strategies for quantitative proteomic comparisons

Most biological studies involve the measurement of protein quantities in a sample beyond generating long lists of protein identifications. DDA proteomics also provides the opportunity to extensively quantify the identified proteins relative to other samples. Though, it is important to note that bottom-up studies suffer from the technical limitation i.e., equal amounts of multiple peptides will not generate an equal MS signal, due to differences in ionization efficiencies that lead to ion suppression. To overcome this challenge, various quantitative proteomic strategies have been developed that are based on chemical isotopic labeling [81, 82], metabolic isotopic labeling, isobaric tagging [83]. The underlying quantification strategy takes advantage of the similar physical properties between isotopomers which result in minimal differences in ionization and near-identical ion suppression effects. Therefore, the labeling would provide good accuracy of quantification. The labelled spike-in proteins/peptides have an important application in measuring the absolute abundances of endogenous proteins in biological samples, assuming the labelled standard is present in the same amount in all the samples [84].

Stable isotope labeling by amino acids in cell culture (SILAC) is a popular MS1-based approach that is extensively applied to cell-culture based relative quantification experiments that provides highest accuracy and precision in a shotgun proteomics study. MS2- based relative quantitation approaches like isobaric tags have been commercialized as relative and absolute quantitation (iTRAQ) or tandem mass tag (TMT) reagents. These tags provide multiplexing capability and can analyze up to 16 samples at once. Although, the high throughput comes at the cost of ratio distortion which may lead to co-isolated precursor

contaminating the tandem spectra. Technological advancements as well as mature data analysis platforms help curb this problem [85].

### 3.2.2. Label-free feature extraction and quantitation

Despite label-based approaches being the gold-standard for MS-based quantitative proteomics methods [86], label-free quantitation (directly measuring the MS1 or MS2 responses) has gained momentum in the last decade. In the label free quantitation workflow, no label is introduced in the sample, all the samples are run in separate LC MS experiments, and independent peptide parameters are then used to carry out analysis of the relative protein abundances. Label free quantification can be carried out either by using algorithms based on extracted ion chromatograms (XIC) e.g., using peak areas computed from MS1 intensity measures or spectral counting algorithms based on MS2 spectra [87]. Label free methods are easy to implement, require no additional sample processing steps and are not limited by the number of samples. Although this approach does require complicated normalization and chromatographic alignment to account for retention time shifts and false matching of peptides across samples, it has become more popular and versatile owing to its flexibility, developments in high-resolution and fast scanning orbitraps [73] and advancements in software solutions for alignment of multiple MS runs [88].

### 3.3. Data processing and bioinformatics

As the mass spectrometric instruments have advanced over the years, the data generated from discovery proteomic experiments has grown increasingly intricate. The bioinformatics analysis of this raw MS data requires sophisticated algorithms that can confidently assign peptide identifications, through untargeted spectrum-centric search, to protein identifications [89, 90]. The bioinformatic analysis of DDA proteomic data involves transformation of the MS raw data into suitable formats to be used as an input for the search algorithms, the experimental spectra are then assigned identifications by matching to a protein sequence database or via a spectral library search or by de-novo sequencing. Different database search engines employ unique spectrum algorithms for peak picking, peptide sequencing and scoring MS2 spectra. The final step involves assigning peptide and protein identifications by applying statistical criteria. The threshold of protein detection is then determined by calculation of a false discovery rate (FDR). The most common method of calculating FDR includes construction of a decoy database which essentially consists of the same protein sequences as the regular database, but these entries have been reversed or scrambled. The FDR is usually fixed at 1% or 5% for typical DDA database searches. Confident protein identification depends on instrument mass accuracy, statistical methods applied for data normalization, missing value imputation and probabilistic scoring used for peptide to protein correlation [56, 59, 60]. This process of matching raw spectral data within a database or spectral library is defined as a “spectrum-centric search” methodology [80].

For DIA peak assignments, a targeted chromatogram extraction “peptide-centric” data query strategy [91] was implemented. This approach relies on prior acquisition of MS2 spectral libraries that are then used to uniquely isolate the target peptide by extraction of the product ion chromatograms and assessing the peptide coelution. Theoretically, any target peptide from the newly acquired data can be searched in this “peptide-centric” approach from the DIA data sets. This approach is intended to address the complicated, chimeric DIA MS2 spectra, as only the chromatographic characteristics like coelution, peak shape, intensity match with the assay library is being utilized to assign peptide identity. Therefore, issue of coelution and co-fragmentation of several peptides in MS1 becomes irrelevant [80].

### 3.4. Analytical instrument considerations

Reversed-phase LC is the most common mode of peptide separations employed in proteome studies. It is critical to carry out effective LC separation of peptides for a variety of reasons. First, the precursor intensity needs to be above the background noise and high enough for it to be selected for MS2 in DDA analyses. Second, effective resolution of peptides with similar  $m/z$  values is necessary to generate of unique and non-chimeric MS2 spectra for peptide identification and to limit potential opportunities for ion suppression. In this regard, one of the most significant enhancements in sensitivity was achieved by coupling nano electrospray [92] to reverse phase nanoflow LC (nLC). Lowering of the flow rate produces smaller ESI droplets, thereby leading to more effective desolvation, ionization and mass spectrometric sampling. The online coupling of nLC to nESI-MS may lead to a 100-1000-fold improvement in sensitivity because ESI is a concentration-dependent ionization technique. When sample amount is a limiting factor, lowering the LC column inner diameter and buffer flow to nanoliter per minute provides the largest gains in the global proteome coverage and quantitation by improved peak capacity and ionization efficiency [93].

Established chromatographic technological improvements such as ultrahigh pressure liquid chromatography (UPLC) [94], reduction in resin particle size [95], introduction of column frits [96] have also benefited the shotgun proteomic analysis. Monolithic nanoLC columns packed with continuous porous beds provide a rapid and efficient peptide separation without the back-pressure limitations associated with the smaller particle size. Silica-modified monolithic columns have demonstrated superlative performance by robust identification of over 6000 proteins from HeLa cell lysates [97]. Application of long-gradient (LG) separations coupled to MS has also been proved to be an effective strategy to diving deeper into the proteome. A LG-SRM study using a 150 cm long LC column and a gradient time of 300 min showed significant reduction of background interference and an 8-100-fold improvement in LOQ of target analytes in serum [98]. A systematic analysis of LG-SRM showed superior performance of the method by demonstrating lower ng/ml LOQs in abundant protein depleted plasma [99].

To improve the depth of proteome coverage, a multidimensional separation approach is often utilized to better resolve peptides from complex mixtures and reduce sample complexity [100]. The two-dimensional LC separation before MS (2D-LC MS) provides separation orthogonality, allowing resolution of components that would have otherwise co-eluted in a single dimension and thereby improving the overall peptide peak capacity [101]. The first application of 2D-LC MS utilized strong cation exchange (SCX) in the first dimension. SCX can be carried out using a salt gradient [102], where separation of the peptides based on their total charge or using a pH gradient [103] where peptides are resolved according to their isoelectric points (pI). SCX offers the ease of use because of its buffer compatibility with MS and can therefore be setup as an online 2D-LC platform [104]. RP-RP has also been used a 2D-LC approach were altering the mobile phase pH in the first dimension can improve orthogonality of separation [105]. Usually, the selectivity of separation is altered by using a high pH in the first dimension followed by a low pH in the second. Hydrophilic interaction chromatography (HILIC) has been used to separate polar and hydrophilic peptides and results in an elution profile opposite to that of RPLC [106]. Zwitterionic HILIC (ZIC-HILIC) was developed and evaluated as a multidimensional separation approach and was found to resemble the 2D SCX separation [107]. Electrostatic repulsion-hydrophilic interaction chromatography (ERLIC), a variation of HILIC has been shown to perform better than SCX as an orthogonal separation technique [108]. ERLIC has also been applied to explore the phospho- as well as glycoproteome from the rat kidney [109].

### 3.5. Emerging technologies

Alternate post-ionization separation techniques like ion mobility (IM) have recently started gaining prominence in the field of proteomics. Two variations of IM interfaces, namely high-field asymmetric waveform ion mobility spectrometry (FAIMS) and trapped ion mobility spectrometry (TIMS) have been adapted on orbitraps and Q-TOFs respectively that aim to provide an additional dimension of separation to facilitate a more comprehensive proteome coverage. FAIMS has been reported to overcome the challenges presented by co-fragmentation and co-isolation of ions in large-scale proteomics, thereby showing an improved sensitivity for low-abundant species [110]. TIMS coupled with Parallel Accumulation- SERIAL Fragmentation (PASEF) [111] is the mode of operation for the new timsTOF Pro mass spectrometer and has demonstrated a higher sequencing speed without the loss in sensitivity when applied to a single-shot proteomic experiment [112]. IM offers the ability to overcome co-eluting isobaric precursors as well as the challenges associated with co-isolated and fragmented “chimeric” spectra. This can in turn facilitate more accurate quantification for survey scan-based approaches by eliminating the presence of confounding ions.

Alternate data-acquisition techniques, like BoxCar, have also been recently introduced that showed significant improvement in the ion-beam sampling capacity at the MS1 level by distributing the ions over multiple narrow mass-to-charge segments in the C-trap to increase ion-injection times for lowly abundant species. Using this technique, the authors showed a 90% coverage of cancer cell proteome [113] without requiring extensive offline fractionation.

In the last decade, capillary electrophoresis coupled to mass spectrometry (CE-MS) [114, 115] has emerged as a complementary separation technology and has been successfully applied to proteomics for peptide measurements [116], intact protein analysis [117] and biomarker discovery [118]. CE offers many advantages over the traditional RP LC, namely, lower overall run time, lower sample consumption and better sensitivity. The lower sample consumption has enabled CE to be successfully applied for PTM quantification e.g., phosphorylation without enrichment [119] and two-plex stable isotope labelled quantification of N-glycans [120].

#### **4. Biofluid sample preparation**

##### **4.1. Considerations for sample integrity**

One of the reasons why the use of CSF biomarkers is challenging in the clinical practice is due to the high variability found in the concentrations calculated between different clinical centers. Such lack of reproducibility can be partly attributed to adoption of different pre-analytical procedures and sample handling steps. To eliminate the confounding effect of the sample processing steps, there is a need to establish standardized operating procedures.

Pre-analytical factors like specimen collection, hemolytic contamination of samples, sample handling, storage conditions and temperature, sample stability over freeze-thaw prior to processing and batch-to batch variability of the kits utilized are critical methodological issues that need consideration (see reviews [121-125]). The volume of CSF collected can influence the protein concentrations, although spinal cord gradient effect on A $\beta$ 42 has not been demonstrated, it is suggested to standardize the volume of CSF withdrawal to 12 ml [126].

Use of polypropylene plastic tubes for CSF storage has also been emphasized over polystyrene to avoid adsorption losses of “sticky” proteins like A $\beta$  [127]. As an example, lower values of all the core AD biomarkers were found when stored in polystyrene or glass tubes [128]. One study showed significant differences in levels of CSF A $\beta$ 42, tau and phospho tau (threonine-181) when polypropylene tubes from different vendors were employed. It was found that the polypropylene tubes introduced polyethylene contamination to the

clinical specimens. Moreover, CSF A $\beta$ 42 adsorptive loss in tubes differs significantly from phospho-tau owing to the varying degrees of protein hydrophobicity and hydrophilicity [129]. These data emphasize the need for standardization of the type of tube used for CSF biomarker storage and processing, as adsorptive losses can introduce sample handling artifacts into the clinical specimens. Use of 96-well polypropylene plates have shown similar losses of CSF biomarkers. The manufacturer recommendation of using a pre analysis plate before finally transferring all the samples to the final analysis plate hardly improves the protein losses. In fact, it was shown in two independent studies that using a pre-analysis plate significantly reduced the levels of A $\beta$ 42. These variations further contribute to lowering the interlaboratory reproducibility for the CSF diagnostic tests.

Blood contamination of the CSF is a significant issue and can impact the biomarker analysis because CSF protein amount is only 0.5% when compared to blood [130]. It has been previously demonstrated that blood contamination of CSF can lead to protein degradation [130] and lowering of A $\beta$ 42 levels due to binding of the protein to plasma proteins [131]. Several labs have also evaluated the effect of time delay between CSF collection and storage. Although, studies have reported stable CSF A $\beta$ 42, tau and phospho tau levels over 72 hours [132] to five [133] to seven days [134] after CSF collection, proteomics experiments have reported that the issue is more severe for serum/plasma proteins than CSF [135].

Multicenter studies have been conducted to analyze the impact of other CSF pre-analytical confounding factors in AD biomarker quantitation [122]. As an example, one study tested different spinning conditions, storage volume variations and effects of freeze-thaw cycles. It was found that spinning speeds generally do not influence protein levels except for samples where the total CSF protein content is relatively high. As for the CSF storage volume, reduction up to half of the tube capacity led to significant lowering of the A $\beta$ 42 concentration, while no effect was seen for tau or phospho tau. Also, it was concluded that tau and phospho tau can tolerate up to five freeze-thaws while levels of A $\beta$ 42 tend to decrease if freeze-thawed over three times [122].

#### 4.2. CSF versus blood

Biofluids do practically meet the expectations for an ideal biomarker, as in they are relatively easy to acquire and more economical to implement. CSF surrounds both brain and the spinal cord and continuously receives a stream of molecules secreted from neurons, glia, synapses, axons, in fact 20% of all CSF proteins are estimated to originate in the brain [136]. Given its close contact with the brain and constant interaction with the brain interstitial fluid, CSF, is the main fluid of choice as source of proteins/peptides biomarker candidates [137]. CSF is collected in clinical setting via lumbar puncture or spinal tap, and overall safety has been repeatedly demonstrated by exhaustive meta-analyses [138, 139]. Although, the occurrence of post procedure headaches and other side effects remain substantially low, the procedure suffers from a poor reputation and a general reservation from patients and clinicians. In fact, it was shown in prospective studies with patients exhibiting cognitive symptoms, the incidence of the well-publicized post-lumbar puncture headache is quite low (<2%) and comparable to headache risk associated with amyloid PET (1.8%) [140, 141].

CSF comprises of fluid formed by ultrafiltration of plasma via choroidal capillaries and partly by secretion of ependymal cells in the brain choroid plexuses. Although, CSF composition matches that of the blood plasma, the total protein content is 50-100-fold lower (CSF protein concentrations range from 0.2-0.7 mg/ml when compared to 50-70 mg/ml in plasma/serum [142]). Blood-based protein measures offer a less invasive and cost-effective strategy but present several challenges for proteomic analyses. From a technical perspective, biofluid proteome is highly heterogenous because of the presence of post-translational modifications, alternative splicing and conformational and structural proteoforms

[92]. Second, the dynamic range of protein expression in plasma/serum ranges from albumin at 50 mg/ml to signaling molecules in the range of lower pg/ml (10-12 orders of magnitude), which can further be aggravated by CNS specific protein abundances being washed out by proteins from peripheral sources [92, 93]. Most notably, over 80-90% of the CSF/plasma proteome [143] is dominated by high-abundance proteins (albumin, IgG, transthyretin, transferrin,  $\alpha$ 1-antitrypsin, apolipoprotein A, alpha-1-acid glycoprotein, haptoglobin,  $\alpha$ 2-macroglobulin, complement C3 [94]) which severely conceal the more informative low-abundance proteins [95]. CSF content is dominated by albumin which makes up to 60% of the total protein and is much higher than that in plasma. Moreover, CSF, contains other highly abundant proteins e.g., cystatin C and prostaglandin D2 synthase, that are being synthesized in CNS [144]. Therefore, because of the large differences in protein concentration and their well-known complexity, current proteomic techniques are limited to a small portion of biomarker appropriate biofluid proteome [96]. Global initiatives through Alzheimer's Association's International Society to Advance Alzheimer's Research and Treatment (ISTAART) have been undertaken to standardize both preanalytical as well as analytical guidelines and tools for AD blood-based biomarkers as part of the Alzheimer's Association Professional Interest Area (PIA) on Blood Based Biomarkers (BBB-PIA) [19, 103].

### 4.3. Immunodepletion

To improve the depth of the biofluid proteomic analysis, different strategies for the removal of the most abundant proteins are employed [145]. Affinity depletion of the most abundant proteins and immunoenrichment of the medium and lowly abundant proteins using combinatorial peptide ligand library are the two most applied approaches. The immunodepletion strategy involves removing the highly abundant proteins by loading the biofluid samples onto a depletion column, where antibodies specific to these proteins capture them, while the lowly abundant proteins are collected in the flow through. Immunodepletion kits have been available in the spin column format for decades now, e.g., Vivapure anti-HAS kit (Sartorius), and Qproteome (Qiagen). Among the most employed affinity depletion columns [146] are Multiple Affinity Removal Systems (MARS, Agilent) for depleting 2, 6, 7 or 14 most abundant proteins, Seppro-IgY-14 (Sigma-Aldrich) and ProteoPrep 20 (Sigma-Aldrich). It has been shown that the LC based techniques are far more superior to the spin columns in terms of replication, amount of carryover and non-specific binding [147]. The alternate strategy of relatively enriching the low and medium abundance proteins is based on concept of combinatorial peptide ligand library approach, available as ProteoMiner (Bio-Rad) [148], which is conceptually complementary to the depletion columns. In this approach, biofluid proteins will bind to immobilized hexapeptide ligands based on the bead capacity. Since the highly abundant proteins will quickly saturate their binding ligands, therefore limiting their binding capacity leading to the excess copies being washed off. In comparison, the lowly abundant analytes, do not saturate their ligands and will become relatively "concentrated" on the beads. This leads to reduced dynamic range of the biofluid proteome or "equalizing of the protein concentrations" [149]. Both approaches, affinity depletion and ligand library have been successfully applied to improve the CSF as well as plasma proteome coverage [145]. A study performed a side-by-side comparison of these two pre-analytical techniques using CSF and found that the MARS 14 column (773 identified CSF proteins) shows a superior qualitative performance over ProteoMiner (611 identified CSF proteins) [150].

## 5. Experimental design

### 5.1. Analytical validation, optimization, and quality control

Despite heavy investment of resources in the area, very few of these candidates pass clinical validation and make it into the clinical practice. Common to all 'omics studies, discovery proteomics generates several features (proteins) and are generally labor and time intensive which leads to analysis of a limited number of samples. These shortcomings, put together, present challenges both at the step of sample selection as well as statistical analysis of the resulting data. Therefore, to conduct robust biomarker discovery experiments, special emphasis is needed on defining the study rationale and statistical experimental design.

Much has been written on the importance of experimental design in the clinical biomarker discovery through 'omics technologies [104, 105]. Special attention needs to be paid on sample selection to avoid biases and to answer the defined clinical question. Choice of the proteomics method and establishing method performance are of critical importance. Taking appropriate steps to monitor inter-day instrument performance and demonstrating instrument reliability through proper quality control (QC) sample injections is key, as shifts in the liquid chromatographic or mass spectrometric performance can introduce severe interference and add to the experimental variability [106]. Several QC retention time calibrators are currently available to maintain LC MS performance which can be logged over time in automatic system suitability platforms like AutoQC in Panorama [107]. Given the depth of the analysis and experimental objective of the study, a few to several stages of method validation may be needed to determine the figures of merit (limit of detection and quantification, accuracy, precision).

To streamline the targeted proteomics measurements, a "fit-for-purpose" approach was used to define three tiers of targeted assays to define the analytical goals of the experiment and the empirical evidence needed to establish that the developed assays perform as intended [108]. The experimental design parameters and assay characteristics (e.g. inclusion of internal standards or reference standards) of each tier were described; with Tier 3 being "exploratory" in nature and low to moderate degree of analytical validation required, Tier 2 was intended for research grade protein/peptide quantification assays requiring moderate to high degree of validation and Tier 1 was described for clinical/diagnostic grade laboratory assays with requirements of high degree of validation as well as high performance metrics like precision, accuracy and repeatability. Based on these recommendations and to further ensure transparency and encourage reproducible research, the journal Molecular Cellular Proteomics introduced a new set of guidelines for manuscripts describing targeted proteomic measurements [109].

## 5.2. Biomarker validation

### 5.2.1. Validation of discovery proteomic data using targeted quantification

DDA based shotgun proteomics provides an exceptional speed and depth of protein profiling from complex mixtures. This unbiased approach may lead to identification of novel biomarkers, not based on existing hypotheses, and in turn may help generate new hypotheses on disease processes. Explorative proteomic studies using label free quantification has been used to identify several candidate AD markers. Although, the label free technology requires minimal sample preparation, it suffers from several disadvantages, e.g., this approach is not amenable to multiplexing therefore it results in long LC MS instrument times. This may in turn lead to lower analytical performance of the instrument over time (retention time shifts, ESI instability leading to inefficient ion transmission). This may be why CSF biomarker studies carried out via label free method have lower sample sizes (n=10-20) resulting in lower statistical confidence in the identified markers (A summary of a select discovery proteomic studies for AD biomarker discovery is presented in Table 1). Some of these limitations are overcome by isobaric labeling techniques like TMT, which can greatly eliminate run-to-run differences and where multiplexing is feasible thereby significantly reducing LC MS analysis time. Although highly reliable, isobaric labeling

techniques have some caveats as well, e.g., isotope-labeled reagents can add to the overall cost of the experiment in a large clinical study. Also, additional experimental steps like peptide pre-fractionation and verification of labeling completion, may be needed to avoid peptide co-isolation in MS/MS and systematic biases in quantitation.

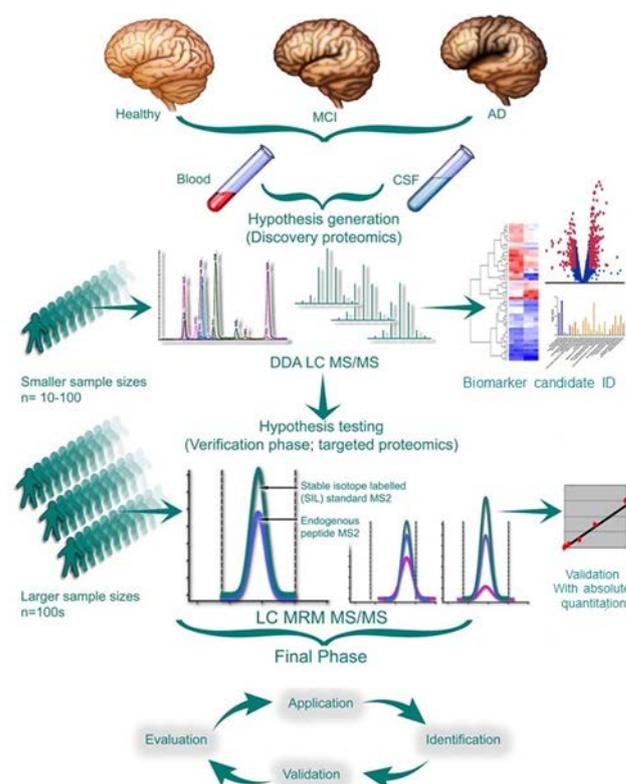
Because of the inherent limitations of DDA and its stochastic nature, only the highly abundant part of the proteome is reproducibly quantified, and the lower abundant species result in large number of missing data points. Moreover, the accuracy and consistency of quantification data further suffers from small sample sizes that do not possess the predictive power of these putative candidates for larger populations. Table 1 presents the summary of a few select proteomic studies that have claimed to identify potential biomarkers of AD in CSF. The relative quantitative data obtained from these discovery studies is often unreliable as follow-on validations are not carried out. A recent meta-analysis of 47 unbiased explorative proteomics studies of CSF biomarkers in AD unveiled a panel of 27 proteins and 21 peptides that were highly disrupted in AD in at least three publications [151]. To realize the full potential of these investigations, robust targeted quantitative methods are needed to validate these findings in multiple large independent populations across different stages of the disease.

### 5.2.2. Higher throughput quantitative assays for use in validation study cohorts

The validation phase of the biomarker discovery projects aims to overcome some of these limitations of the DDA proteomics by allowing for accurate and consistent quantification of the candidate molecules across large sample cohorts. Follow up studies are required to validate putative biomarker candidates identified in the discovery data sets. These measures typically use orthogonal targeted data acquisition schemes like SRM/MRM, parallel reaction monitoring (PRM) or data independent acquisition (DIA) methods to verify both analyte identification and changes in expression levels with improved quantification [76].

SRM-based targeted proteomics acquisition do not involve full MS<sup>2</sup> scans, but rather require selection of pre-determined precursor and product ion pairs, known as transitions. The predefined list of transitions for a targeted peptide/protein provides reproducible qualitative and quantitative data for hypothesis testing and biomarker validation studies [77]. Although SRM/MRM analyses provide a high degree of quantitation confidence, they do come at a higher cost of front-end assay development [78]. Standard instruments for SRM/MRM acquisitions include triple quadrupole instruments and quadrupole linear ion traps (QTRAP) operated as triple quadrupoles. PRM provides a compelling alternate to SRM/MRM that provides the same degree of quantitative specificity but a lower burden of method development [79, 80]. A PRM experiment involves a full MS<sup>2</sup> scan on an orbitrap instrument and lets users pick the optimal fragment ions post-acquisition.

The explorative CSF biomarker studies have benefited greatly from the development of robust and highly sensitive targeted MRM/PRM methods as verification assays (Figure 1, Table 1). These assays can be developed relatively quickly when compared to immunoassays and remain unbiased from matrix effects and antibody cross-specificity. MRM/PRM methods offer high analyte multiplexing capabilities, although can be time expensive when several stages of validation need to be undertaken.



**Figure 1: Discovery to validation to clinical implementation of biofluid protein biomarkers in AD:** proteomic biomarker discovery pipeline enables an unbiased and untargeted exploration of the biofluid proteome leading to identification of a list of potential biomarker candidates, that can be validated in larger sample cohorts. After undergoing several rounds of identification, verification and evaluation, candidate biomarkers can be applied to clinical practice.

**Table 1: Synopsis of published discovery and targeted proteomic studies for assay development and biomarker discovery in AD.** The sample size (n) for each condition is mentioned. The results from discovery studies mentioned were not validated further or were verified using other biochemical techniques like western blotting or ELISA

Discovery proteomics studies			
Sample	LC MS technique	Summary	Reference
Plasma AD (n=17) MCI (n=12) Control (n=11)	IP-MS coupled to MALDI-TOF	Immuno Affinity purification (IP) MS method developed to measure A $\beta$ s; A $\beta$ 1-40 and A $\beta$ 1-42) and A $\beta$ approximate peptides. APP/A $\beta$ (-3-40)/A $\beta$ 1-42 ratio was increased in amyloid PET-positive AD patients and was proposed as biomarker to surrogate cerebral amyloid deposition.	[152]
CSF AD (n=100) MCI (n=40) Control (n=80)	Label free LC MS	Anti-neurogranin antibodies were developed and used to show a marked increased level of neurogranin in AD dementia as well as MCI.	[153]
CSF AD patients (n=14) Control (n=14)	IP-PRM-MS	Significantly higher levels of CSF lysosomal protein LAMP2 was reported in AD patients when compared to controls	[154]
CSF Familial AD mutation carries (PSEN1 and APP, n=14) Non-carriers (n=5)	Label free LC MS	Comparative analysis identified 56 significantly differentially expressed proteins between groups. 14 of these aligned with the previous findings. Novel proteins reported include calyntenin-3, $\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor, CD99 antigen, di-N-acetyl-chitobiase, and secreted phosphoprotein-1. Protein expression changes in symptomatic and asymptomatic mutation carriers overlapped with those seen in late-onset AD.	[155]
CSF AD patients (n=8) Controls (n=8)	TMT labeling coupled to LC MS	The integrated proteomic and endopeptidomic approach simultaneously analyzed the abundances of 437 endogenous peptides and 374 proteins. The proteins that differed between groups include mesothelin, Ig alpha-1 chain C region, neu-rixin-1-beta, N-acetyllactosaminide beta-1,3-N-acetylglucosaminyltransferase, neurosecretory protein VGF, isoform 3 of neurotrimin, metalloproteinase inhibitor 2, and UPF0454 protein C12orf49.	[156]

CSF and cultured cells AD patients (n=3) Control (n=3)	CSF and cells combined in the same TMT multiplexed workflow	The optimized TMTcalibrator workflow allowed identification of lowly abundant peptides. Of the 77 proteins identified, 41 that are regulated in AD, hadn't been previously reported.	[157]
CSF AD patients (n=20) Control (n=20)	Endopeptidomic approach with stepwise protein and peptide precipitation followed by MALDI TOF MS and nLC MS	Peptides from VGF nerve growth factor-inducible precursor and $\alpha$ -2-HS-glycoprotein were downregulated in AD and a peptide from complement C4 factor and an O-glycosylated peptide from $\alpha$ -2-HS glycoprotein were found to be elevated	[158]
CSF Healthy volunteers (n=50)	Endopeptidomic approach with TMT labeling coupled to LC MS	Changes in CSF peptidome were measured longitudinally following administration of a $\gamma$ -secretase inhibitor. Many peptides showed dose-dependent changes in expression including one derived from APP and one from amyloid precursor-like protein-1, which are known $\gamma$ -secretase substrates.	[159]
CSF Pooled aliquots (n=14)	Label free LC MS	Quantitative Label free proteomic technique coupled to multi-affinity fractionation was used to assess technical variability as well as inter-subject variation. The technique was also evaluated for its ability to distinguish samples based on the dried biomarker criteria	[160]
CSF Dementia patients(n=159) Controls (n=17)	CE MS to identify differential peptide pattern for early differential diagnosis of various dementias	Using CSF measurements of A $\beta$ 42, tau and phospho tau, the AD pattern was diagnosed with a sensitivity of 87% and a specificity of 83%. Potential synaptic biomarkers identified: Apo-J, chromogranin A, phospholemmann, synaptic protein-like proSAAS and neuronal secretory protein VGF	[161]
CSF AD patients(n=4) Controls (n=22)	Label free LC MS	A $\beta$ 42 to A $\beta$ 40 ratio was estimated in PSEN1 mutant AD using surrogate amyloid precursor-like protein-1-derived A $\beta$ -like peptide (APL1 $\beta$ ), including APL1 $\beta$ 28. Relatively high ratio of CSF A $\beta$ 42 surrogate in PSEN1 mutant AD without an increase of A $\beta$ 42 secretion in the brain.	[162]
CSF (n=2)	Label free LC MS for extracellular vesicles (EV) characterization	Exosomal markers identified were alixand syntenin-1, heat shock proteins and tetraspanins and several brain -derived proteins. Known biomarkers of neurodegeneration were also identified in the EV fractions., e.g., amyloid precursor protein, the prion protein and DJ-1	[163]
CSF Postmortem CSF (n=4) Antemortem CSF (n=4)	TMT 6-plex coupled to LC MS	Discovery analyses found 78 identified proteins to be significantly upregulated in postmortem CSF samples when compared to antemortem. Previously identified brain damage biomarkers were identified like glial fibrillary acidic protein (GFAP), protein S100B, and protein DJ-1 (PARK7)	[164]
Plasma Non-demented controls (ND, n=36) Non demented subjects with AD family history (ND-FH, n=44) AD (n=40)	Label free LC MS	A $\beta$ -binding proteins circulating in the plasma were isolated and identified by LC MS. Many apolipoproteins were identified, i.e., apoA-I, apoB-100, apoC-III, and apoE. ApoA-I was reduced in AD and was proposed as an AD biomarker. ApoC-III was reduced in both ND-FH and AD and was proposed as a predictive marker for AD	[165]
Plasma Cohort 1 AD (n=24) MCI (n=261) Control (n=411) Cohort 2 MCI (n=180) Control (n=153)	iTRAQ coupled to LC MS	AD relevant biological pathways enriched in MCI included complement system, the coagulation cascade, lipid metabolism, and metal and vitamin D and E transport. Significant downregulation of potential markers fibronectin and C1 inhibitor was seen in the MCI cohorts.	[166]
Plasma AD (n=15) MCI (n=15) Control (n=15)	Isobaric labeling coupled to LC MS	Plasma levels of gelsolin was found to be decreased in AD subjects when compared to controls. This finding was validated via western blotting in the bigger validation cohort.	[167]

Validation cohort AD (n=60) Control (n=35)		However, additional validation from three different regions of the brain failed to replicate this finding.	
Plasma AD (n=15) Control (n=15)	iTRAQ coupled to LC MS	Differential expression of zinc-alpha-2-glycoprotein (AZGP1), fibulin-1 (FBLN1), platelet basic protein (PPBP), thrombospondin-1 (THBS1), S100 calcium-binding protein A8 (S100A8), and S100 calcium-binding protein A9 (S100A9) seen in the AD patients when compared to controls.	[168]
Plasma Stable MCI (n=58) Progressive MCI (n=34) Control (n=23) AD (n=31)	Label free LC MS	Both inflammation mediating proteins and pro-inflammatory IgG Fc glycoforms were significantly upregulated in AD subjects.	[169]
CSF Delirium (n=17) AD (n=17) Control (n=8)	iTRAQ coupled to LC MS	Discovery analyses of patients with delirium, a risk factor for development of dementia and patients with mild AD identified several interesting protein families, including apolipoproteins, secretogranins, chromogranins, clotting factors, serine protease inhibitors and acute-phase response elements.	[170]
<b>Targeted proteomic studies</b>			
CSF AD dementia (n=8), MCI (n=11), controls (n=19)	TMT coupled to IP-MS	Robust assay developed for parallel relative quantification of 27 A $\beta$ peptides in CSF. Although no statistically difference was seen between diseased and control groups.	[171]
CSF and blood AD patients (n=39) Control patients (n=38)	SRM: Absolute quant with heavy isotope standards	ApoE proteoforms quantified using stable isotope dilution. Total ApoE in CSF or blood doesn't distinguish AD from non-AD subjects. ApoE e4 carriers have lower blood ApoE irrespective of clinical diagnosis.	[172]
Plasma Case-control (n=669)	SRM-MS	Total ApoE and ApoE e4 proteoform quantified. ApoE e4 specific peptide contained a single methionine, which was chemically oxidized after tryptic digestion, completeness of oxidation was thoroughly evaluated. Chemical oxidation allowed unbiased monitoring of ApoE e4 unique proteotypic peptide. Neither total ApoE and ApoE e4 levels nor ApoE/APOE e4 ratio consistent with AD diagnosis	[173]
Serum DLB patients (n=47) AD patients (n=97)	SpotLight Melon Gel kit enriches polyclonal IgGs.	De novo sequencing identifies peptides from variable regions of IgGs and uncovers "hidden proteome". SpotLight peptide quantification generated a predictive model with 95% accuracy to distinguish AD and dementia with Lewy bodies	[174]
CSF (Sample size not listed)	IP-MS with heavy isotope internal standards coupled to MALDI-TOF. Confirmation carried out on LIT-FT ICR	Affinity purification MS method optimized for A $\beta$ using A $\beta$ specific crosslinked antibodies. Two novel Ab peptides identified: A $\beta$ 2-17 and A $\beta$ 3-17 (probable cleavage products of neprilysin and ECE) The developed assay facilitated target engagement clinical studies [175, 176]	[177]
CSF (three separate cohorts) Cohort 1: AD subjects (n=9), prodromal AD (n=7), non-demented controls (n=9) Cohort 2: AD (n=10), non-demented controls (n=6) Cohort 3: AD (n=17), non-demented controls (n=17) Brain tissue Autopsy confirmed AD patients (n=15) Age-matched controls (n=15)	IP-SRM-MS with heavy isotope internal standards	Affinity purification MS method developed to measure levels of the presynaptic protein synaptosomal-associated protein 25 (SNAP-25) in CSF. SNAP-25 levels were significantly higher in prodromal AD and AD when compared to controls. CSF SNAP-25 differentiated AD from controls and was proposed as novel biomarker for synapse degeneration.	[178]
CSF (2 cohorts) Cohort 1: CSF AD dementia (n=15), MCI (n=5), controls (n=17) Cohort 2: CSF AD (n=24), MCI (n=18), controls (n=36)	IP-PRM-MS with heavy isotope internal standards	Affinity purification MS method developed to measure levels of the presynaptic vesicle protein synaptotagmin-1 in CSF. Synaptotagmin-1 levels were significantly higher in MCI AD and AD dementia when compared to controls. CSF synaptotagmin-1 was proposed as a biomarker of synaptic dysfunction and degeneration in AD	[179]
CSF and plasma AD patients (n=43) Control (n=43)	SRM MS with heavy isotope internal standards	Previously developed ApoE quantification assay was used to measure ApoE proteoforms ApoE2, ApoE3 and ApoE4. No distinction was found between AD patients and controls.	[180]

CSF AD patients (n=37) Control (n=22) Validation cohort AD patients (n=24) Control (n=16)	SRM with heavy isotope internal stand- ards	Significantly higher concentration of soluble triggering receptor expressed on myeloid cells 2 (sTREM2) was found in AD patients when compared to controls. This finding was replicated in the validation sample set. sTREM2 was found to correlate with markers of neurodegeneration and glial activation.	[181]
--	--	---	-------

### 5.3. Multisite variability assessment: quantitative proteomic data reporting, sharing and the need for standardization

In the last few years, we have seen a concerted effort towards establishing accepted protocols, guidelines, and criteria for testing, reporting and qualification of mass spectrometry based proteomic biomarkers. Multi-site replication studies are needed for the assessment of intra-lab and inter-lab variability utilizing the same or alternative proteomic technique and data processing platforms. Research transparency is encouraged through data sharing to scientists, regulators, clinicians through online repositories, such as ProteomeXchange (Pride [115, 116]) and Panorama Public [117]. LC MS AD biomarker discovery research can borrow conventional wisdom from the more mature oncology initiatives like Clinical Proteomic Tumor Analysis Consortium (CPTAC). CPTAC is a web-based assay portal that aims at advancing the development and dissemination of thoroughly validated mass spectrometry based targeted proteomic assays. CPTAC also actively supports proteogenomic translational research to integrate, visualize and analyze cancer biology across multiple 'omics data dimensions like genomics, transcriptomics and proteomics [119]. CPTAC data portal provides a central database for sharing and re-use of data across the research community to advance cancer biomarker discovery and clinical translation [118]. Moreover, CPTAC has provided a validation guidance document to help researchers carry out MRM assay characterization thus promoting high standards for assay inclusion. The validation document outlines five experiments to evaluate assay parameters like linearity, intra-assay precision, upper and lower limit of quantification, repeatability, selectivity, and internal standard stability [182].

### 6. Case study: Longitudinal Proteomic Changes in CSF from ADNI: Towards Better Defining the Trajectory of Early Alzheimer's Disease

Alzheimer's Disease Neuroimaging Initiative (ADNI), setup in 2004, is an ongoing multi-center project that has come a long way to identify changes in brain structure and function and cognition related changes [183]. One of the major objectives of ADNI has been to validate MRI, PET and CSF/blood biomarkers as outcome predictors to be used in AD clinical trials. ADNI has made a profound impact in the field by development and implementation of standardized protocols across multiple sites and improved clinical trial efficiency by identification of sensitive outcome measures for patient stratification [184, 185]. The overarching goal of ADNI is to detect AD at the earliest possible stage and implement ways to stage AD progression using robust biomarkers [186]. Another defining characteristic of ADNI is its innovative data-access policy and the commitment to allow embargo free data sharing to stimulate collaborations and further scientific investigations to answer unresolved questions about competing hypotheses regarding the disease pathophysiology [187]. Laboratory of Neuroimaging (LONI) at the University of Southern California hosts all the archives of the data generated via ADNI studies. Owing to the open data sharing approach, presently, over 1000 publications have utilized ADNI data in AD research as well as in fields outside of AD [186]. ADNI data has also been used by big data projects and a number of consortia like Enhancing Neuro Imaging Genetics through Meta-Analysis (ENIGMA) consortium [188] as well as Dialogue on Reverse Engineering Assessment and Methods (DREAM) Alzheimer's disease Big Data Challenge #1[189].

ADNI samples have been utilized in one of its kind, multi-phased effort, using targeted multiplexed mass spectrometry-based assays to identify diagnostic and predictive CSF based biomarkers in AD. This project presented the analysis of many candidate peptide markers from a well characterized AD patient sample cohort. Phase 1 of this study was conducted as a feasibility analysis to assess MRM assay characteristics e.g., reproducibility of sample processing, analytic variability, and the ability to detect a variety of analytes of interest. The CSF multiplex MRM panel was built with proteins and peptides curated from a variety of published and unpublished AD biomarker studies in CSF, brain and cell lines and previous results from the Rules Based Medicine (RBM) multiplex ADNI CSF immunoassays. The phase 1 of the study was completed using 25 human CSF samples that included 5 technical replicates, that underwent abundant protein depletion followed by tryptic digestion and were detected by LC MRM MS. Of the 510 peptides/267 proteins queried, 198 peptides/121 proteins were detectable in the CSF, measuring 2 transitions per peptide. After the successful completion of the pilot study, phase 2 was initiated. The final MRM panel was built by supplementing the detectable peptides from the pilot study with many other peptides particularly from proteins of interest either those that are inflammatory markers or those of interest in the RBM assay. The final MRM assay consisted of 567 peptides representing 222 proteins. System suitability test of the LC MRM MS system was carried out using synthetic standard peptide solutions to assess instrument reproducibility as well as sensitivity. Several filtering criteria was employed to refine the final MRM transitions using parameters like co-elution with the standard, retention time stability over runs, reproducibility of the intensity ratios, setting a detection threshold. The refined MRM assays were applied to 306 ADNI-1 CSF samples (85 healthy controls, 66 AD patients and 134 MCI patients) that included 16 blinded technical replicates. A variety of statistical approaches like univariate association/prediction analyses as well as multivariate exploratory/supervised analyses were employed to assess whether the analytes were associated with clinical pathology like MCI or AD vs. healthy controls or associated with progression vs. non-progression from MCI to AD. The results included several peptides with potential “diagnostic” utility mainly from Hemoglobin A (HBA), Hemoglobin B (HBB) and superoxide Dismutase (SODE) as well as peptides with “predictive” utility from neuronal pentraxin-2 (NPTX2), neurosecretory protein VGF (VGF), and secretogranin-2 (SCG2) [190]. This study utilized extensive quality control and data normalization measures and applied the level of rigor that has not been previously demonstrated in the proteomic biomarker research.

The next phase of this effort focuses on within-subject alterations over time in patients with MCI, AD, and healthy controls. This longitudinal sample set has at least three CSF samples from each patient drawn over a period of three years or more. The aim of this project was to evaluate the ability of a panel of peptides to distinguish the disease states and show changes in a longitudinal manner. Proteins and peptides were selected based on their previous detection in CSF, AD relevance and previous results from RBM as well as earlier MRM assays. A total of 5 primary (Fatty Acid Binding Protein-3 (FABP3), SCG2, VGF, NPXT2 and Chromogranin A (CHGA)) targets were chosen for absolute quantification and 121 targets were chosen for relative quantification. The final MRM assay consisted of 278 peptides representing 126 proteins. The refined MRM assays were applied to 750 unique CSF samples that comprises of 730 longitudinal samples from ADNI-1, ADNI-2 and ADNI-GO and 20 blinded replicates. The samples were processed in a randomized and blinded manner. Due to the presence of endogenous proteins in CSF, the standard curve for absolute quantitation was built using recombinant proteins (FABPH, SCG2, VGF, NPXT2 and CHGA) in BSA buffer. Quality control samples were prepared using a CSF pool of 300 individual CSF samples supplied by ADNI. The blinded phase of the project has been complete (unpublished data) and further analysis using linear mixed effects modeling is underway as ADNI subject IDs are un-blinded. The primary objective of this study is to retrospectively investigate analytes that exhibit intra-individual trajectories

corresponding to various stages of disease including cognitively normal, MCI and AD dementia. The data indicated that neuronal pentraxin 2 (NPTX2) could be a potential disease progression marker in AD. NPTX2 exhibited robust association with baseline clinical diagnosis of MCI for the intra-subject evaluations. Moreover, degradation of NPTX2 showed positive correlation with cognitive decline. These findings suggested that NPTX2 could be a potential prognostic biomarker of accelerated cognitive decline in a subset of AD patients [191].

These ADNI studies, demonstrate the potential for the discovery of novel biomarkers that can be used as endpoints in clinical trials of early AD, can help monitor treatment effects, enable more efficient trial design and further help understand the therapeutic mechanism of action.

## 7. The Promise of Fluid Biomarkers of CNS-Related Diseases

The paucity of early detection and disease progression biomarkers seems to present a major obstacle to AD drug development. The current established readouts based on expression levels of amyloid beta, tau and phospho tau have shown many discrepancies in patient samples when linked to disease progression. Current diagnostic measures based on neuroimaging-based techniques like structural and functional Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) are expensive, require advanced on-site technologies and expertise. There is an urgent need to identify new markers from biofluids like CSF and blood that can facilitate early detection of the disease and provide pharmacodynamic readouts for new drugs being tested in clinical trials. These biofluid biomarkers can be attained through less invasive means and measures are amenable to routine clinical laboratory workflows. Traditional biochemical techniques generally focus on only a few key genes/proteins to investigate molecular processes, the resulting data often provides an insufficient and somewhat incomplete understanding of complex intricacies of AD pathology. Mass spectrometry based multiplexed targeted proteomic assays can be used as an efficient strategy to analyze many biofluid candidate markers with diagnostic and prognostic value for AD drug development. As demonstrated by the large-scale ADNI studies, the promise of novel CSF AD biomarkers can be realized by carrying out large-scale, longitudinal, well-characterized, robust studies with good statistical power.

**Author Contributions:** Conceptualization, N.G.H. and D.S.S. ; Manuscript Preparation, S.A and N.G.H.

All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Mattsson, N., et al., Revolutionizing Alzheimer's disease and clinical trials through biomarkers. *Alzheimer's Dement (Amst)*, 2015. 1(4): p. 412-9.
2. Lewczuk, P., et al., Cerebrospinal fluid and blood biomarkers for neurodegenerative dementias: An update of the Consensus of the Task Force on Biological Markers in Psychiatry of the World Federation of Societies of Biological Psychiatry. *World J Biol Psychiatry*, 2018. 19(4): p. 244-328.
3. <https://www.alz.org/alzheimers-dementia/facts-figures>
4. Glenner, G.G., et al., The amyloid deposits in Alzheimer's disease: their nature and pathogenesis. *Appl Pathol*, 1984. 2(6): p. 357-69.
5. Hardy, J.A. and G.A. Higgins, Alzheimer's disease: the amyloid cascade hypothesis. *Science*, 1992. 256(5054): p. 184-5.
6. Masters, C.L. and D.J. Selkoe, Biochemistry of amyloid beta-protein and amyloid deposits in Alzheimer disease. *Cold Spring Harb Perspect Med*, 2012. 2(6): p. a006262.
7. Kosik, K.S., C.L. Joachim, and D.J. Selkoe, Microtubule-associated protein tau (tau) is a major antigenic component of paired helical filaments in Alzheimer disease. *Proc Natl Acad Sci U S A*, 1986. 83(11): p. 4044-8.
8. Grundke-Iqbal, I., et al., Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. *Proc Natl Acad Sci U S A*, 1986. 83(13): p. 4913-7.
9. Iqbal, K., et al., Mechanism of neurofibrillary degeneration and pharmacologic therapeutic approach. *J Neural Transm Suppl*, 2000. 59: p. 213-22.
10. Clavaguera, F., et al., Transmission and spreading of tauopathy in transgenic mouse brain. *Nat Cell Biol*, 2009. 11(7): p. 909-13.
11. Reitz, C., C. Brayne, and R. Mayeux, Epidemiology of Alzheimer disease. *Nat Rev Neurol*, 2011. 7(3): p. 137-52.
12. Dubois, B., et al., Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol*, 2007. 6(8): p. 734-46.
13. Petersen, R.C., Mild cognitive impairment as a diagnostic entity. *J Intern Med*, 2004. 256(3): p. 183-94.
14. Albert, M.S., et al., The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*, 2011. 7(3): p. 270-9.
15. McKhann, G.M., et al., The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*, 2011. 7(3): p. 263-9.
16. Storandt, M., et al., Longitudinal course and neuropathologic outcomes in original vs revised MCI and in pre-MCI. *Neurology*, 2006. 67(3): p. 467-73.
17. Duara, R., et al., Pre-MCI and MCI: neuropsychological, clinical, and imaging features and progression rates. *Am J Geriatr Psychiatry*, 2011. 19(11): p. 951-60.
18. Migliaccio, R., et al., Mapping the Progression of Atrophy in Early- and Late-Onset Alzheimer's Disease. *J Alzheimers Dis*, 2015. 46(2): p. 351-64.
19. Koss, E., et al., Clinical and neuropsychological differences between patients with earlier and later onset of Alzheimer's disease: A CERAD analysis, Part XII. *Neurology*, 1996. 46(1): p. 136-41.
20. Dubois, B., et al., Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *Lancet Neurol*, 2014. 13(6): p. 614-29.
21. Jack, C.R., Jr., et al., NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement*, 2018. 14(4): p. 535-562.
22. Jack, C.R., Jr., et al., A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers. *Neurology*, 2016. 87(5): p. 539-47.
23. Olsson, B., et al., CSF and blood biomarkers for the diagnosis of Alzheimer's disease: a systematic review and meta-analysis. *Lancet Neurol*, 2016. 15(7): p. 673-684.
24. Motter, R., et al., Reduction of beta-amyloid peptide<sub>42</sub> in the cerebrospinal fluid of patients with Alzheimer's disease. *Ann Neurol*, 1995. 38(4): p. 643-8.
25. Strozzyk, D., et al., CSF Aβ<sub>42</sub> levels correlate with amyloid-neuropathology in a population-based autopsy study. *Neurology*, 2003. 60(4): p. 652-6.

26. Blennow, K., et al., Tau protein in cerebrospinal fluid: a biochemical marker for axonal degeneration in Alzheimer disease? *Mol Chem Neuropathol*, 1995. 26(3): p. 231-45.
27. Shaw, L.M., et al., Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. *Ann Neurol*, 2009. 65(4): p. 403-13.
28. Hansson, O., et al., Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impairment: a follow-up study. *Lancet Neurol*, 2006. 5(3): p. 228-34.
29. Barthélemy, N.R., et al., A soluble phosphorylated tau signature links tau, amyloid and the evolution of stages of dominantly inherited Alzheimer's disease. *Nat Med*, 2020. 26(3): p. 398-407.
30. Barthélemy, N.R., et al., P1-023: MASS SPECTROMETRY-BASED MEASUREMENT OF LONGITUDINAL CSF TAU IDENTIFIES DIFFERENT PHOSPHORYLATED SITES THAT TRACK DISTINCT STAGES OF PRESYMPTOMATIC DOMINANTLY INHERITED AD. *Alzheimer's & Dementia*, 2018. 14(7S\_Part\_4): p. P273-P274.
31. Barthélemy, N.R., et al., Blood plasma phosphorylated-tau isoforms track CNS change in Alzheimer's disease. *J Exp Med*, 2020. 217(11).
32. Mattsson-Carlsson, N., et al., Longitudinal plasma p-tau217 is increased in early stages of Alzheimer's disease. *Brain*, 2020. 143(11): p. 3234-3241.
33. Wennström, M., et al., Cellular localization of p-tau217 in brain and its association with p-tau217 plasma levels. *Acta Neuropathol Commun*, 2022. 10(1): p. 3.
34. Hampel, H., et al., Biomarkers for Alzheimer's disease: academic, industry and regulatory perspectives. *Nat Rev Drug Discov*, 2010. 9(7): p. 560-74.
35. Food and Drug Administration (2018) Early Alzheimer's disease: developing drugs for treatment; draft guidance for industry. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/alzheimers-disease-developing-drugs-treatment-guidance-industry>
36. European Medicines Agency, Committee for Medicinal Products for Human Use (2018) Guideline on the clinical investigation of medicines for the treatment of Alzheimer's disease. [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2018/02/WC500244609.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2018/02/WC500244609.pdf)
37. Bertram, L. and H. Hampel, The role of genetics for biomarker development in neurodegeneration. *Prog Neurobiol*, 2011. 95(4): p. 501-4.
38. Zetzsche, T., et al., Advances and perspectives from genetic research: development of biological markers in Alzheimer's disease. *Expert Rev Mol Diagn*, 2010. 10(5): p. 667-90.
39. Teipel, S.J., et al., Relevance of magnetic resonance imaging for early detection and diagnosis of Alzheimer disease. *Med Clin North Am*, 2013. 97(3): p. 399-424.
40. Ewers, M., et al., Neuroimaging markers for the prediction and early diagnosis of Alzheimer's disease dementia. *Trends Neurosci*, 2011. 34(8): p. 430-42.
41. Blennow, K., et al., Clinical utility of cerebrospinal fluid biomarkers in the diagnosis of early Alzheimer's disease. *Alzheimers Dement*, 2015. 11(1): p. 58-69.
42. Snyder, H.M., et al., Developing novel blood-based biomarkers for Alzheimer's disease. *Alzheimers Dement*, 2014. 10(1): p. 109-14.
43. Teipel, S.J., et al., Perspectives for multimodal neurochemical and imaging biomarkers in Alzheimer's disease. *J Alzheimers Dis*, 2013. 33 Suppl 1: p. S329-47.
44. Beach, T.G., et al., Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on Aging Alzheimer Disease Centers, 2005-2010. *J Neuropathol Exp Neurol*, 2012. 71(4): p. 266-73.
45. Archer, M.C., P.H. Hall, and J.C. Morgan, ACCURACY OF CLINICAL DIAGNOSIS OF ALZHEIMER'S DISEASE IN ALZHEIMER'S DISEASE CENTERS (ADCS). *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 2017. 13(7): p. P800-P801.
46. Franklin, E.E., et al., Brain collection, standardized neuropathologic assessment, and comorbidity in Alzheimer's Disease Neuroimaging Initiative 2 participants. *Alzheimers Dement*, 2015. 11(7): p. 815-22.
47. Hjalmarsson, C., et al., Neuronal and glia-related biomarkers in cerebrospinal fluid of patients with acute ischemic stroke. *J Cent Nerv Syst Dis*, 2014. 6: p. 51-8.
48. Lattanzio, F., et al., Prion-specific and surrogate CSF biomarkers in Creutzfeldt-Jakob disease: diagnostic accuracy in relation to molecular subtypes and analysis of neuropathological correlates of p-tau and A $\beta$ 42 levels. *Acta Neuropathologica*, 2017. 133(4): p. 559-578.
49. Robinson, M.R., R.A. Miller, and D.S. Spellman, Mass Spectrometry-Based Biomarkers in Drug Development. *Adv Exp Med Biol*, 2019. 1140: p. 435-449.

- 
50. Hosp, F. and M. Mann, A Primer on Concepts and Applications of Proteomics in Neuroscience. *Neuron*, 2017. 96(3): p. 558-571.
  51. Hartwell, L.H., et al., From molecular to modular cell biology. *Nature*, 1999. 402(6761 Suppl): p. C47-52.
  52. Wilkins, M.R., et al., Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol Genet Eng Rev*, 1996. 13: p. 19-50.
  53. Aebersold, R. and M. Mann, Mass spectrometry-based proteomics. *Nature*, 2003. 422(6928): p. 198-207.
  54. Cravatt, B.F., G.M. Simon, and J.R. Yates, 3rd, The biological impact of mass-spectrometry-based proteomics. *Nature*, 2007. 450(7172): p. 991-1000.
  55. Mann, M., et al., The coming age of complete, accurate, and ubiquitous proteomes. *Mol Cell*, 2013. 49(4): p. 583-90.
  56. Yates, J.R., C.I. Ruse, and A. Nakorchevsky, Proteomics by Mass Spectrometry: Approaches, Advances, and Applications. *Annual Review of Biomedical Engineering*, 2009. 11(1): p. 49-79.
  57. Mann, M. and O.N. Jensen, Proteomic analysis of post-translational modifications. *Nat Biotechnol*, 2003. 21(3): p. 255-61.
  58. Doll, S. and A.L. Burlingame, Mass spectrometry-based detection and assignment of protein posttranslational modifications. *ACS Chem Biol*, 2015. 10(1): p. 63-71.
  59. Gingras, A.C., et al., Analysis of protein complexes using mass spectrometry. *Nat Rev Mol Cell Biol*, 2007. 8(8): p. 645-54.
  60. Zhang, Y., et al., Protein Analysis by Shotgun/Bottom-up Proteomics. *Chemical Reviews*, 2013. 113(4): p. 2343-2394.
  61. Kelstrup, C.D., et al., Performance Evaluation of the Q Exactive HF-X for Shotgun Proteomics. *J Proteome Res*, 2018. 17(1): p. 727-738.
  62. Michalski, A., et al., Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Mol Cell Proteomics*, 2011. 10(9): p. M111.011015.
  63. Yates, J.R., 3rd, The revolution and evolution of shotgun proteomics for large-scale proteome analysis. *J Am Chem Soc*, 2013. 135(5): p. 1629-40.
  64. Fenn, J.B., et al., Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 1989. 246(4926): p. 64-71.
  65. Yamashita, M. and J.B. Fenn, Electrospray ion source. Another variation on the free-jet theme. *The Journal of Physical Chemistry*, 1984. 88(20): p. 4451-4459.
  66. Tanaka, K., et al., Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, 1988. 2(8): p. 151-153.
  67. Aebersold, R. and M. Mann, Mass-spectrometric exploration of proteome structure and function. *Nature*, 2016. 537: p. 347.
  68. Tolley, L., J.W. Jorgenson, and M.A. Moseley, Very high pressure gradient LC/MS/MS. *Anal Chem*, 2001. 73(13): p. 2985-91.
  69. Hunt, D.F., et al., Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science*, 1992. 255(5049): p. 1261-3.
  70. Zhang, Y., et al., Protein analysis by shotgun/bottom-up proteomics. *Chem Rev*, 2013. 113(4): p. 2343-94.
  71. Olsen, J.V., et al., Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods*, 2007. 4(9): p. 709-12.
  72. Steen, H. and M. Mann, The ABC's (and XYZ's) of peptide sequencing. *Nat Rev Mol Cell Biol*, 2004. 5(9): p. 699-711.
  73. Eliuk, S. and A. Makarov, Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annu Rev Anal Chem (Palo Alto Calif)*, 2015. 8: p. 61-80.
  74. Andrews, G.L., et al., Performance characteristics of a new hybrid quadrupole time-of-flight tandem mass spectrometer (TripleTOF 5600). *Anal Chem*, 2011. 83(13): p. 5442-6.
  75. Beck, S., et al., The Impact II, a Very High-Resolution Quadrupole Time-of-Flight Instrument (QTOF) for Deep Shotgun Proteomics. *Mol Cell Proteomics*, 2015. 14(7): p. 2014-29.
  76. Harper, J.W. and E.J. Bennett, Proteome complexity and the forces that drive proteome imbalance. *Nature*, 2016. 537(7620): p. 328-38.
  77. Schwanhauser, B., et al., Global quantification of mammalian gene expression control. *Nature*, 2011. 473(7347): p. 337-42.
  78. Schwenk, J.M., et al., The Human Plasma Proteome Draft of 2017: Building on the Human Plasma PeptideAtlas from Mass Spectrometry and Complementary Assays. *J Proteome Res*, 2017. 16(12): p. 4299-4310.
  79. Michalski, A., J. Cox, and M. Mann, More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. *J Proteome Res*, 2011. 10(4): p. 1785-93.

80. Gillet, L.C., A. Leitner, and R. Aebersold, Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing. *Annu Rev Anal Chem (Palo Alto Calif)*, 2016. 9(1): p. 449-72.
81. Ong, S.E., et al., Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics*, 2002. 1(5): p. 376-86.
82. Ong, S.E., L.J. Foster, and M. Mann, Mass spectrometric-based approaches in quantitative proteomics. *Methods*, 2003. 29(2): p. 124-30.
83. Thompson, A., et al., Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem*, 2003. 75(8): p. 1895-904.
84. Gerber, S.A., et al., Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci U S A*, 2003. 100(12): p. 6940-5.
85. O'Connell, J.D., et al., Proteome-Wide Evaluation of Two Common Protein Quantification Methods. *J Proteome Res*, 2018. 17(5): p. 1934-1942.
86. Tabb, D.L., et al., Reproducibility of Differential Proteomic Technologies in CPTAC Fractionated Xenografts. *Journal of proteome research*, 2016. 15(3): p. 691-706.
87. Nahnsen, S., et al., Tools for Label-free Peptide Quantification. *Molecular & Cellular Proteomics*, 2013. 12(3): p. 549.
88. Cox, J. and M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*, 2008. 26(12): p. 1367-72.
89. Claassen, M., et al., Generic Comparison of Protein Inference Engines. *Molecular & Cellular Proteomics*, 2012. 11(4): p. O110.007088.
90. Nesvizhskii, A.I. and R. Aebersold, Interpretation of shotgun proteomic data: the protein inference problem. *Mol Cell Proteomics*, 2005. 4(10): p. 1419-40.
91. Ting, Y.S., et al., Peptide-Centric Proteome Analysis: An Alternative Strategy for the Analysis of Tandem Mass Spectrometry Data. *Mol Cell Proteomics*, 2015. 14(9): p. 2301-7.
92. Wilm, M. and M. Mann, Analytical properties of the nanoelectrospray ion source. *Anal Chem*, 1996. 68(1): p. 1-8.
93. Zhou, F., et al., Nanoflow low pressure high peak capacity single dimension LC-MS/MS platform for high-throughput, in-depth analysis of mammalian proteomes. *Anal Chem*, 2012. 84(11): p. 5133-9.
94. MacNair, J.E., K.C. Lewis, and J.W. Jorgenson, Ultrahigh-pressure reversed-phase liquid chromatography in packed capillary columns. *Anal Chem*, 1997. 69(6): p. 983-9.
95. Motoyama, A., et al., Automated ultra-high-pressure multidimensional protein identification technology (UHP-MudPIT) for improved peptide identification of proteomic samples. *Anal Chem*, 2006. 78(14): p. 5109-18.
96. Chen, C.J., et al., Tunnel frit: a nonmetallic in-capillary frit for nanoflow ultra high-performance liquid chromatography-mass spectrometry applications. *Anal Chem*, 2012. 84(1): p. 297-303.
97. Iwasaki, M., et al., Human proteome analysis by using reversed phase monolithic silica capillary columns with enhanced sensitivity. *J Chromatogr A*, 2012. 1228: p. 292-7.
98. Shi, T., et al., Long-gradient separations coupled with selected reaction monitoring for highly sensitive, large scale targeted protein quantification in a single analysis. *Anal Chem*, 2013. 85(19): p. 9196-203.
99. Burgess, M.W., et al., Simplified and efficient quantification of low-abundance proteins at very high multiplex via targeted mass spectrometry. *Mol Cell Proteomics*, 2014. 13(4): p. 1137-49.
100. Nakamura, T., J. Kuromitsu, and Y. Oda, Evaluation of comprehensive multidimensional separations using reversed-phase, reversed-phase liquid chromatography/mass spectrometry for shotgun proteomics. *J Proteome Res*, 2008. 7(3): p. 1007-11.
101. Gilar, M., et al., Orthogonality of separation in two-dimensional liquid chromatography. *Anal Chem*, 2005. 77(19): p. 6426-34.
102. Essader, A.S., et al., A comparison of immobilized pH gradient isoelectric focusing and strong-cation-exchange chromatography as a first dimension in shotgun proteomics. *Proteomics*, 2005. 5(1): p. 24-34.
103. Dai, J., et al., Proteomic analysis with integrated multiple dimensional liquid chromatography/mass spectrometry based on elution of ion exchange column using pH steps. *Anal Chem*, 2005. 77(18): p. 5793-9.
104. Zhou, H., et al., A fully automated 2-D LC-MS method utilizing online continuous pH and RP gradients for global proteome analysis. *Electrophoresis*, 2007. 28(23): p. 4311-9.
105. Gilar, M., et al., Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions. *J Sep Sci*, 2005. 28(14): p. 1694-703.
106. Boersema, P.J., S. Mohammed, and A.J. Heck, Hydrophilic interaction liquid chromatography (HILIC) in proteomics. *Anal Bioanal Chem*, 2008. 391(1): p. 151-9.

107. Boersema, P.J., et al., Evaluation and optimization of ZIC-HILIC-RP as an alternative MudPIT strategy. *J Proteome Res*, 2007. 6(3): p. 937-46.
108. Hao, P., et al., Novel application of electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) in shotgun proteomics: comprehensive profiling of rat kidney proteome. *J Proteome Res*, 2010. 9(7): p. 3520-6.
109. Hao, P., T. Guo, and S.K. Sze, Simultaneous analysis of proteome, phospho- and glycoproteome of rat kidney tissue with electrostatic repulsion hydrophilic interaction chromatography. *PLoS One*, 2011. 6(2): p. e16884.
110. Pfammatter, S., et al., A Novel Differential Ion Mobility Device Expands the Depth of Proteome Coverage and the Sensitivity of Multiplex Proteomic Measurements. *Mol Cell Proteomics*, 2018. 17(10): p. 2051-2067.
111. Meier, F., et al., Parallel Accumulation–Serial Fragmentation (PASEF): Multiplying Sequencing Speed and Sensitivity by Synchronized Scans in a Trapped Ion Mobility Device. *Journal of Proteome Research*, 2015. 14(12): p. 5378-5387.
112. Meier, F., et al., Online Parallel Accumulation–Serial Fragmentation (PASEF) with a Novel Trapped Ion Mobility Mass Spectrometer. *Molecular & Cellular Proteomics*, 2018. 17(12): p. 2534.
113. Meier, F., et al., BoxCar acquisition method enables single-shot proteomics at a depth of 10,000 proteins in 100 minutes. *Nat Methods*, 2018. 15(6): p. 440-448.
114. Stolz, A., et al., Recent advances in capillary electrophoresis-mass spectrometry: Instrumentation, methodology and applications. *Electrophoresis*, 2019. 40(1): p. 79-112.
115. Gomes, F.P. and J.R. Yates, 3rd, Recent trends of capillary electrophoresis-mass spectrometry in proteomics research. *Mass Spectrom Rev*, 2019. 38(6): p. 445-460.
116. Jansson, E.T., Strategies for analysis of isomeric peptides. *J Sep Sci*, 2018. 41(1): p. 385-397.
117. Haselberg, R., G.J. de Jong, and G.W. Somsen, CE-MS for the analysis of intact proteins 2010-2012. *Electrophoresis*, 2013. 34(1): p. 99-112.
118. Pontillo, C., et al., CE-MS-based proteomics in biomarker discovery and clinical application. *Proteomics Clin Appl*, 2015. 9(3-4): p. 322-34.
119. Faserl, K., et al., Investigating capillary electrophoresis-mass spectrometry for the analysis of common post-translational modifications. *Electrophoresis*, 2018. 39(9-10): p. 1208-1215.
120. Varadi, C., et al., Quantitative twoplex glycan analysis using (12)C6 and (13)C6 stable isotope 2-aminobenzoic acid labelling and capillary electrophoresis mass spectrometry. *Anal Bioanal Chem*, 2016. 408(30): p. 8691-8700.
121. del Campo, M., et al., Recommendations to standardize preanalytical confounding factors in Alzheimer's and Parkinson's disease cerebrospinal fluid biomarkers: an update. *Biomark Med*, 2012. 6(4): p. 419-30.
122. Leitao, M.J., et al., Chasing the Effects of Pre-Analytical Confounders - A Multicenter Study on CSF-AD Biomarkers. *Front Neurol*, 2015. 6: p. 153.
123. Fourier, A., et al., Pre-analytical and analytical factors influencing Alzheimer's disease cerebrospinal fluid biomarker variability. *Clin Chim Acta*, 2015. 449: p. 9-15.
124. Le Bastard, N., P.P. De Deyn, and S. Engelborghs, Importance and impact of preanalytical variables on Alzheimer disease biomarker concentrations in cerebrospinal fluid. *Clin Chem*, 2015. 61(5): p. 734-43.
125. Comstock, G.W., et al., Effects of repeated freeze-thaw cycles on concentrations of cholesterol, micronutrients, and hormones in human plasma and serum. *Am J Epidemiol*, 2008. 168(7): p. 827-30.
126. Bateman, R.J., et al., Fluctuations of CSF amyloid- $\beta$  levels. Implications for a diagnostic and therapeutic biomarker, 2007. 68(9): p. 666-669.
127. Vanderstichele, H.M., et al., Optimized Standard Operating Procedures for the Analysis of Cerebrospinal Fluid Abeta42 and the Ratios of Abeta Isoforms Using Low Protein Binding Tubes. *J Alzheimers Dis*, 2016. 53(3): p. 1121-32.
128. Lewczuk, P., et al., Effect of Sample Collection Tubes on Cerebrospinal Fluid Concentrations of Tau Proteins and Amyloid  $\beta$  Peptides. *Clinical Chemistry*, 2020. 52(2): p. 332-334.
129. Perret-Liaudet, A., et al., Risk of Alzheimer's disease biological misdiagnosis linked to cerebrospinal collection tubes. *J Alzheimers Dis*, 2012. 31(1): p. 13-20.
130. You, J.S., et al., The impact of blood contamination on the proteome of cerebrospinal fluid. *Proteomics*, 2005. 5(1): p. 290-6.
131. Bjerke, M., et al., Confounding factors influencing amyloid Beta concentration in cerebrospinal fluid. *Int J Alzheimers Dis*, 2010. 2010.
132. Schoonenboom, N.S., et al., Effects of processing and storage conditions on amyloid beta (1-42) and tau concentrations in cerebrospinal fluid: implications for use in clinical practice. *Clin Chem*, 2005. 51(1): p. 189-95.

133. Vanderstichele, H., et al., Standardization of preanalytical aspects of cerebrospinal fluid biomarker testing for Alzheimer's disease diagnosis: a consensus paper from the Alzheimer's Biomarkers Standardization Initiative. *Alzheimers Dement*, 2012. 8(1): p. 65-73.
134. Zimmermann, R., et al., Preanalytical sample handling and sample stability testing for the neurochemical dementia diagnostics. *J Alzheimers Dis*, 2011. 25(4): p. 739-45.
135. West-Nielsen, M., et al., Sample handling for mass spectrometric proteomic investigations of human sera. *Anal Chem*, 2005. 77(16): p. 5114-23.
136. Hokfelt, T., et al., Neuropeptides--an overview. *Neuropharmacology*, 2000. 39(8): p. 1337-56.
137. Reiber, H., Dynamics of brain-derived proteins in cerebrospinal fluid. *Clin Chim Acta*, 2001. 310(2): p. 173-86.
138. Engelborghs, S., et al., Consensus guidelines for lumbar puncture in patients with neurological diseases. *Alzheimers Dement (Amst)*, 2017. 8: p. 111-126.
139. Nath, S., et al., Atraumatic versus conventional lumbar puncture needles: a systematic review and meta-analysis. *Lancet*, 2018. 391(10126): p. 1197-1204.
140. Peskind, E., et al., Safety of lumbar puncture procedures in patients with Alzheimer's disease. *Curr Alzheimer Res*, 2009. 6(3): p. 290-2.
141. Zetterberg, H., et al., Low incidence of post-lumbar puncture headache in 1,089 consecutive memory clinic patients. *Eur Neurol*, 2010. 63(6): p. 326-30.
142. Yuan, X. and D.M. Desiderio, Proteomics analysis of human cerebrospinal fluid. *Journal of Chromatography B*, 2005. 815(1): p. 179-189.
143. Schutzer, S.E., et al., Establishing the Proteome of Normal Human Cerebrospinal Fluid. *PLOS ONE*, 2010. 5(6): p. e10980.
144. Kroksveen, A.C., et al., Proteomics of human cerebrospinal fluid: Discovery and verification of biomarker candidates in neurodegenerative diseases using quantitative proteomics. *Journal of Proteomics*, 2011. 74(4): p. 371-388.
145. Wu, C., et al., Contributions of immunoaffinity chromatography to deep proteome profiling of human biofluids. *Journal of Chromatography B*, 2016. 1021: p. 57-68.
146. Pieper, R., et al., Multi-component immunoaffinity subtraction chromatography: An innovative step towards a comprehensive survey of the human plasma proteome. *Proteomics*, 2003. 3(4): p. 422-432.
147. Liu, T., et al., Evaluation of multiprotein immunoaffinity subtraction for plasma proteomics and candidate biomarker discovery using mass spectrometry. *Molecular and Cellular Proteomics*, 2006. 5(11): p. 2167-2174.
148. Boschetti, E. and P.G. Righetti, The ProteoMiner in the proteomic arena: A non-depleting tool for discovering low-abundance species. *Journal of Proteomics*, 2008. 71(3): p. 255-264.
149. Righetti, P.G., et al., Combinatorial peptide ligand libraries as a "trojan horse" in deep discovery proteomics. *Analytical Chemistry*, 2015. 87(1): p. 293-305.
150. Jankovska, E., et al., Affinity depletion versus relative protein enrichment: a side-by-side comparison of two major strategies for increasing human cerebrospinal fluid proteome coverage. *Clinical Proteomics*, 2019. 16(1): p. 9.
151. Pedrero-Prieto, C.M., et al., A comprehensive systematic review of CSF proteins and peptides that define Alzheimer's disease. *Clinical Proteomics*, 2020. 17(1): p. 21.
152. Kaneko, N., et al., Novel plasma biomarker surrogating cerebral amyloid deposition. *Proc Jpn Acad Ser B Phys Biol Sci*, 2014. 90(9): p. 353-64.
153. Kvarstberg, H., et al., Cerebrospinal fluid levels of the synaptic protein neurogranin correlates with cognitive decline in prodromal Alzheimer's disease. *Alzheimers Dement*, 2015. 11(10): p. 1180-90.
154. Sjodin, S., et al., Targeting LAMP2 in human cerebrospinal fluid with a combination of immunopurification and high resolution parallel reaction monitoring mass spectrometry. *Clin Proteomics*, 2016. 13: p. 4.
155. Ringman, J.M., et al., Proteomic changes in cerebrospinal fluid of presymptomatic and affected persons carrying familial Alzheimer disease mutations. *Arch Neurol*, 2012. 69(1): p. 96-104.
156. Holtta, M., et al., An integrated workflow for multiplex CSF proteomics and peptidomics-identification of candidate cerebrospinal fluid biomarkers of Alzheimer's disease. *J Proteome Res*, 2015. 14(2): p. 654-63.
157. Russell, C.L., et al., Combined tissue and fluid proteomics with Tandem Mass Tags to identify low-abundance protein biomarkers of disease in peripheral body fluid: An Alzheimer's Disease case study. *Rapid Commun Mass Spectrom*, 2017. 31(2): p. 153-159.
158. Wijte, D., et al., A novel peptidomics approach to detect markers of Alzheimer's disease in cerebrospinal fluid. *Methods*, 2012. 56(4): p. 500-7.
159. Holtta, M., et al., A single dose of the gamma-secretase inhibitor semagacestat alters the cerebrospinal fluid peptidome in humans. *Alzheimers Res Ther*, 2016. 8(1): p. 11.

160. Perrin, R.J., et al., Quantitative label-free proteomics for discovery of biomarkers in cerebrospinal fluid: assessment of technical and inter-individual variation. *PLoS One*, 2013. 8(5): p. e64314.
161. Jahn, H., et al., Peptide fingerprinting of Alzheimer's disease in cerebrospinal fluid: identification and prospective evaluation of new synaptic biomarkers. *PLoS One*, 2011. 6(10): p. e26540.
162. Tagami, S., et al., Relative ratio and level of amyloid-beta 42 surrogate in cerebrospinal fluid of familial Alzheimer disease patients with presenilin 1 mutations. *Neurodegener Dis*, 2014. 13(2-3): p. 166-70.
163. Chiasserini, D., et al., Proteomic analysis of cerebrospinal fluid extracellular vesicles: A comprehensive dataset. *Journal of Proteomics*, 2014. 106: p. 191-204.
164. Dayon, L., et al., Relative Quantification of Proteins in Human Cerebrospinal Fluids by MS/MS Using 6-Plex Isobaric Tags. *Analytical Chemistry*, 2008. 80(8): p. 2921-2931.
165. Shih, Y.H., et al., Apolipoprotein C-III is an amyloid-beta-binding protein and an early marker for Alzheimer's disease. *J Alzheimers Dis*, 2014. 41(3): p. 855-65.
166. Muenchhoff, J., et al., Plasma protein profiling of mild cognitive impairment and Alzheimer's disease across two independent cohorts. *J Alzheimers Dis*, 2015. 43(4): p. 1355-73.
167. Guntert, A., et al., Plasma gelsolin is decreased and correlates with rate of decline in Alzheimer's disease. *J Alzheimers Dis*, 2010. 21(2): p. 585-96.
168. Shen, L., et al., Proteomics Analysis of Blood Serums from Alzheimer's Disease Patients Using iTRAQ Labeling Technology. *J Alzheimers Dis*, 2017. 56(1): p. 361-378.
169. Lundstrom, S.L., et al., Blood plasma IgG Fc glycans are significantly altered in Alzheimer's disease and progressive mild cognitive impairment. *J Alzheimers Dis*, 2014. 38(3): p. 567-79.
170. Poljak, A., et al., Quantitative proteomics of delirium cerebrospinal fluid. *Transl Psychiatry*, 2014. 4: p. e477.
171. Rogeberg, M., et al., Isobaric Quantification of Cerebrospinal Fluid Amyloid-beta Peptides in Alzheimer's Disease: C-Terminal Truncation Relates to Early Measures of Neurodegeneration. *J Proteome Res*, 2015. 14(11): p. 4834-43.
172. Rezeli, M., et al., Quantification of total apolipoprotein E and its specific isoforms in cerebrospinal fluid and blood in Alzheimer's disease and other neurodegenerative diseases. *EuPA Open Proteomics*, 2015. 8: p. 137-143.
173. Simon, R., et al., Total ApoE and ApoE4 isoform assays in an Alzheimer's disease case-control study by targeted mass spectrometry (n=669): a pilot assay for methionine-containing proteotypic peptides. *Molecular & cellular proteomics : MCP*, 2012. 11(11): p. 1389-1403.
174. Lundström, S.L., et al., SpotLight Proteomics: uncovering the hidden blood proteome improves diagnostic power of proteomics. *Scientific Reports*, 2017. 7(1): p. 41929.
175. Portelius, E., et al., A novel A $\beta$  isoform pattern in CSF reflects gamma-secretase inhibition in Alzheimer disease. *Alzheimers Res Ther*, 2010. 2(2): p. 7.
176. Portelius, E., et al.,  $\beta$ -site amyloid precursor protein-cleaving enzyme 1 (BACE1) inhibitor treatment induces A $\beta$ 5-X peptides through alternative amyloid precursor protein cleavage. *Alzheimer's Research & Therapy*, 2014. 6(5): p. 75.
177. Portelius, E., et al., Characterization of amyloid beta peptides in cerebrospinal fluid by an automated immunoprecipitation procedure followed by mass spectrometry. *J Proteome Res*, 2007. 6(11): p. 4433-9.
178. Brinkmalm, A., et al., SNAP-25 is a promising novel cerebrospinal fluid biomarker for synapse degeneration in Alzheimer's disease. *Molecular Neurodegeneration*, 2014. 9(1): p. 53.
179. Ohrfelt, A., et al., The pre-synaptic vesicle protein synaptotagmin is a novel biomarker for Alzheimer's disease. *Alzheimers Res Ther*, 2016. 8(1): p. 41.
180. Martinez-Morillo, E., et al., Total apolipoprotein E levels and specific isoform composition in cerebrospinal fluid and plasma from Alzheimer's disease patients and controls. *Acta Neuropathol*, 2014. 127(5): p. 633-43.
181. Heslegrave, A., et al., Increased cerebrospinal fluid soluble TREM2 concentration in Alzheimer's disease. *Mol Neurodegener*, 2016. 11: p. 3.
182. <https://proteomics.cancer.gov/sites/default/files/assay-characterization-guidance-document.pdf>
183. Mueller, S.G., et al., The Alzheimer's Disease Neuroimaging Initiative. *Neuroimaging Clinics of North America*, 2005. 15(4): p. 869-877.
184. Weiner, M.W., et al., The Alzheimer's disease neuroimaging initiative: progress report and future plans. *Alzheimer's & dementia : the journal of the Alzheimer's Association*, 2010. 6(3): p. 202-11.e7.
185. Weiner, M.W., et al., Impact of the Alzheimer's Disease Neuroimaging Initiative, 2004 to 2014. *Alzheimer's & dementia : the journal of the Alzheimer's Association*, 2015. 11(7): p. 865-884.
186. Weiner, M.W., et al., The Alzheimer's Disease Neuroimaging Initiative 3: Continued innovation for clinical trial improvement. *Alzheimer's & dementia : the journal of the Alzheimer's Association*, 2017. 13(5): p. 561-571.

- 
187. Toga, A.W. and K.L. Crawford, The Alzheimer's Disease Neuroimaging Initiative informatics core: A decade in review. *Alzheimer's & dementia : the journal of the Alzheimer's Association*, 2015. 11(7): p. 832-839.
  188. Thompson, P.M., et al., The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. *Brain imaging and behavior*, 2014. 8(2): p. 153-182.
  189. "Big Data Challenge for Alzheimer's Disease Launches in Global Effort to Use Innovative Open Science Techniques to Improve Diagnosis and Treatment | Global CEO Initiative on Alzheimer's Disease". Archived from the original on 2017-01-06.
  190. Spellman, D.S., et al., Development and evaluation of a multiplexed mass spectrometry based assay for measuring candidate peptide biomarkers in Alzheimer's Disease Neuroimaging Initiative (ADNI) CSF. *Proteomics Clin Appl*, 2015. 9(7-8): p. 715-31.
  191. Libiger, O., et al., Longitudinal CSF proteomics identifies NPTX2 as a prognostic biomarker of Alzheimer's disease. *Alzheimers Dement*, 2021. 17(12): p. 1976-1987.