Novel Fluid Biomarkers to Differentiate Frontotemporal Dementia and Dementia with Lewy Bodies from Alzheimer’s Disease: A Systematic Review

Aiysha Chaudhry 1, Henry Houlden 1, Mie Rizig 1

1 UCL Queen Square Institute of Neurology, University College London, Queen Square, London WC1N 3BG, 020 3456 7890

Corresponding author: m.rizig@ucl.ac.uk

Abstract

Rationale: Frontotemporal dementia (FTD) and dementia with Lewy bodies (DBL) are two common forms of neurodegenerative dementia, subsequent to Alzheimer’s disease (AD). AD is the only dementia that includes clinically validated cerebrospinal fluid (CSF) biomarkers in the diagnostic criteria. FTD and DBL often overlap with AD in their clinical and pathological features, making it challenging to differentiate between these conditions.

Aim: This systematic review aimed to identify if novel fluid biomarkers are useful in differentiating FTD and DBL from AD. Increasing the certainty of the differentiation between dementia subtypes would be advantageous clinically and in research.

Methods: PubMed and Scopus were searched for studies that quantified and assessed diagnostic accuracy of novel fluid biomarkers in clinically diagnosed patients with FTD or DBL, in comparison to patients with AD. Meta-analyses were performed on biomarkers that were quantified in 3 studies or more.

Results: The search strategy yielded 614 results, from which, 27 studies were included. When comparing bio-fluid levels in AD and FTD patients, neurofilament light chain (NfL) level was often higher in FTD, whilst brain soluble amyloid precursor protein β (sAPPβ) was higher in patients with AD. When comparing bio-fluid levels in AD and DBL patients, α-synuclein ensued heterogeneous findings, while the noradrenaline metabolite (MHPG) was found to be lower in DBL. Ratios of Aβ42/Aβ38 and Aβ42/Aβ40 were lower in AD than FTD and DBL and offered better diagnostic accuracy than raw amyloid-β (Aβ) concentrations.

Conclusions: Several promising novel biomarkers were highlighted in this review. Combinations of fluid biomarkers were more often useful than individual biomarkers in distinguishing subtypes of dementia. Considering the heterogeneity in methods and results between the studies, further validation, ideally with longitudinal prospective designs with large sample sizes and unified protocols, are fundamental before conclusions can be finalised.

Key words

Dementia with Lewy bodies; frontotemporal dementia; biomarkers; cerebrospinal fluid; blood; Alzheimer’s disease; bio-fluid

Abbreviations

Dementia with Lewy bodies (DBL); Alzheimer’s disease (AD); frontotemporal dementia (FTD); Lewy bodies (LB); frontotemporal lobar degeneration (FTLD); Amyloid beta (Aβ); behavioural variant FTD (bvFTD); primary progressive aphasia (PPA); semantic dementia (SD); corticobasal degeneration (CBD); progressive supranuclear palsy (PSP); Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2); confidence interval (CI); interquartile range (IQR); standard deviation (SD); standardised mean difference (SMD); area under curve value (AUC)
1. Introduction

Neurodegenerative dementias are a group of syndromes distinguishable by their underlying proteinopathy, and involve an interplay of molecular pathways, resulting in synaptic loss, gliosis, inflammation, and cell death. This progressively disrupts networks for cognition, behaviour or sensorimotor functions (Lashley, 2018; Elahi, 2017). The most common is Alzheimer’s disease (AD), followed by dementia with Lewy bodies (DLB) and frontotemporal dementia (FTD) (Prince, 2014; NICE, 2017).

AD is characterised by a dual proteinopathy: amyloid plaques are composed of amyloid beta peptides (Aβ), the most prevalent being Aβ1-40 (Aβ40) and Aβ1-42 (Aβ42), and neurofibrillary tangles (NFTs) are composed of highly phosphorylated microtubule-associated protein tau (MAPT) (Hardy, 1992; Barage, 2015; Brion, 1998). FTD is characterised by frontotemporal lobar degeneration (FTLD), which includes three main pathological subtypes: FTLD-tau, FTLD-TDP, and FTLD-FUS (Seelar, 2011; Boxer, 2014; Mackenzie, 2016; Mackenzie, 2011). DLB, Parkinson’s disease (PD) and PD dementia (PDD), are Lewy body diseases (LBD), involving α-synuclein neuronal inclusions: Lewy bodies (LB) and Lewy neurites (LNs) (Donaghy, 2014; Beyer, 2009). The main clinico-pathological features and genetic characteristics of AD, FTD, and DLB are shown in figure (1) and diagnostic criteria are presented in supplementary (1).

The National Institute of Neurological and Communicative Disorders and Stroke—Alzheimer’s Disease and Related Disorders Association (NINCDS-ADRDA) proposed AD diagnostic criteria in 1984, which were revised in 2011 to integrate cerebrospinal fluid (CSF) biomarkers (McKhann, 1984; McKhann, 2011). FTLD diagnostic criteria were proposed in 1998 including behavioural (FTD) and language syndromes: progressive nonfluent aphasia (PNFA), and semantic dementia (SD) (Neary, 1998). This was revised into separate criteria in 2011. Rascovsky, et al characterised behavioural variant FTD (bvFTD), whilst Gorno-Tempini, et al described the three variants of primary progressive aphasia (PPA) (Rascovsky, 2011; Gorno-Tempini, 2011). Definitive diagnosis relies on post-mortem histopathology or the presence of a known mutation. FTD overlaps with motor neuron disease (MND) or atypical Parkinsonism, i.e. corticobasal degeneration/syndrome (CBD/CBS), progressive supranuclear palsy (PSP) (Sivasthiaseelan, 2019). The DLB Consortium proposed diagnostic criteria in 1996 and revised these in 2005 (McKieth, 1996, McKieth, 2005). Biomarkers were integrated in 2017. DLB is characterised by fluctuating cognition, visual hallucinations, and parkinsonism. Definitive diagnosis relies on post-mortem histopathology (McKieth, 2017; Huey, 2015).
Figure 1: Main features of AD, FTD, & DLB

**AD**
- Amnestic presentation has progressive learning & memory impairments.
- Non-amnestic presentation can involve executive, language, or visuospatial dysfunction.
- Familial AD: ~5% of AD (typically arise following mutations APP, PSEN1, or PSEN2 genes)
- Sporadic AD: ~90% of AD
- \(\text{A}_{\beta_{42}} \) & \(\text{A}_{\beta_{40}}\)
- Amyloid plaques
- Neurofibrillary tangles

**FTD**
- Behavioural variant (bvFTD) involves behavioural disinhibition & personality changes.
- Language phenotypes (PPA) can influence word memory, speech production, comprehension, and grammar.
- A third to a half of FTD has autosomal dominant inheritance. Mutations may occur in GRN, MAPT, VCP, TARDP, FUS, and CHMP2B.
- Tau
- Hyperphosphorylated tau inclusions
- FUS
- FUS inclusions
- TDP-43
- TDP-43 inclusions

**DLB**
- Typically presents with fluctuating cognition, visual hallucinations, REM sleep behaviour disorder, & features of Parkinsonism.
- Lewy bodies & neurites
- \(\alpha\)-synuclein

---

Abbreviations: APP = amyloid precursor protein; PSEN1/2 = presenilin 1/2; GRN = progranulin; MAPT = microtubule associated protein tau; VCP = Valosin-containing protein; TARDP = TAR-DNA binding protein; FUS = fused in sarcoma; CHMP2B = chromatin-modifying protein 2B; SNCA = \(\alpha\)-synuclein; LRRK2 = leucine-rich repeat kinase 2; SNCB = \(\beta\)-synuclein

Information for clinical features from: McKhann, 2011; Rascovsky, 2011; Gorno-Tempini, 2011; McKieth, 2017.
The National Institute for Health and Care Excellence (NICE) currently recommend validated clinical criteria for AD, FTD, and DLB (NICE, 2018; McKhann, 2011; Rascovský, 2011; Gorno-Tempini, 2011; McKieth, 2017). Whilst detecting dementia is typically straightforward, diagnosing subtypes remain challenging (Bayer, 2018).

There are several overlapping clinical features between FTD and AD. For example, hallucinations, delusions, executive dysfunction, and episodic memory problems present similarly (Karantzoulis, 2011). Memory performance analysis showed bvFTD and early AD patients had similar memory test scores, which was confirmed by a meta-analytic review of 94 studies that found AD and FTD had overlapping cognitive test performance. (Hornberger, 2010; Hutchinson, 2007). Another study that evaluated behavioural and neuropsychological symptoms found both groups showed unawareness in most domains (Salmon, 2008). Impairments in semantic and phonematic fluency, backwards digit span, and emotional recognition also often present similarly (Ruel, 2017). Moreover, AD neuropathology is prominent in clinically diagnosed FTD (Irwin, 2013). Post-mortem analysis of a group of FTLD patients found 16.7% had primarily AD neuropathology (Forman, 2013). In two retrospective studies, 7.1% and 19.7% of bvFTD, and 44.1% of PNFA patients had primary AD neuropathology (Alladi, 2007; Balasa, 2015). This reflects a discrepancy between clinical and neuropathological diagnosis, indicating that AD may be initially misdiagnosed as FTD. Prospective evaluation of patients diagnosed with FTD, found 12.7% were diagnosed with AD after two years (Mendez, 2007). Similarly, post-mortem evaluation of bvFTD patients found the majority of false-positive cases had AD (Harris, 2013).

There are also overlapping clinical features between DLB and AD, including executive, visuospatial, episodic memory, language and social cognition deficits (Bousiges, 2019). Cognitive evaluation of DLB and AD shows patients have similar immediate total recall scores, and impairments to naming/repetition, comprehension, verbal fluency, and visuospatial function (Noe, 2004; Kyung, 2011). Further, similar Mini-Mental State Examination (MMSE) and Clinical Dementia Rating scores have been observed (Noe, 2004; Walker, 2012). Several studies find overlapping neuropathology between AD and DLB (Karantzoulis, 2011; Ballard, 2006). One study found 50% of prospectively evaluated DLB patients had overlapping LB and AD at post-mortem (Lopez, 2002). This contributes towards misdiagnosis, which is indicated in two studies that analysed brain tissue. In one, of 19 cases with mixed AD/DLB neuropathology, 8 had been clinically diagnosed with DLB, and 8 with AD before death (Walker, 2015). In the other, of 88 cases with prominent LB neuropathology, only 33 were correctly diagnosed with DLB before death. 54 had been diagnosed with AD (Weisman, 2007). More recently, a meta-analysis of DLB diagnostic criteria found 20% of diagnoses were incorrect, with AD the most frequent misdiagnosis (Rizzo, 2018).

While AD can be misdiagnosed as FTD or DLB, there is also evidence of FTD and DLB being misdiagnosed as AD. One study found that AD clinical diagnosis based on McKhann 1984 criteria was wrong in ~20% of cases, with mismatched cases commonly exhibiting LBD, FTD, CBD, or PSP neuropathology (Beach, 2012). Misdiagnosis of AD can have serious implications, for example, a longitudinal study found that 18.18% of patients misdiagnosed with AD were given inappropriate medication (Gaugler, 2013). Whilst anticholinesterases are often prescribed for AD, these fail to benefit FTD patients, and can even harm them (DeLozier, 2016). A 6-month follow-up study found that a third of donepezil-treated FTD patients experienced increased disinhibition and compulsions (Mendez, 2009). Similarly, DLB patients have increased sensitivity to anticholinergics and antipsychotics, possibly increasing morbidity and mortality (Rizzo, 2018; Gaugler, 2013). Accurate diagnosis is essential in ensuring correct treatment regimens are administered.

Biomarkers are objectively measurable indicators of normal or pathological biological processes. In dementia, they can be broadly divided into imaging and biological fluid biomarkers (Ahmed, 2014). Neuroimaging techniques have beneficial clinical utility, providing means to monitor age-related and pathophysiological mechanisms causing structural, connectivity, and functional decline (Varghese, 2013). This review focuses on biomarkers found biological fluid (bio-fluid), which can be valuable for the purposes of diagnosis, subtype classification, and monitoring prognosis or therapeutic responses.
(Lonneborg, 2008). The cerebrospinal fluid (CSF) is a well-established source of biomarkers, with the value of having direct contact with the brain and spinal cord, hence providing a representation of biochemical and metabolic changes. However, as the lumbar puncture required to retrieve CSF coincides with mild side effects, identifying biomarkers in blood, urine, or saliva, would offer minimally invasive and cheaper alternatives (Zetterberg, 2019; Sharma, 2016). Based on consensus criteria for AD molecular and biochemical markers, a biomarker should ideally be precise, non-invasive, inexpensive, reproducible, with sensitivity/specificity above 80% (Growdon, 1998).

The identification of AD biomarkers in CSF that were incorporated in the AD diagnostic criteria (McKhann, 2011) illustrates the successful translation of pathophysiological understanding to the clinical setting. The core AD biomarkers are decreased Aβ1–42 and increased total-tau (t-tau) and phosphorylated-tau181 (p-tau181). Decreased Aβ42 indicates increased plaque load, increased t-tau reflects neuroaxonal degeneration, and increased p-tau reflects NFT pathology (Zetterberg, 2017). Details of amyloid and tau as the core fluid biomarkers of AD are outlined in review articles (Lee, 2019; Khoury, 2019; Zetterberg, 2020).

Whilst the core AD biomarkers were incorporated for clinical use in McKhann 2011 diagnostic criteria, they were subsequently incorporated into a framework for observational and interventional research purposes, proposed by the National Institute on Aging and Alzheimer’s Association (NIA-AA). The A/T/N classification system encompasses the correlation between neuroimaging findings and fluid biomarkers of AD to reflect pathological changes related to Aβ, tau, and neuronal injury, helping to classify varying clinical presentations, including preclinical and prodromal AD (Jack, 2018). Hence, CSF biomarkers for AD are making significant contributions both clinically and in research, but progress has not yet developed to this point for FTD and DLB.

The overlapping pathology between FTD/DLB and AD coincides with pathological levels of AD CSF biomarkers in these patients, creating misinterpretations. CSF tau in DLB and FTD were at intermediate levels between control and AD subjects in one study (Van Harten, 2011). Another found 13% and 30% of FTD and DLB patients had pathological Aβ42:p-tau ratios, respectively (Skillback, 2015). The AD profile was also seen in 30% and 47% of FTD and DLB CSF samples, respectively, challenging the differential diagnosis (Schoonenboom, 2012). Hence, despite fluid biomarkers being incorporated into AD diagnostic criteria, those for other dementias are less developed (Ahmed, 2014).

As FTD and DLB share overlapping symptoms, pathology and CSF profiles with AD, improving discriminatory power between the syndromes with novel fluid biomarkers would be clinically valuable (Neimantsverdriet, 2017). This could enhance diagnostic accuracy without neuropathological validation, guaranteeing disease-specific interventions are applied during life. Fluid markers may enhance our pathophysiological understanding, prognostic monitoring and clinical trial recruitment process (Ahmed, 2014; Lilford, 2018, Bayer, 2018).

The purpose of this systematic review is to compare novel biomarker levels between FTD and AD, and DLB and AD patient groups, and to evaluate the diagnostic accuracy of these biomarkers in differentiating the patients. For the purpose of this review, “novel” biomarkers are defined as those not yet utilised clinically. Hence, the core CSF biomarkers of AD that have already been incorporated into diagnostic criteria (Aβ1–42, t-tau, and p-tau) will not be included. Outcome measure 1: To compare novel fluid biomarkers levels between AD and FTD patients, and AD and DLB patients. Outcome measure 2: To consider the diagnostic value of novel fluid biomarkers to differentiate AD and FTD, and AD and DLB.
Methodology

This systematic review was performed in conjunction with Preferred Reporting Items for Systematic Review and Meta-analysis (PRISMA) guidelines (Moher, 2009). The Population, Intervention, Comparators, Outcome, and Study design (PICOS) framework was used to form the review question.

- **Participants**: clinically diagnosed FTD (specifically bvFTD and PPA) or DLB patients.
- **Intervention**: quantification of biomarker levels in fluid sample (CSF, urine, saliva, or blood).
- **Comparison**: individuals with clinically diagnosed AD.
- **Outcome measure**: the difference of biomarker levels between AD and FTD, or AD and DLB patients, and the diagnostic value of these biomarkers to differentiate them.
- **Study design**: original, peer-reviewed observational studies.

Articles were included if clinical groups comprised ≥8 subjects and if patients were diagnosed using validated consensus criteria. For FTD, this included Neary 1998 criteria, Rascovsky, 2011 bvFTD and Gorno-Tempini, 2011 PPA criteria (Neary, 1998; Gorno-Tempini, 2011; Rascovsky, 2011). For DLB, this included 1996, 2005, and 2017 McKhann criteria (McKhann, 1996, 2005, 2017). For AD, this included 1984 or 2011 McKhann criteria (McKhann, 1984, 2011). Old and revised criteria were included as many institutions use original criteria. Post-mortem studies, reviews, studies without specified ethical approval, case reports, editorials, commentaries, studies on one dementia type, studies only measuring core AD biomarkers (Aβ1–42, t-tau, and p-tau), and studies comparing AD to a collection of “other dementias” were excluded.

FTD compared to AD and DLB compared to AD searches were conducted separately on PubMed and Scopus. The full search syntax is presented in supplementary (2). Searches were restricted to 2009–2019 and to the English language. Retrieved results were screened based on rapid review of titles and abstracts. Potentially eligible studies were then read in full to assess against the eligibility criteria. Additional sources including Open Grey, DARE, Trip Medical, and reviews from widely cited authors were also searched for primary literature.

Data of interest including patient demographics, study design, technique to measure biomarkers, biomarker levels, and the diagnostic accuracy for the novel biomarker (either individually or in combination with other biomarkers) were extracted into tables. Bias and applicability of the included studies were assessed by Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2). The “reference standard” was the clinical diagnostic criteria, and the “index test” was the novel biomarker (Harrison, 2016). Answers to QUADAS-2 signalling questions were recorded, and scores were given based on the number of domains in which there was risk of bias or applicability concern.

Meta-analysis was performed using Review Manager (RevMan 5). Forest-plots were created using the random-effects model for biomarkers measured in ≥3 studies to find the standardised mean difference (SMD) and 95% confidence interval (CI) between AD vs. FTD and AD vs. DLB. Statistical heterogeneity was quantified using $I^2$: 25%, 50%, and 75% indicates low, medium and high heterogeneity, respectively (Higgins, 2003). For studies where median and interquartile range (IQR) were given, the method by Wan, et al was used to estimate the mean and standard deviation (SD) (Wan, 2014). Biomarker diagnostic accuracy was assessed using receiver operator curve (ROC) analysis. Area under curve values (AUC), sensitivity/specificity were provided for AD vs FTD or AD vs DLB. Details on diagnostic accuracy are shown in supplementary (3). The relationship of AUC with diagnostic accuracy are: 0.5–0.6=fail; 0.6–0.7=poor; 0.7–0.8=fair; 0.8–0.9=good; 0.9–1.0=excellent (Xia, 2013).
3. Results

The search protocol yielded 614 results from PubMed and Scopus (June, 2019). After screening against the inclusion and exclusion criteria, 27 studies were eligible. Reasons for exclusions are presented in supplementary (4). A PRISMA flow diagram illustrating the selection process is shown in figure (2).

The proportion of studies with high/low/unclear risk of bias or applicability concern in each QUADAS-2 domain are displayed in figure (3). The patient selection domain was affected with a high risk of selection bias as many studies failed to implement consecutive or random sampling. Only 5 studies used consecutive sampling (Paterson, 2018; Kapaki, 2013; Chiasserini, 2017; Aerts, 2011; Bostrom, 2009). The index test domain was affected with a high risk of detection bias as many studies failed to implement blinding to clinical diagnosis when analysing the biomarkers. 8 studies used blinding to diagnosis (Steinacker, 2018; Hampel, 2018; Baldacci, 2017; Chiasserini, 2017; Mulugeta, 2011; Nutu, 2011; Herbert, 2014; Van Steenoven, 2018). There was mostly a low risk of bias affecting the reference standard domain; biomarker levels could not influence the diagnosis in 14 studies, as the recruitment required diagnosis prior to biomarker analysis (Steinacker, 2018; Alcolea, 2017; Goetzl, 2016; Oeckl, 2019; Schneider, 2018; Podlesniy, 2013; Kapaki, 2013; Kasuga 2010; Chiasserini, 2017; Mulugeta 2011; Herbert 2014; Aerts 2011; Wennstrom, 2015; Bostrom 2009). Blinding to biomarker levels when performing clinical diagnosis was carried out in 6 studies (Paterson, 2018; Fernecczy, 2011; Bibl, 2012; Boban, 2010; Bibl, 2010; Van Steenoven, 2018). There was a low risk of bias in the flow and timing domain, except for a few cases in which biomarker levels were not available for all patients in the study.
Applicability concerns relate to the practical applicability of each domain. For the reference standard, this was influenced by the use of older diagnostic criteria rather than revised ones. For the index test, applicability was influenced by the use of in-house assays rather than commercial ones. The full results of the QUADAS-2 assessment are presented in supplementary (5).

13 of the 27 studies compared fluid biomarkers in AD to FTD only, 12 compared AD to DLB only, and two studies: Paterson (2018) and Struyfs (2015) included all three groups. All studies were conducted in Europe and North America, except two studies: one in Japan and one in Argentina. None of the 27 studies analysed saliva or urine, 4 analysed blood, and the remaining studies analysed CSF. In total, 1242 patients with AD were compared to 399 patients with DLB and 424 patients with FTD. Details of the characteristics of the included studies can be found in figure (4).

In summary, the fluid biomarkers quantified in the 27 studies can be organised into six main groups: (1) β-amyloid peptides; (2) soluble amyloid precursor protein α/β (sAPPα/β); (3) neurofilament light chain (NfL); (4) α-synuclein; (5) markers of neuroinflammation and gliosis; (6) Other miscellaneous biomarkers, including synaptic proteins, circulating mitochondrial DNA (mtDNA), MHPG, fatty acid binding protein B (FABP3). Findings from each group are discussed below.
Figure 4: Characteristics of the 27 included studies

(A) Diagnostic criteria used for AD, FTD, DLB
14 studies included DLB groups. 15 studies included FTD groups. Of these, Neary (1998) criteria was used in 8. Of these, 6 included behavioural FTD and 2 included behavioural and language phenotypes. Newer 2011 FTD criteria was used in 7 studies. Of these, 4 included only bvFTD and 1 included bvFTD and PNFA. Newer criteria were also used in 2 studies to include FTLD phenotypes, including bvFTD, PNFA, CBS, PSP, and MND.

(B) Study design

(C) Patient demographics

1 One study (Paterson, 2018) included a semantic dementia sample; this was not included due to the sample size of n=7.

2 One study (Struyfs, 2015) was not accounted for as the authors did not specify genders of all groups.

3 One study (Schneider, 2018) was not accounted for as the authors did not specify ages.

(D) Geographical distribution of the studies
3.1. β-amyloid biomarkers

Novel CSF β-amyloid peptides were compared in AD versus FTD patients in 4 studies: Gabelle (2011) and Bibl (2012), Paterson (2018) and Struyfs (2015), and in AD versus DLB patients in 5 studies: Bibl (2010), Nutu (2013) and Mulugeta (2011), Paterson (2018) and Struyfs (2015). *Meso Discovery Scale (MSD)* electrochemiluminescence multi-array was used to quantify CSF β-amyloid in all studies, except Bibl (2010) and Bibl (2012), which used Aβ-SDS-PAGE immunoblotting. Full results for β-amyloid biomarkers can be found in supplementary (6): tables 1 and 2.

The studies showed that CSF Aβ peptides: Aβ38, Aβ40 and Aβ37, were lower in FTD and DLB than AD (and controls) but derived no more than “fair” discriminatory potential. Amyloid ratios provided improved diagnostic accuracy.

3.1.1. Aβ42/Aβ40

4 studies measured Aβ42/Aβ40 in FTD and AD. Bibl (2012), Struyfs (2015), and Paterson (2018) found Aβ42/Aβ40 was decreased in AD compared to FTD (and controls), achieving “fair” to “good” discriminatory power. AUC, sensitivity, and specificity values ranged from 0.797-0.86, 75.5-91%, and 65-85%, respectively. Contrastingly, Gabelle (2011) found Aβ42/Aβ40 was higher in AD compared to FTD, achieving “good” discriminatory power (AUC=0.85), with sensitivity and specificity of 79% and 76%, respectively. The results from these studies were meta-analysed and found Aβ42/Aβ40 was higher in FTD than AD, with a SMD of 2.91 (-0.74 – 6.56) 95% CI. However, this was not statistically significant (p=0.12), and there was significant heterogeneity (I²=99%). (Figure 5A)

3 studies measured Aβ42/Aβ40 in DLB and AD. Paterson (2018), Struyfs (2015), and Nutu (2013) found levels were lower in AD compared to DLB (and controls), achieving “fair” discriminatory power. AUC, sensitivity and specificity values ranged from 0.73-0.759, 89.9-90%, 47-58.8%, respectively. The results from these studies were meta-analysed and found that Aβ42/Aβ40 levels were significantly higher in DLB. The SMD was 3.44 (2.7 – 4.17) 95% CI (p<0.00001), but there was significant heterogeneity (I²=72%). (Figure 5B)

3.1.2. Aβ42/Aβ38

3 studies measured Aβ42/Aβ38 in FTD and AD. Struyfs (2015) and Bibl (2012) found Aβ42/Aβ38 was lower in AD compared to FTD (and controls), achieving “good” to “excellent” diagnostic accuracy. AUC, sensitivity and specificity values ranged from 0.815-0.917, 81.6-82%, and 68.8-82%, respectively. Contrastingly, in Gabelle (2011), Aβ42/Aβ38 was higher in AD compared to FTD, resulting in “good” discriminatory power (AUC=0.87), with sensitivity and specificity of 88% and 86%, respectively. The results from these studies were meta-analysed and found a slight increase in FTD compared to AD, with a SMD 0.39 (-2.30 – 3.08) 95% CI. However, this was not statistically significant (p=0.78), and there was significant heterogeneity (I²=99%). (Figure 5C)

2 studies measured Aβ42/Aβ38 in DLB and AD. Mulugeta (2011) and Struyfs (2015) found levels were lower in AD compared to DLB (and controls), achieving “fair” to “good” diagnostic accuracy. AUC, sensitivity and specificity values ranged from 0.765-0.843, 78-95.9%, and 67-70.6%, respectively.

3.1.3. Aβ42/37

Struyfs (2015) measured Aβ42/37 in FTD, DLB and AD. It was found to be lower in AD and provided better discriminatory power than Aβ37 alone, achieving “good” AUC values. AUC, sensitivity and specificity values to distinguish AD and FTD were 0.851, 69.4%, and 94.1%, respectively. AUC, sensitivity, and specificity values to distinguish AD and DLB were 0.832, 89.8%, 70.6%, respectively.

3.1.4. Aβ38/40

2 studies (Struyfs, 2015; Gabelle, 2011) found that Aβ38/40 levels were not significantly different between AD and FTD and achieved “fail” AUC values. It showed better discriminatory power between
AD and DLB; Struyfs (2015) found Aβ38/40 was decreased in DLB compared to AD (and controls). AUC, sensitivity and specificity values were 0.826, 61.2%, 94.1%, respectively.

3.1.5. Other amyloid biomarkers
To distinguish AD and FTD, Bibl (2012) found that amino-terminally truncated Aβ2-42, Aβ2-42/Aβ1-38, and Aβ2-42/Aβ1-40 were decreased in AD compared to FTD (and controls), offering “good” to “excellent” discriminatory potential, with sensitivity/specificity above 80%. Aβ2-42/Aβ1-38 was particularly useful, with 100% sensitivity. To distinguish AD and DLB, Bibl (2010) found the percentage of oxidised Aβ40 (Aβ40ox) was elevated in DLB compared to AD (and controls). This gave “excellent” discriminatory power, with sensitivity/specificity over 80%.

Figure 5: Forest plots for CSF levels of novel amyloid biomarkers measured in ≥3 studies
A – Meta-analysis of Aβ42/Aβ40 levels in AD & FTD groups from 4 studies.
B – Meta-analysis of Aβ42/Aβ40 in AD & DLB groups from 3 studies.
C – Meta-analysis of Aβ42/Aβ38 levels in AD & FTD groups from 3 studies.
3.2. Soluble amyloid precursor protein (sAPPα/β)

Three studies compared CSF sAPPβ levels between AD and FTD groups (supplementary 7). Alcolea (2017) and Perneczky (2011) used IBL ELISA; Gabelle (2011) used MSD electrochemiluminescence. Meta-analysing findings from these studies showed that CSF sAPPβ was significantly higher in AD compared to FTD, providing a SMD of -1.13 (-1.84 – -0.42) 95% CI (p=0.002), however, there was significant heterogeneity (I²=84%) (Figure 6A). CSF sAPPα was also measured in 1 study; Gabelle (2011) found no significant differences between FTD and AD.

Whilst Alcolea (2017) found sAPPβ gave “good” discriminatory potential (AUC=0.86) to distinguish AD and FTLD, Gabelle (2011) found discriminatory potential of “fail” (AUC=0.67). It is worth noting that Gabelle (2011) did not provide a control group or sAPPα units, challenging the capacity to evaluate and reproduce results from this study. sAPPβ with tau fulfilled criteria of an ideal biomarker in Perneczky (2011), achieving “excellent” discriminatory potential (AUC=0.92), with sensitivity and specificity of 95.2% and 81.2%, respectively.

A) CSF sAPPβ levels in AD & FTD groups

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>FTD Mean</th>
<th>SD</th>
<th>Total</th>
<th>AD Mean</th>
<th>SD</th>
<th>Total</th>
<th>Weight</th>
<th>Std. Mean Difference</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcolea 2017</td>
<td>546.6</td>
<td>243.3</td>
<td>68</td>
<td>1,015.5</td>
<td>346.7</td>
<td>72</td>
<td>35.2%</td>
<td>-1.15 [1.03, -1.17]</td>
<td></td>
</tr>
<tr>
<td>Gabelle 2011</td>
<td>38,895</td>
<td>8,599</td>
<td>34</td>
<td>39,407</td>
<td>8,813</td>
<td>52</td>
<td>34.9%</td>
<td>-0.51 [0.95, -0.07]</td>
<td></td>
</tr>
<tr>
<td>Perneczky 2011</td>
<td>680.032</td>
<td>258.93</td>
<td>18</td>
<td>1,200.29</td>
<td>452.4</td>
<td>21</td>
<td>23.9%</td>
<td>-1.35 [2.06, -0.65]</td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>120</td>
<td>145</td>
<td>100%</td>
<td>115 [-1.84, -0.42]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.33; Chi² = 12.74, df = 2 (P = 0.002); I² = 84%
Test for overall effect: Z = 3.11 (P = 0.002)

B) CSF NfL levels in AD & FTD groups

Median (IQR) values from Hampel (2010) & Paterson (2018) were converted to mean (SD) using Wan (2014) method.

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>FTD Mean</th>
<th>SD</th>
<th>Total</th>
<th>AD Mean</th>
<th>SD</th>
<th>Total</th>
<th>Weight</th>
<th>Std. Mean Difference</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcolea 2017</td>
<td>2,174.4</td>
<td>2,394.9</td>
<td>66</td>
<td>1,051.8</td>
<td>395.4</td>
<td>72</td>
<td>25.4%</td>
<td>0.66 [0.002, 1.00]</td>
<td></td>
</tr>
<tr>
<td>Hampel 2008</td>
<td>1,043</td>
<td>248.3</td>
<td>9</td>
<td>1,497.5</td>
<td>157.8</td>
<td>35</td>
<td>24.7%</td>
<td>-2.50 [-3.41, -1.58]</td>
<td></td>
</tr>
<tr>
<td>Nikkadi 2019</td>
<td>2,009.5</td>
<td>1,544.9</td>
<td>13</td>
<td>1,333</td>
<td>432.6</td>
<td>14</td>
<td>24.9%</td>
<td>1.62 [0.21, 1.83]</td>
<td></td>
</tr>
<tr>
<td>Paterson 2018</td>
<td>2,226.8</td>
<td>518.7</td>
<td>17</td>
<td>1,206.1</td>
<td>157.0</td>
<td>150</td>
<td>25.0%</td>
<td>5.14 [4.40, 5.88]</td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>107</td>
<td>277</td>
<td>100%</td>
<td>110 [-1.45, 3.63]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 6.00; Chi² = 179.65, df = 3 (P < 0.00001); I² = 96%
Test for overall effect: Z = 0.64 (P = 0.40)

C) CSF α-synuclein levels in AD & DLB groups

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>DLB Mean</th>
<th>SD</th>
<th>Total</th>
<th>AD Mean</th>
<th>SD</th>
<th>Total</th>
<th>Weight</th>
<th>Std. Mean Difference</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chiassone 2017</td>
<td>1,751.1</td>
<td>1,105.3</td>
<td>40</td>
<td>2,450.8</td>
<td>871.24</td>
<td>48</td>
<td>25.9%</td>
<td>-0.70 [-1.14, -0.27]</td>
<td></td>
</tr>
<tr>
<td>Kapiesi 2013</td>
<td>174</td>
<td>103.5</td>
<td>16</td>
<td>92.1</td>
<td>42.4</td>
<td>18</td>
<td>23.1%</td>
<td>1.03 [0.31, 1.76]</td>
<td></td>
</tr>
<tr>
<td>Kasuga 2010</td>
<td>6.2</td>
<td>4.2</td>
<td>34</td>
<td>12.2</td>
<td>5.8</td>
<td>31</td>
<td>28.3%</td>
<td>-0.79 [-1.29, -0.28]</td>
<td></td>
</tr>
<tr>
<td>Van Steenoven 2018</td>
<td>106</td>
<td>34</td>
<td>41</td>
<td>89</td>
<td>30</td>
<td>35</td>
<td>25.7%</td>
<td>0.58 [0.12, 1.04]</td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>131</td>
<td>132</td>
<td>100%</td>
<td>132 [-0.84, 0.85]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.67; Chi² = 32.41, df = 3 (P < 0.00001); I² = 91%
Test for overall effect: Z = 0.02 (P = 0.99)

Figure 6: Forest plots for CSF levels of sAPPβ, NfL, and α-synuclein levels

A – Meta-analysis of sAPPβ levels in AD & FTD groups in 3 studies.
B – Meta-analysis of NfL levels in AD and FTD groups in 4 studies.
C – Meta-analysis of α-synuclein levels in AD & DLB groups in 4 studies.
3.3. Neurofilament light chain (NfL)

NfL was quantified in AD and FTD groups in 5 studies, as shown in supplementary (8). Using Quanterix single molecule array (Simoa), Steinacker (2018) found serum NfL was higher bvFTD patients compared AD, hence this offered “poor” diagnostic accuracy (AUC=0.678). Of the 4 studies that quantified CSF NfL, Niikado (2019), Alcolea (2017), and Paterson (2018) found NfL was higher in FTD than AD. Opposing results arose from Hampel (2018), which was dissimilar to the others, having used older diagnostic criteria and a small FTD sample. CSF NfL from these 4 studies (specifically bvFTD in Alcolea (2017)), were meta-analysed (Figure 6B). This found an overall increase of NfL in FTD compared to AD, however, not significantly. The SMD was 1.09 (-1.45 – 3.63) 95% CI (p=0.40). There was also significant heterogeneity (I² = 98%).

Regarding the diagnostic accuracy of CSF NfL, Niikado (2019) found “fair” accuracy (AUC=0.736) to differentiate sporadic bvFTD and AD. In Paterson (2018), ROC analysis was given only for PNFA vs AD, resulting in a “good” value (AUC=0.84), but specificity was 50%. Alcolea (2017) recruited FTLD (including CBS and PSP) and AD; the large sample has increased statistical power, however, overinterpretations are possible as NfL was not measured in all subjects. NfL provided “poor” accuracy (AUC=0.67); whilst for NfL/sAPPβ, it was “good” (AUC=0.85). In Hampel (2018), NfL with Aβ1-42 and p-tau gave “fair” diagnostic accuracy (AUC=0.796).

3.4. α-synuclein

Five studies quantified α-synuclein in AD and DLB groups, as shown in supplementary (9), and found conflicting results. Two studies found CSF α-synuclein was higher in DLB compared to AD. In Kapaki (2013), total α-synuclein levels provided “fair” discriminatory power (AUC=0.73), with sensitivity and specificity of 50% and 94.4%, respectively. Van Steenoven (2018) quantified the early soluble oligomer aggregates of α-synuclein, α-α-synuclein, with investigators blinded to diagnoses. In combination with tau, this provided “good” discriminatory power (AUC=0.84), sensitivity and specificity of 81% and 74%, respectively. In contrast, two studies for CSF and 1 for serum found α-synuclein was higher in AD compared to DLB. For total CSF α-synuclein levels, Chiasserini (2017) found this provided “fair” discriminatory power (AUC=0.78), and Kasuga (2010) found this provided sensitivity and specificity of 72.4% and 61.8%, respectively. Serum α-synuclein was quantified in Laske (2011) using Invitrogen ELISA and found DLB had significantly lower α-synuclein than AD, which offered “fair” discriminatory power (AUC=0.723), with sensitivity and specificity of 70% and 65%.

The contradictory findings are reflected by the meta-analysis of CSF α-synuclein levels, including total and α-α-synuclein. This found no difference between the AD and DLB groups, giving a SMD of 0.01 (-0.84 – 0.85) 95% CI, with significant heterogeneity (I²=91%) (Figure 6C).

3.5. Markers of neuroinflammation & gliosis

3.5.1. Chitinase-3-like protein 1 (YKL-40)

CSF YKL-40 was quantified in AD and FTD groups in 3 studies, as shown in supplementary (10). 2 studies, Hampel (2018) and Baldacci (2017) used R&D Systems ELISA and found YKL-40 was higher in AD compared to FTD. These studies were conducted by the same group, and although not confirmed to be the same patients, it is worth noting the methods, demographics, and results are the same. Hence, a meta-analysis was not conducted for the YKL-40 studies. Baldacci (2017) found “fair” (AUC=0.71) discriminatory power of YKL-40, whilst Hampel (2018) found “good” (AUC=0.813) discriminatory power when YKL-40 was combined with Aβ1-42 and p-tau. Using MicroVue ELISA, Alcolea (2017) found YKL-40 was higher in AD than bvFTD but lower than PPA. YKL-40 had “fair” discriminatory power but was improved to “good” (AUC=0.84) when combined with sAPPβ. CSF YKL-40 was also higher in AD than DLB in 1 study: Wennstrom (2015) found it provided “fair” (AUC=0.736) discriminatory power between the groups (supplementary 13).

3.5.2. Glial fibrillary acidic protein (GFAP)
One study quantified serum GFAP in AD and FTD: Oeckl (2019) used Quanterix Simoa and found GFAP was increased in AD compared to bvFTD, providing “good” discriminatory power (AUC=0.85), with sensitivity and specificity of 89% and 79%, respectively (supplementary 12).

3.6. Other novel biomarkers

3.6.1. FTD compared to AD

Synaptic proteins
Two studies quantified synaptic protein levels in AD and FTD groups, as shown in supplementary (11). Hampel (2018) used in-house ELISA to quantify CSF neurogranin, which was higher in AD than FTD. Goetzl (2016) quantified plasma synaptic proteins in bvFTD and AD using American Research Products or Biomatik ELISA. In contrast to Hampel (2018), AD had significantly lower synaptotagmin, synaptopodin, synaptophysin, neurogranin and GAP-43 than bvFTD. To distinguish AD and FTD, CSF neurogranin with Aβ1-42 and YKL-40 provided “good” (AUC=0.802) diagnostic accuracy. Plasma proteins had prominent capacity, achieving “perfect” (AUC=1) accuracy, and sensitivity/specificity of 1. “Good” discriminatory power was found for synaptotagmin (AUC=0.85) and neurogranin (AUC=0.88), and “excellent” for synaptopodin (AUC=0.94).

miR-632
One study quantified miR-632: Schneider (2018) used quantitative real-time PCR (qRT-PCR) to find the expression of 72 miRNAs in CSF of AD and FTD (bvFTD, bvFTD/ALS, PPA, PPA/ALS) (supplementary 12). MiR-632 was significantly decreased in FTD compared to AD and provided “good” discriminatory power (AUC=0.88).

Circulating mtDNA
One study, Podlesniy (2013), quantified CSF mtDNA in AD and FTLD through qRT-PCR and found circulating cell-free mitochondrial DNA (mtDNA) was decreased in AD compared to FTLD (supplementary 12). MtDNA distinguished the groups with “excellent” (AUC=0.98) diagnostic accuracy, with sensitivity and specificity of 92% and 87%, respectively.

3.6.2. DLB compared to AD

MHPG
Two studies conducted by the same research group but with different groups of patients that quantified CSF 3-methoxy-4-hydroxyphenylethleneglycol (MHPG) using liquid chromatography: Aerts (2011) and Herbert (2014) found MHPG was decreased in DLB compared to AD (supplementary 13). Aerts (2011) found MHPG discriminatory power was “good” (AUC=0.81), with sensitivity and specificity of 74.4% and 78.3%, respectively. Combining MHPG with AD core biomarkers gave “excellent” discriminatory power (AUC=0.99), with sensitivity and specificity of 97.6% and 95%. In Herbert (2014), combining MHPG with AD core biomarkers, provided “good” discriminatory power (AUC=0.85), sensitivity and specificity of 64.6% and 100%, respectively.

FABP3
One study quantified CSF fatty acid binding protein 3 (FABP3): Chiasserini (2017) utilised Hycult Biotech ELISA and found FABP3 was increased in AD compared to DLB (supplementary 13). Although FABP3 found “fail” (AUC=0.54) discriminatory potential, this improved to “excellent” (AUC=0.92) when combined with other biomarkers, including p-tau and α-synuclein.

Ca, Cu, Mg
One study quantified CSF levels of Cu, Ca, and Mg using mass spectrometry: Bostrom (2009) found Ca and Mg were significantly higher in DLB than AD (supplementary 13). Cu was not significantly higher. The discriminatory potential for Mg was “excellent” (AUC=0.92), sensitivity and specificity of 93% and 81%, respectively, whilst Ca achieved “good” potential (AUC=0.84), with sensitivity and specificity of 93% and 63%, respectively.
4. Discussion

4.1. Contextualizing the findings

4.1.1. Novel indicators of amyloid pathology

Amyloid precursor protein (APP) processing occurs through two pathways and generates soluble APP (sAPPα/β), as shown in figure (7). SAPPβ is a product of APP amyloidogenic processing, marking a critical step towards Aβ generation (Perneczky 2011). Increased CSF sAPPβ have previously correlated to Aβ peptides in AD (Lewczuk, 2010; Gabelle, 2010). As shown in the meta-analysis, CSF sAPPβ was significantly higher in AD compared to FTD, and intriguingly, a consequent study by the Perneczky group found sAPPβ was also higher in AD than FTD in plasma samples (Perneczky, 2013). To this end, sAPPβ is worth investigating as a prospective biomarker to distinguish AD and FTD, particularly when combined with other markers.

Aβ42, the major component of amyloid plaques, only accounts for ~10% of the total Aβ concentration (Otto, 2008). Considering this, and the overlapping CSF Aβ42 levels between AD and non-AD dementias, there is interest to discover other isoforms to improve the distinction (Bibl, 2012). In this review, amyloid ratios were typically superior for distinguishing DLB and FTD from AD compared to raw peptides, in particular Aβ42/Aβ38 and Aβ42/Aβ40. This concurs with literature where they provide improved differentiation between AD and non-AD dementias, detection of amyloid pathology and were more interpretable than AD core markers (Janelidze, 2016; Spies, 2010; Welge, 2009). Aβ42/Aβ40 may also overcome the confounding effect of Aβ42 during preanalytical processing, implying it offers more reproducible measurements (Willemse, 2017). Although ROC analysis showed Aβ42/Aβ38 had slightly better discriminatory power than Aβ42/Aβ40, literature is vaster on the benefits of Aβ42/Aβ40 over Aβ42 (Baldeiras, 2018; Lewczuk, 2015; Dumurgier, 2015; Hansson, 2019). Additional efforts to explore both Aβ42/Aβ38 and Aβ42/Aβ40 are desirable.

Single studies also highlighted promise for other amyloid biomarkers, ensuing sensitivity/specificity above 80%. The decreased levels of amino-truncated Aβ2-42 in AD compared to FTD shown in Bibl (2012) interestingly corresponds to evidence of elevated Aβ2-42 in AD brains post-mortem, suggesting Aβ2-42 may be significant in AD pathology (Wiltfang, 2001). Similarly, the elevated Aβ40 in DLB compared to AD shown in Bibl (2010) has also been observed in pathologically confirmed patients (Bibl, 2011). Although the pathophysiological significance of these biomarkers remains uncertain, Aβ40 could possibly be a DLB-specific marker.

4.1.2. NfL as an FTD biomarker

NfL is the lightest of 3 major polypeptides composing neurofilaments, abundant in axoplasm of large myelinated neurons, and pivotal for cytoskeletal function (Zetterberg, 2016). Upon axonal damage, NfL releases into interstitial fluid, communicating with CSF and consequently with blood (Gaetani, 2019). NfL is a marker of axonal degeneration and is increased in CSF/blood of patients with vascular dementia, PD, amyotrophic lateral sclerosis, Huntington’s disease, and multiple sclerosis (Zhao, 2019; Backstrom, 2015; Lu, 2015; Constantinescu, 2009; Varhag, 2019).

Whilst the meta-analysis found that CSF NfL was higher in FTD groups compared to AD, the difference was not significant. It is worth considering that none of these studies utilised the Simoa technique, which has 126-fold-higher sensitivity than ELISA; using this may have retrieved different outcomes (Kuhle, 2016). Further, the diagnostic accuracy of NfL found in these studies offered conflicting results, however, the accuracy improved notably when combined with other biomarkers.

More convincing evidence of NfL as an FTD marker arises from another meta-analysis, confirming CSF NfL was higher in FTD compared to other dementias (Bridel, 2019). Cultivating literature also suggests NfL is a credible FTD prognostic marker, correlating to disease severity and short survival (Skillback, 2014; Skillback, 2017; Rohrer, 2016). The differences in NfL fluid levels between AD and FTD, and the association with disease progression, may be explained by pathology. For example, in
transgenic mice models, CSF/serum NfL was higher in the tauopathy models compared to AD (Bacioglu, 2016). Thus, NfL may not only be a prospective diagnostic marker for FTD, but also a prognostic marker. This would be clinically influential, but future investigations implementing longitudinal Simoa measurements are needed to determine the true potential.

4.1.3. α-synuclein as a DLB marker

Fluid α-synuclein are considered markers of LB pathology, and this review identified conflicting results for whether it is higher in AD or in DLB. Increased CSF α-synuclein levels in DLB were hypothesised by Kapaki (2013) to signify pathological release into the extracellular space following intracellular aggregation and neuronal damage, whilst Van Steenoven (2013) postulated it is due to failed clearance pathways. Meanwhile, the decreased CSF α-synuclein levels in DLB compared to AD was hypothesised by Kasuga (2010) to correspond to increased α-synuclein in the brain. This particular concept is strengthened by post-mortem findings whereby increased α-synuclein in AD compared to DLB was also observed in pathologically confirmed patients (Mollenhauer, 2011).
Although the pathological significance is unverified, the contradictory findings may be explained by the fact most studies used in-house ELISA to quantify α-synuclein, each with varied protocols and antibodies, creating heterogeneous results and lack of reproducibility (Mollenhauer, 2010). To reliably compare α-synuclein between AD and DBL, quantification platforms must be standardised across centres. Further, confounders may also be influential, for example, α-synuclein may differ between genders (Wennstrom, 2012). Hence, factors such as gender differences and disease duration are important to consider in future studies.

4.1.4. Neuroinflammation and gliosis

Neuroinflammation contributes to AD pathophysiology, and involves activated microglia and astrocytes (Heneka, 2015). YKL-40 glycoprotein is a neuroinflammation marker, abundant in reactive glia, with elevated levels often seen in MS and traumatic brain injury. (Llorens, 2017; Bonnhe-Barkay, 2010). Based on the findings presented, increased YKL-40 appeared to be more specific towards AD compared to FTD and DBL. This coincides with post-mortem evidence where YKL-40 was elevated in AD but not DBL and increased in DBL brains specifically with AD pathology (Llorens, 2017; Lleo, 2019). Thus, it was proposed that YKL-40 is a preclinical AD marker, representing pathophysiology (Antonell, 2014; Olsson, 2013). To this end, YKL-40 has been found correlating with neuroinflammation, gliosis, Aβ, axonal degeneration, and cognitive decline (Bos, 2019).

YKL-40 appeared to result in contradictory outcomes for diagnostic accuracy analyses. For example, YKL-40 diagnostic accuracy ranged from fail–fair to distinguish AD and FTD. However, this could be explained by the dissimilarities between the patients; atypical Parkinsonism was included in the Alcolea (2017) FTLD group, whilst the other studies included behavioural FTD only. Some post-mortem evidence has also established that YKL-40 correlates with tau, and levels differ between FTLD-TDP and FTLD-tau groups (Alcolea, 2019; Del Campo, 2018). Hence, the heterogeneity in FTD syndromes and pathology perhaps influenced the results. Despite the inadequate accuracy shown by YKL-40 to distinguish AD from FTD, it had improved value when combined with sAPPβ or p-tau and Aβ1-42.

GFAP was also found to be increased in AD compared to bvFTD. GFAP is an astrocytic cytoskeletal filament protein, a marker of astrocytosis and is released during neurodegeneration (Oeckl, 2019). The single pilot study by Oeckl, et al (2019) proposes preliminary evidence for a blood biomarker with sufficient ability to distinguish AD and FTD that is worth exploring further in a larger cohort. The study also coincides with prior findings that CSF GFAP is elevated in AD (Fukuyama, 2001). This may be attributed to histopathological evidence of reactive astrocytes surrounding amyloid plaques, suggesting GFAP expression correlates to AD plaque load (Kamphuis, 2014).

4.1.5. Other promising biomarkers to differentiate FTD and AD

Synaptic degeneration is central to neurodegenerative diseases (Lleo, 2019). Early evidence conveys synaptic protein loss in AD brains, including presynaptic (synaptotagmin and synaptophysin), membrane (GAP-43 and synaptobrevin) and postsynaptic (neurogranin and synaptotodin) proteins (Reddy, 2005). This concurs with elevated CSF synaptic proteins, which are considered markers of synaptic degeneration (Thorsell, 2010; Portelius, 2015; Ohrfelt, 2016).

In this review, one study found that CSF synaptic proteins were higher in AD than FTD, whilst another study found plasma synaptic proteins were higher in FTD than AD. The discrepancy between plasma and CSF findings in fact coincides with a study where increased CSF, but unchanged plasma neurogranin levels were found in AD. This seems to convey the complex relationship of analytes between CSF and blood, or that plasma changes perhaps occur at a slower rate (De Vos, 2015). Differences in synaptic protein levels between AD and FTD may reflect pathological differences leading to synaptic degeneration. For example, synaptic loss has also been observed histopathologically in FTLD, but specifically in the frontal regions, with no significant changes found in parietal areas, when compared to AD (Brun, 1995; Lui, 1996). The differences in the extent of synaptic degeneration between FTLD and AD possibly elucidates distinctive synaptic protein levels. MiRNA dysregulation may contribute to neurodegenerative pathogenesis by promoting toxic protein accumulation or altering expression of proteins that inhibit cell survival (Eacker, 2009). The release of
miRNA into body fluid is considered a marker of dysregulated cellular communication (Denk, 2018). This review identified a single study in which miR-632 was decreased in FTD compared to AD and achieved “good” diagnost ic accuracy. Although the pathophysiological connection is yet unknown, aberrant RNA processing in FTLD is a topic of growing research (Piscopo, 2016).

Mitochondrial DNA (mtDNA) in the extracellular space may signify mtDNA turnover in the brain. Podlesny, et al (2013) found decreased mtDNA in AD compared to FTD, which perhaps relates to altered bioenergetics and mitochondrial dysfunction in AD pathophysiology (Moreira, 2010; Lagouge, 2013). However, Podlesny (2013) has limited generalisability due to the small sample size, and the conclusion is challenged by a consequent study that found mtDNA was higher in AD than controls (Cervera-Carles, 2017). Despite this, Podlesny (2013) encouraged further work into circulating mtDNA, with altered levels also observed in MS and PD (Varhaug, 2017; Pyle, 2015). MtDNA as a biomarker is in the premature stages of research, but is advantageous over proteins, being more resistant to endonuclease degradation, and PCR amplification techniques to detect them are more accurate than immunoassays (Podlesny, 2018). Hence, whether mtDNA depletion in AD can differentiate from FTD is worth investigating further.

4.1.6. Other promising biomarkers to differentiate DLB and AD

MHPG is the primary metabolite of norepinephrine/ noradrenaline (NE/NA) and can be used as a CSF index of NE metabolism (Chase, 1973). MHPG was decreased in DLB compared to AD in two studies. This coincides with post-mortem evidence where decreased MHPG was observed in 8/11 brain regions in DLB patients compared to AD (Vermeiren, 2015). This relates to low NE in the neocortex and putamen in DLB compared to AD, and noradrenergic neuron degradation (Szot, 2006; Ohara, 1998). LB pathology severely affects the locus coeruleus, the main NE producing nucleus (Del Tredici, 2013). Thus, the NE neuron loss due to LB explains reduced NE, and hence, reduced MHPG in DLB. Considering this, and the fact CSF MHPG in combination with AD biomarkers showed good diagnostic accuracy, MHPG may be useful for differentiating AD and DLB. Further, an earlier study discovered salivary MHPG correlated to CSF MHPG (Reuster, 2002). Hence, MHPG may be a potentially valuable non-invasive marker to study the NA system and distinguish DLB from AD.

FABP3 is a cytosolic protein that regulates lipid composition/fluidity of the brain membrane and may influence synapse formation and neuronal activity (Sepe, 2018). This review found FABP3 was higher in AD compared to DLB, which may be linked to neurodegeneration, as FABP3 has also been found to correlate with entorhinal cortex atrophy and tau (Desikan, 2013; Bjerke, 2016). The diagnostic accuracy of FABP3 was improved substantially in combination with other biomarkers, which overlaps with an earlier study where serum FABP3 with CSF tau was helpful in differentiating AD and DLB (Mollenhauer, 2007). Hence, CSF FABP3 in AD may represent neurodegeneration and can differentiate from AD when used alongside other biomarkers.

Metal dyshomeostasis is important in dementia, implicated in cellular metabolism, antioxidation, and inflammation (Huut, 2019). Metal imbalances have been identified in PD and AD (Bocca, 2006; McAllum, 2016). In this review, one study highlighted identified Mg and Ca levels in DLB compared to AD, suggesting raised metal levels are more dominant in DLB. It was postulated that Ca represents disrupted blood brain barrier integrity and Mg represents the cellular damage by α-synuclein in DLB. This coincides with in-vitra findings whereby Mg and Ca accelerate α-synuclein formation (Lowe, 2009; Nielsen, 2001). However, it is ambiguous whether metal dyshomeostasis is the cause or consequence of LB pathology.

4.2. Limitations and future directions

Having been conducted by an unblinded author, this review ensues the risk of reporting bias and human error, and language bias is sourced by excluding non-English written studies. Due to the small number of included studies, meta-analyses were not conducted for all biomarkers, generating inconsistent outcome measures, perhaps challenging the interpretations. As definitive diagnosis generally relies on
post-mortem confirmation, it is possible some of the clinically diagnosed patients included in this review were misdiagnosed. Having excluded post-mortem studies also comes with the limitation of not being aware of the extent of underlying neuropathology. Mixed neuropathologies are a common cause of dementia, hence it is possible that neuropathological heterogeneity between patients has a confounding effect, whereby the varying and mixed levels of AD, FTD, and DLB pathology influenced the levels of fluid biomarkers (Boyle, 2018).

Whilst this review focusses specifically on biomarkers found in biological fluid, it can be recognised as a limitation that the association with neuroimaging findings was not explored. The importance of this is reflected in the A/T/N research framework proposed by the NIA-AA, which incorporates both imaging and fluid biomarkers to define AD by pathological stages (Jack, 2018). Few studies identified in this review investigated both imaging and fluid biomarkers, for example, Niikado et al (2019) found that NfL correlated with left orbitofrontal cortex thickness in some FTD patients, whilst Steinacker et al (2018) found that NfL correlated with atrophy in frontal and subcortical areas in FTD. As imaging modalities are paramount to providing reflections of disease in living subjects, a worthwhile future investigation is that of multi-modal biomarkers including neuroimaging and fluid biomarkers to help strengthen the differentiation of FTD and DLB from AD.

There are limitations of the included studies. Most of the studies included in this review are cross-sectional, from which we cannot infer causality (Levin, 2006). There was notable risk of selection and detection bias, creating potentially subjective sampling and outcome reporting. Applicability concerns relates to the use of older diagnostic criteria; the consistent use of revised criteria across centres would increase the likelihood of correctly classifying the dementias and the generalisability between studies. Confounders may reduce the internal validity of the studies that failed to account for influence of dementia duration/severity, comorbidities and medications on biomarker levels. Further, there is evidence that cerebrovascular risk factors and diseases may directly contribute to neurodegeneration (Wolters, 2019). As such, varying extents of underlying vascular pathology between patients may have confounding effect on the biomarker levels found in the included studies. Heterogeneous results may also be explained by the variable methods between the studies, such as pre-analytical/analytical differences. This issue is exacerbated by the fact some studies did not fully report details of methodological processes, reducing the reproducibility.

Governing the true utility of fluid biomarkers involves combating some issues. At the study level, investigators can reduce bias by consecutive recruitment, blinding, and controlling confounders. Following consensus guidelines for fluid sampling e.g. Teunissen, et al (2009) would enhance comparability between studies. Beyond control of the studies, is the lack of standardised methods to detect novel biomarkers, and other than Uman Diagnostics NfL ELISA, they are mainly research grade. Validated methods would be a step towards systematising procedures. Ideally, recruiting large AD and DLB/FTD cohorts diagnosed using newer clinical criteria, and quantifying biomarkers longitudinally with homogenous techniques across centres could enhance our knowledge of novel biomarkers.

4.3. Conclusions

This review depicted where research is positioned in the pursuit of biomarkers to distinguish FTD and DLB from AD. The novel biomarkers identified in this review are illustrated in figure (8). The following was determined from the literature:

1. Research focussed on CSF and to a lesser extent, blood. Biomarkers in non-invasively obtained fluid: urine and saliva, were not found, suggesting research to this end perhaps achieves discouraging results, seldom reaching publication.

2. Combining multiple biomarkers, each signifying distinctive pathophysiological processes, consistently outperformed in ROC analysis compared to individual biomarkers, demonstrating better capacity to distinguish the dementias.
3. Increased NfL may be an FTD-specific biomarker. It is hypothesised that $\uparrow$NfL, $\uparrow$Aβ42/Aβ38 & Aβ42/Aβ40, and $\downarrow$sAPPβ CSF levels in FTD compared to AD may be helpful in distinguishing them.

4. Results for the value of α-synuclein for DLB were contradictory; a commercial assay may help to accurately quantify levels. MHPG had convincing evidence as a DLB biomarker of noradrenergic degeneration. It is hypothesised that $\downarrow$MHPG and $\uparrow$Aβ42/Aβ38 & Aβ42/Aβ40 CSF levels in DLB compared to AD may be helpful in distinguishing them.

To conclude, ample progress has been made in the pursuit of novel biomarkers to increase the certainty of FTD/DLB classification from AD. Many have been highlighted here, which may be markedly beneficial when used in combination with other biomarkers, such as other novel biomarkers, or core AD biomarkers including p-tau, t-tau or Aβ42. Clinically, a cheap, non-invasive fluid biomarker could guarantee accurate patient diagnosis, reducing the economic cost of invasive investigations, and ensure application of dementia-specific management. They can be employed in research for precise clinical trial recruitment, and monitoring prognosis and treatment response. Before we fully harness their worth, overarching concerns must be resolved, including methodological issues, and the need to develop validated quantification procedures. The literature presented paves avenues for subsequent probing to validate the significance of several promising biomarkers.

**Figure 8: Illustration of neuron with surrounding novel biomarkers**

CSF α-synuclein and p-tau are markers of intracellular Lewy body pathology and neurofibrillary tangles (NFTs), respectively. CSF mtDNA is believed to be indicative of mitochondrial dysfunction. CSF miR-632 represents dysregulated miRNA processing. CSF FAPB3 is a cytosolic protein and possibly a marker of neurodegeneration. CSF YKL-40 is a marker of neuroinflammation, abundantly expressed by reactive microglia (green) and astrocytes (purple). Serum GFAP is a marker of astrocytosis (activated astrocytes) and neuronal injury. Levels of CSF sAPPβ and Aβ are indicative of amyloidogenic APP processing leading to extracellular amyloid pathology (red). CSF Ca, Cu, Mg represent metal dyshomeostasis, which can have intracellular or extracellular effects, contributing to processes such as oxidative stress or inflammation. The synapse is represented in the circle. Noradrenaline (NA) is released from NA presynaptic nerve endings and is metabolised in an extra-neuronal cell (green), forming the MHPG metabolite. CSF MHPG is thought to represent NA neuronal degeneration. Synaptic degeneration is detected through CSF/plasma synaptic proteins.

Information on NA metabolism & source of MHPG from: Eisenhofer, 1998

Representation of neuron adapted from: Molinuevo, 2018; De Deyn, 2015; Blennow, 2012; Zetterberg, 2016; Zetterberg, 2019

*Diagram created on BioRender.*
Acknowledgements
We would like to thank Henrik Zetterberg for reviewing and advising us on the write-up of this manuscript.

Funding declarations
During the period in which this study was undertaken, Dr. Mie Rizig received funding from the UCL Grand Challenges Small Grant scheme and the Michael J Fox Foundation. Professor Henry Houlden received funding from the Medical Research Council and Michael J Fox Foundation.

References


Levin K. (2006) Study design III: Cross-sectional studies. Evidence Based Dentistry. 7:24-25. DOI: https://doi.org/10.1038/sj.ebd.6400375


**Included studies**


