CSF and blood biomarkers for Alzheimer's disease: a systematic review and meta-analysis

Bob Olsson PhD[a], Ronald Lautner MD[a], Ulf Andreasson PhD[a], Annika Öhrfelt PhD[a], Erik Portelius PhD[a], Maria Bjerke PhD[a,b], Mikko Hölttä PhD[a], Christoffer Rosén MD[a], Caroline Olsson PhD[c], Gabrielle Strobel[d], Elizabeth Wu[d], Kelly Dakin PhD[d], Max Petzold PhD[e,f], Kaj Blennow MD[a] and Henrik Zetterberg MD[a,g].

[a]Department of Psychiatry and Neurochemistry, the Sahlgrenska Academy at the University of Gothenburg, Mölndal, Sweden. [b]Department of Biomedical Sciences, University of Antwerp, Belgium. [c]Department of Radiation Physics, the Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden. [d]Alzforum, Cambridge, USA. [e]Unit for Health Metrics, Department of Medicine, Institute of Medicine, the Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden. [f]School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. [g]Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, United Kingdom.

Abstract count: 294, Word count: 2874, Figures/Tables: 8, References: 31, Supplemental information: 1

Running title: Fluid biomarkers for Alzheimer's disease

Corresponding author: Bob Olsson, PhD, Department of Psychiatry and Neurochemistry, University of Gothenburg, V-house, Sahlgrenska University Hospital Mölndal, 431 80

Mölndal, Sweden, Phone, +46 31 343 24 06, Fax + 46 31 343 24 26, E-mail: bob.olsson@neuro.gu.se.

Summary

Background Alzheimer's disease (AD) biomarkers are important for early diagnosis in clinical routine and trials. Three core AD cerebrospinal fluid (CSF) biomarkers (Aβ42, T-tau, and P-tau) have been evaluated in numerous studies and there are several emerging AD markers. However, there is no comprehensive meta-analysis of their diagnostic performance.

Methods We screened PubMed and Web of Science for articles on CSF and blood biomarkers reflecting neurodegeneration (T-tau, NFL, NSE, VLP-1, and HFABP), amyloid precursor protein (APP) metabolism (Aβ42, Aβ40, Aβ38, sAPPα, and sAPPβ), tangle pathology (P-tau), blood-brain-barrier function (albumin ratio) and glial activation (YKL-40, MCP-1, and GFAP). Based on inclusion criteria, 231 of 13,319 screened articles were included. Biomarker performance was rated on fold change between AD and controls and between mild cognitive impairment (MCI) who later converted to AD (MCI-AD) and stable MCI who had at least a follow-up time of two years.

Findings Core biomarkers differentiated AD from controls with excellent performance; CSF T-tau (2·54, CI 2·44-2·64, p<0·0001, 152 studies), P-tau (1·88, CI 1·79-1·97, p<0·0001, 91 studies), and Aβ42 (0·56, CI 0·55-0·58, p<0·0001, 131 studies). Differentiation between MCI-AD and stable MCI was also strong (0·66-1·81). Furthermore, CSF NFL (2·35, 95% CI 1·90-2·91, p<0·0001) and plasma T-tau (1·95, 95% CI 1·12-3·38, p=0·02) had large effect sizes, while CSF NSE, VLP-1, HFABP, and YKL-40 were more moderate (1·28-1·47). Remaining biomarkers had marginal effect sizes or were non-significant.

Interpretation Core CSF AD biomarkers and NFL, as well as plasma T-tau, are strongly associated with AD. Emerging biomarkers CSF NSE, VLP-1, HFABP, and YKL-40 are moderately associated with AD, while plasma Aβ42 and Aβ40 are not.

Funding Swedish Research Council, Swedish State Support for Clinical Research, Alzheimer's Association, the Knut and Alice Wallenberg Foundation, the Torsten Söderberg Foundation, the Alzheimer Foundation (Sweden), and the Biomedical Research Forum, LLC.

Introduction

Alzheimer's disease (AD) is the most common neurodegenerative disease. Its neuropathological hallmarks are amyloid β (Aβ)-containing plaques, and tau-containing neurofibrillary tangles.[1] It is generally acknowledged that several drug candidates, most targeting Aβ, have failed in large, multi-centre clinical trials in part because a large proportion of clinically diagnosed patients show no evidence of amyloid pathology on PET scans, and thus do not have AD.[2] This diminishes the possibility of identifying a clinical benefit of the tested drug and points to the need for validated biomarkers in drug development and clinical practice.

Over the past 25 years, three core AD cerebrospinal fluid (CSF) biomarkers have been identified and tested in hundreds of studies. These are the 42 amino acid form of Aβ (Aβ42) with low levels due to cortical amyloid deposition, total tau (T-tau) with high levels due to cortical neuronal loss,[3-6] and phosphorylated tau (P-tau) with high levels reflecting cortical tangle formation.[7,8] High diagnostic accuracy of these biomarkers has been demonstrated for AD, with sensitivity and specificity reaching 85-90%, and also for prodromal AD in the mild cognitive impairment stage of the disease (MCI-AD).[9] Biomarkers have also been incorporated in modern diagnostic criteria. [10] However, individual studies vary greatly and no comprehensive meta-analysis has evaluated their diagnostic performance. Nor has emerging biomarkers reflecting neurodegeneration, glial cell activation, and amyloid precursor protein (APP) metabolism, which show promise as diagnostic tools,[11] yet been meta-analyzed.

In the present paper, we have examined the literature on 15 biomarkers in both CSF and blood, covering core AD biomarkers as well as other markers of neurodegeneration, glial and blood-brain-barrier (BBB) function, and APP metabolism, to determine which ones separate AD from controls and MCI-AD from stable MCI.

Methods

Search strategy and selection criteria

We did this study according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[12] We screened articles published in PubMed and Web of Science between 1984 (when the first diagnostic criteria were published)[13] and June 30th, 2014, which reported data on biomarkers of neurodegeneration (T-tau,[3-6] neurofilament light protein (NFL),[14] neuron-specific enolase (NSE),[15] visinin-like protein 1 (VLP-1),[16] and heart fatty acid binding protein (HFABP)[17]), APP metabolism (Aβ42, Aβ40, Aβ38, α and β cleaved soluble amyloid precursor protein (sAPPα, and sAPPβ)), tangle pathology (P-tau),[7,8] BBB function (CSF/serum albumin ratio) and glial activation (YKL-40,[28,29] monocyte chemotactic protein 1 (MCP-1),[18] and glial fibrillary acidic protein (GFAP)[19]) in CSF and blood (serum or plasma) in AD vs. controls or MCI-AD vs. stable MCI (Table 1). It is important to note that only a few of these markers have been validated against neuropathology and that the precise process their CSF levels reflect is tentative until such studies have been performed.[20] For detailed search terms, see Supplemental Methods. The review protocol can be found at http://www.alzforum.org/alzbiomarker/about-alzbiomarker/methods. Additional records were identified through reference lists of included articles (Figure 1). Data was independently extracted by ten authors. Plasma and serum analyses for each biomarker were meta-analyzed together. Data were curated from cross-sectional cohort studies as well as from baseline measurements in longitudinal studies with clinical follow-up. All entered data was checked by

two researchers. Stable MCI was defined as having a follow-up of at least 2 years without meeting dementia diagnosis. MCI-AD was defined as baseline MCI with progression to AD dementia at follow-up. Controls included cognitively normal volunteers (spouses or volunteers recruited by other means from the population) or individuals who were sampled when admitted to hospital, *e.g.*, for a surgical procedure that required spinal anaesthesia, and who proved cognitively normal when tested.

Exclusion criteria for entire studies or individual cohorts were: 1) Articles not written in English. 2) Studies not containing AD and a control cohort, or MCI-AD and a stable MCI cohort. 3) Cohorts with fewer than ten individuals. 4) Data reported in a format other than mean±SD or mean±SEM (study authors were contacted and asked to supply this information). 5) Biomarker data from urine, saliva, or cellular blood fractions. 6) Studies using non-quantitative methods, *e.g.*, explorative proteomics or western blot. 7) Cohorts lacking diagnostic criteria for AD or MCI, or cohorts representing a mix of diagnoses. 8) Cohorts of stable MCI with less than two years follow-up. 9) Cohorts of individuals younger than 18 years of age. 10) Studies lacking appropriately referenced methods, with the exception of CSF/serum albumin ratio which we accepted as a well-established routine analysis. 11) Previously published data. This also includes large initiatives such as ADNI where the same biomarkers are listed in many articles but only measured once and therefore only the first article where a biomarker is presented was included in this study. Finally, for longitudinal cohorts with several publications on the same baseline data, we included the publication that had the longest follow-up. 12) Control cohorts containing subjects with inflammatory, neurological, or psychiatric diagnoses that may affect CSF biomarker levels.

Här måste vi beskriva QUADAS (inklusive ta med QUADAS-referenserna som finns i SBU-dokumentet), ungefär så som du skriver i referee-svaret. Resultatet av QUADAS (hur många som hade high och medium quality) skall stå i Results första paragraf.

Statistical analysis

There is a large variation in how different laboratories establish cut-points. Some use historical in house data, some establish optimal cut-off points within the same cohort or from an external cohort and some use reference limits established on healthy subjects; these are just some of the approaches used. Furthermore, there is significant variability in biomarker levels between laboratories and assays.[21] The application of fold change reduces these problems. Hence, we performed meta-analyses using the AD/control ratios and the MCI-AD/stable MCI ratios; each specific ratio was generated within a single study. In studies with more than one control cohort, only the most cognitively normal (*i.e.*, healthy controls over hospital controls) and age appropriate cohort was used. In studies with more than one AD cohort, all AD cohorts were included and divided by the control group to generate multiple ratios per study. In cohorts where a biomarker was analyzed with more than one assay we included only one of them and chose a commercial assay over an in-house assay. Standard error of the ratio of the mean values was calculated using the delta method.[22,23] Random effects meta-analysis using the method of DerSimonian and Laird with the estimate of heterogeneity taken from the inverse-variance fixed-effect model was applied in Stata 13.1 (metan command sbe24_3).[24] P-values ≤0·05 were considered significant. Publication bias was assessed using funnel plots. The study protocol is listed on http://www.alzforum.org/alzbiomarker/about-alzbiomarker/methods.

Role of the funding source

Funders of the study had no role in study design, data collection, analysis, interpretation, or writing of the report with the exception of the Biomedical Research Forum LLC, employees of which were involved in the development of the AlzBiomarker database, revised the

manuscript, and are co-authors here. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

The initial screening identified 13,319 articles, 675 full-text articles were assessed for eligibility and 231 articles were included (Figure 1).

Core biomarkers

We found most eligible studies for T-tau. A total of 165 AD and 154 control cohorts in 152 studies had data on CSF T-tau, comprising 11,341 AD patients and 7,086 controls. All studies had an AD/control ratio above one, with an average of 2·54 (95% CI 2·44-2·64, p<0·0001; Fig 2A). For CSF P-tau, we combined data from studies using single or multiple epitopes for detection. In 91 studies, a total of 7,498 AD patients from 98 cohorts and 5,126 controls from 91 cohorts were included. As was the case for T-tau, all AD/control P-tau ratios were above one, with an average of 1·88 (95% CI 1·79-1·97, p<0·0001; Fig 2B). In the meta-analysis of Aβ42, we included studies using assays covering the 1-42 as well as the x-42 epitopes. A total of 143 AD cohorts and 135 control cohorts in 131 studies with 9,949 AD patients and 6,841 controls were included. The CSF Aβ42 ratio between AD and controls was below one in all comparisons but one, with an average of 0·56 (95% CI 0·55-0·58, p<0·0001; Fig 2C). The funnel plots suggested publication bias for all three core biomarkers.

Markers of neurodegeneration

Data on NFL in CSF were available from nine AD and eight control cohorts comprising 245 AD patients and 292 controls (average ratio 2·35, 95% CI 1·90-2·91, p<0·0001; Fig 3A), and on NSE from seven cohorts including 258 AD patients and 160 controls (average ratio 1·47, 95% CI 1·08-2·00, p=0·01; Fig 3B). VLP-1 was analyzed in four AD (n=252) and control (n=486) cohorts (average ratio 1·46, 95% CI 1·31-1·62, p<0·0001; Fig 3C) and HFABP in

five AD (n=285) and control (n=297) cohorts (average ratio 1·39, 95% CI 1·24-1·57, p<0·0001; Fig 3D). The funnel plots suggested no publication bias for markers of neurodegeneration.

Glial markers

The microglial and astrocyte marker YKL-40 was analyzed in six AD and five control cohorts with 298 and 330 subjects, respectively. It was significantly elevated with an average ratio of 1·28 (95% CI 1·23-1·35, p<0·0001; Fig 3E). In contrast, the astrocyte marker GFAP did not differ between AD and controls (average ratio 1·12, 95% CI 0·58-2·15, p=0·74, from two AD and control cohorts of 59 and 39 subjects respectively; Fig 3F). The microglial marker MCP-1 had a small but significant effect size in three cohorts of AD and controls with 59 subjects in each group (average ratio 1·12 95% CI 1·06-1·18, p<0·0001; Fig 3G). The funnel plots suggested no publication bias for glial markers.

Marker of BBB

We analyzed CSF/serum albumin ratio in AD (20 cohorts, n=854) vs. controls (16 cohorts, n=441). The ratio was significantly elevated but the effect size was small (average ratio 1·10, 95% CI 1·01-1·20, p=0·04; Fig 3H). The funnel plots suggested no publication bias for the marker of BBB.

Markers of APP metabolism

APP cleavage generates a host of quantifiable products besides Aβ42. Aβ40 was the second-most studied amyloid marker. It was analyzed in 25 AD and 24 control cohorts with 1,079 and 784 subjects, respectively. The average effect size was significant but minor 0·94 (95% CI 0·90-0·99, p=0·02; Supplementary fig 1A). Aβ38 was analyzed in eight studies with 251 AD patients and 195 controls (average ratio 0·99, 95% CI 0·88-1·12, p=0·89, Supplementary fig 1B). Both sAPPα and sAPPβ are cleavage products from APP. Neither sAPPα (average

ratio 1·03, 95% CI 0·93-1·14, p=0·55), from nine AD (n=572) and control (n=415) cohorts (Supplementary fig 1C), nor sAPPβ (average ratio 1·02, 95% CI 0·95-1·09, p=0·61), from ten cohorts of 631 AD and 439 controls (Supplementary fig 1D) differed between AD patients and controls. The funnel plots suggested no publication bias for markers of APP metabolism.

Plasma and serum biomarkers

In contrast to the significantly lower CSF levels of Aβ42 in AD compared with controls, the average ratio of plasma Aβ42 between AD and controls was non-significant and close to one in the analysis of 22 AD and 20 control cohorts comprising 1,872 AD and 3,855 controls (average ratio 1·04, 95% CI 0·96-1·12, p=0·32; Supplementary fig 2A). Besides Aβ42, the plasma and serum literature had sufficient data for meta-analysis only for Aβ40, T-tau, YKL-40, NSE, MCP-1, and HFABP. Aβ40 was not significantly different between AD and controls (average ratio 1·04, 95% CI 0·98-1·11, p=0·17) in 21 AD (n=1,661) and 19 control cohorts (n=3,668; Supplementary fig 2B). Data from six AD and five control cohorts on T-tau, comprising 271 AD and 394 controls, showed a large effect size (average ratio 1·95, 95% CI 1·12-3·38, p=0·02; Supplementary fig 2C). There were no differences between AD and controls for NSE (average ratio 1·00, 95% CI 0·86-1·17, p=0·99) in three AD (n=102) and control cohorts (n=97; Supplementary fig 2D) or for HFABP (average ratio 1·05, 95% CI 0·83-1·33, p=0·69) in two AD (n=55) and control (n=74) cohorts (Supplementary fig 2E). For YKL-40, the effect size was large but just not significant in three AD and two control cohorts with 155 AD and 233 controls, respectively (average ratio 1·95, 95% CI 0·99-3·84, p=0·053; Supplementary fig 2F). However, there was no difference between AD and controls for MCP-1 (average ratio 1·00, 95% CI 0·89-1·13, p=0·99) in six AD (n=540) and control (n=344) cohorts (Supplementary fig 2G). The funnel plots suggested no publication bias for plasma and serum biomarkers.

Biomarkers in MCI

All 12 eligible comparisons of CSF T-tau between MCI-AD (n=307) and stable MCI (n=570) had a ratio above one, with an average of 1·76 (95% CI 1·64-1·89, p<0·0001; Supplementary fig 3A). Likewise, all nine comparisons of CSF P-tau between MCI-AD (n=251) and stable MCI (n=501) had a ratio above one, with an average of 1·72 (95% CI 1·46-2·02, p<0·0001; Supplementary fig 3B). CSF Aβ42 concentrations were also significantly lower in MCI-AD (12 cohorts, n=352) compared with stable MCI (12 cohorts, n=610), albeit with a smaller effect size than between AD and controls (average ratio 0·67, 95% CI 0·63-0·73, p<0·0001; Supplementary fig 3C). In contrast, CSF Aβ40 was not significant between three cohorts each with 152 MCI-AD and 189 stable MCI patients (average ratio 0·98, 95% CI 0·90-1·07, p=0·71; Supplementary fig 3D). Neither was the average ratio of plasma Aβ42 from three MCI-AD (n=308) and stable MCI (n=379) cohorts (average ratio 0·81, 95% CI 0·53-1·24, p=0·32; Supplementary fig 3E). However, the average ratio of plasma Aβ40 was significantly different between three MCI-AD (n=308) and stable MCI (n=379) cohorts, but the effect size was negligible (average ratio 1·07, 95% CI 1·03-1·10, p=0·0002; Supplementary fig 3F). Neither sAPPα nor sAPPβ was significantly changed in CSF between MCI-AD and stable MCI; sAPPα from three MCI-AD (n=118) and stable MCI (n=169) cohorts (average ratio 1·09 95% CI 0·96-1·25, p=0·20; Supplementary fig 3G) and sAPPβ from three MCI-AD (n=118) and stable MCI (n=169) cohorts (average ratio 1·06, 95% CI 0·87-1·28, p=0·59; Supplementary fig 3H). The funnel plots suggested no publication bias for biomarkers in MCI.

Biomarker performance

Head-to-head biomarker performance is shown in Figure 4A-B.

Discussion

Our study provides by far the most comprehensive meta-analysis of the rapidly growing biomarker literature in AD. It shows unequivocally that AD is associated with lower CSF levels of Aβ42 and higher CSF levels of T-tau and P-tau compared with controls. Furthermore, AD is associated with increased CSF levels of NFL, NSE, VLP-1, HFABP, YKL-40, and increased plasma levels of T-tau.

Fluid biomarker measurement in AD is marked by variability between laboratories and batches of commercial assays, which are research-grade but not clinically certified.[21] This precluded a traditional meta-analysis based on existing cut-off levels. Furthermore, our analysis included a multitude of in-house assays, some of which used different epitopes in the same molecule, particularly for Aβ42 and P-tau. Some Aβ42 assays use antibodies targeting the first and last amino acids, *i.e.* Aβ1-42, or a mid-domain epitope in combination with an end-specific antibody, designated Aβx-42. Tau is phosphorylated at multiple sites, but most assays only detect the phosphorylation of one specific amino acid. To circumvent this methodological variability, we used ratios of AD/control and MCI-AD/stable MCI biomarker levels for meta-analysis and combined the analyses of relevant forms of a specific protein into one. Neither for Aβ42 nor for P-tau did the literature indicate superiority of a particular epitope. In our meta-analysis, different assays using different antibody combinations against Aβ42 generated similar results (Supplement). For P-tau, the other epitopes besides P-tau181 were analyzed in too few studies to evaluate differences in performance.

The picture was unanimous for T-tau and P-tau; all studies had an AD/control ratio above one. The results were also remarkably consistent for Aβ42 in CSF, with 139 studies finding an AD/control ratio below one, with just one outlier, Jensen et al,[25] finding a ratio above one. A plausible explanation for this anomaly is that the Aβ42 levels of the controls in their study were exceptionally low, only 40% of the levels in a simultaneously analyzed cohort with

depression. Normally, CSF Aβ42 levels are similar between controls and patients with depression.[26]

The CSF signature of elevated T-tau and P-tau and reduced Aβ42 was also observed in MCI-AD compared with stable MCI, although the effect was somewhat less pronounced. Most likely, some individuals categorized as stable MCI were in fact developing prodromal AD thereby diverging in their biomarker profile from healthy controls. This interpretation is supported by results showing that the drop in CSF Aβ42 levels precedes AD dementia by at least ten years.[27-29] To minimize this problem, we excluded all mixed MCI cohorts and required that stable MCI cohorts had a follow-up period of at least two years with cognitive stability. Even so, this problem cannot be totally avoided since the only way to estimate early non-clinical signs of progression, except for very long follow-up or post-mortem examinations, is through the use of CSF or imaging markers of AD pathology, which would be a form of catch 22 if taken into account. It is also worth to note that clinical AD does not guarantee AD pathology and that normal cognition does not guarantee absence of AD pathology. However, this is an issue that we cannot solve in this study since we have to rely upon the clinical diagnostic criteria that were listed in the studies since only very few have autopsy confirmed diagnoses.

Among the new or less studied CSF biomarkers, NFL showed a large effect size (2·35). NFL is the light protein of neurofilament and, with either the medium or the heavy counterpart, makes up neurofilament bundles that determine axonal caliber and conduction velocity.[14] Thus, our data indicate that axonal destruction is prominent in AD. Furthermore, NSE, VLP-1, HFABP, and YKL-40 had a moderately large (1·28-1·47) and significant effect size between AD and controls. NSE is a neuron-enriched enzyme of the glycolytic pathway,[15] VLP-1 is a calcium-sensor protein found in the neuronal cytoplasm,[16] HFABP is an intracellular fatty acid transport protein expressed in skeletal muscle heart and neurons,[17] and

YKL-40 is a marker of activated microglia and astrocytes.[30,31] None of these markers reflect the core pathology of AD; nevertheless, this meta-analysis suggests that they could be useful in clinical trials of anti-AD drugs as tau- and Aβ-independent measures of neurodegeneration and glial activation. In contrast, although CSF Aβ40 and MCP-1 and CSF/serum albumin ratio were statistically different between AD and controls, their small effect size precludes their use as diagnostic markers. Moreover, our meta-analysis suggests that sAPPα, sAPPβ, and Aβ38 are not useful in AD diagnostics.

Acceptance of lumbar puncture by physicians varies from country to country, therefore plasma is considered a desirable matrix for AD biomarker analyses. When all studies were weighed in, we found no significant differences between Aβ markers in AD and controls, supporting the hypothesis that plasma Aβ levels reflect peripheral Aβ generation more than AD brain pathology. In contrast, plasma levels of T-tau were significantly associated with AD. Variation in the few available studies was large, therefore more data are needed to verify this association.

Despite conducting exhaustive PubMed and the Web of Science searches, we might have missed some eligible studies. Some studies reported data in a format unsuitable for our analysis; for these studies we requested the missing data from the corresponding authors and most responded. All studies for T-tau and P-tau had a ratio above one and all but one a ratio below one for Aβ42 indicating that the results for these biomarkers are solid. However, there was publication bias for all three core biomarkers between AD and controls and therefore the results will have to be interpreted in the light of this. Furthermore, this indicates that the heterogeneity is small for the core biomarkers. In contrast, some heterogeneity can be observed in the forest plots for the other biomarkers in spite of the use of the fold change approach. Potential reasons for heterogeneity in the results of the less established biomarkers

include smaller studies and less established and validated assays compared with the core biomarkers, as well as less knowledge on potential confounding pre-analytical factors.

To extract maximal knowledge from the abundance of fluid biomarker studies published in recent years, data must first be extracted, aggregated, and organized. Therefore, in conjunction with this manuscript, we have developed an open-access database containing the curated data and meta-analyses shown here. The AlzBiomarker database, which is hosted at www.alzforum.org, contains additional curated data, such as mean age of subjects, MMSE scores, and duration of disease. In addition, interactive visuals allow users to compare results, explore promising biomarkers, methods and assays, as well as to identify knowledge gaps. The database will be updated with new data and meta-analyses as additional studies are published.

In conclusion, CSF Aβ42, T-tau, P-tau, and NFL are biomarkers that, at least on a group level, robustly separate AD patients from controls, while CSF NSE, VLP-1, HFABP, and YKL-40 show more moderate differences. Importantly for earlier detection, CSF Aβ42, T-tau, and P-tau also discriminate between MCI-AD and stable MCI. Plasma T-tau is the only blood biomarker that separates AD from controls; this finding needs verification in larger cohorts.

Contributors

BO designed the study, acquired data, undertook statistical analysis and interpretation, drafted and revised the manuscript. HZ and KB designed the study, acquired data, interpreted the data, and revised the manuscript. MP undertook statistical analysis and revised the manuscript. KD and EW developed the database and revised the manuscript. GS revised the manuscript. RL, UA, AÖ, EP, MB, MH, CR, and CO acquired data and revised the manuscript.

Declaration of interests

KD, EW and GS are employees of Biomedical Research Forum, LLC which owns and operates Alzforum. KB has served on advisory boards for Eli Lilly, Kyowa Kirin Pharma, Pfizer, and Roche. KB and HZ are co-founders of Brain Biomarkers Solutions in Gothenburg AB, a GU Venture-based platform company at the University of Gothenburg. All other authors declare no competing interests.

Acknowledgements

The authors thank all colleagues who have generated the extensive biomarker literature that was meta-analyzed here. The study was funded by the Swedish Research Council, Swedish State Support for Clinical Research, Alzheimer's Association, the Knut and Alice Wallenberg Foundation, the Torsten Söderberg Foundation, the Alzheimer Foundation (Sweden), the European Research Council (grant # 681712) and the Biomedical Research Forum, LLC.

References

1. Blennow K, de Leon MJ, Zetterberg H. Alzheimer's disease. *Lancet* 2006; 368: 387-403.

2. Salloway S, Sperling R, Fox NC et al. Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. *N Engl J Med* 2014; 370: 322-33.

3. de Souza LC, Chupin M, Lamari F et al. CSF tau markers are correlated with hippocampal volume in Alzheimer's disease. *Neurobiol Aging* 2012; 33: 1253-7.

4. Tapiola T, Alafuzoff I, Herukka SK et al. Cerebrospinal fluid {beta}-amyloid 42 and tau proteins as biomarkers of Alzheimer-type pathologic changes in the brain. *Arch Neurol* 2009; 66: 382-9.

5. Hesse C, Rosengren L, Andreasen N et al. Transient increase in total tau but not phospho-tau in human cerebrospinal fluid after acute stroke. *Neurosci Lett* 2001; 297: 187-90.

6. Ost M, Nylen K, Csajbok L et al. Initial CSF total tau correlates with 1-year outcome in patients with traumatic brain injury. *Neurology* 2006; 67: 1600-4.

7. Buerger K, Ewers M, Pirttila T et al. CSF phosphorylated tau protein correlates with neocortical neurofibrillary pathology in Alzheimer's disease. *Brain* 2006; 129: 3035-41.

8. Seppala TT, Nerg O, Koivisto AM et al. CSF biomarkers for Alzheimer disease correlate with cortical brain biopsy findings. *Neurology* 2012; 78: 1568-75.

9. Blennow K, Hampel H, Weiner M, Zetterberg H. Cerebrospinal fluid and plasma biomarkers in Alzheimer disease. *Nat Rev Neurol* 2010; 6: 131-44.

10. Schott JM, Petersen RC. New criteria for Alzheimer's disease: which, when and why? *Brain* 2015; 138: 1134-7.

11. Sutphen CL, Fagan AM, Holtzman DM. Progress update: fluid and imaging biomarkers in Alzheimer's disease. *Biol Psychiatry* 2014; 75: 520-6.

12. Liberati A, Altman DG, Tetzlaff J et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009; 339: b2700.

13. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984; 34: 939-44.

14. Friede RL, Samorajski T. Axon caliber related to neurofilaments and microtubules in sciatic nerve fibers of rats and mice. *Anat Rec* 1970; 167: 379-87.

15. Schmechel D, Marangos PJ, Zis AP, Brightman M, Goodwin FK. Brain endolases as specific markers of neuronal and glial cells. *Science* 1978; 199: 313-5.

16. Laterza OF, Modur VR, Crimmins DL et al. Identification of novel brain biomarkers. *Clin Chem* 2006; 52: 1713-21.

17. Ockner RK, Manning JA, Poppenhausen RB, Ho WK. A binding protein for fatty acids in cytosol of intestinal mucosa, liver, myocardium, and other tissues. *Science* 1972; 177: 56-8.

18. Ishizuka K, Kimura T, Igata-yi R, Katsuragi S, Takamatsu J, Miyakawa T. Identification of monocyte chemoattractant protein-1 in senile plaques and reactive microglia of Alzheimer's disease. *Psychiatry Clin Neurosci* 1997; 51: 135-8.

19. Brenner M, Johnson AB, Boespflug-Tanguy O, Rodriguez D, Goldman JE, Messing A. Mutations in GFAP, encoding glial fibrillary acidic protein, are associated with Alexander disease. *Nat Genet* 2001; 27: 117-20.

20. Olsson B, Zetterberg H, Hampel H, Blennow K. Biomarker-based dissection of neurodegenerative diseases. *Prog Neurobiol* 2011; 95: 520-34.

21. Mattsson N, Andreasson U, Persson S et al. CSF biomarker variability in the Alzheimer's Association quality control program. *Alzheimers Dement* 2013; 9: 251-61.

22. In: Armitage P, Colton T, editors. Encyclopedia of Biostatistics Chichester, United Kingdom: John Wiley & Sons; 1998. p. 3731-7.

23. Friedrich JO, Adhikari NK, Beyene J. The ratio of means method as an alternative to mean differences for analyzing continuous outcome variables in meta-analysis: a simulation study. *BMC Med Res Methodol* 2008; 8: 32.

24. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; 7: 177-88.

25. Jensen M, Schroder J, Blomberg M et al. Cerebrospinal fluid A beta42 is increased early in sporadic Alzheimer's disease and declines with disease progression. *Ann Neurol* 1999; 45: 504-11.

26. Osorio RS, Gumb T, Pomara N. Soluble amyloid-beta levels and late-life depression. *Curr Pharm Des* 2014; 20: 2547-54.

27. Buchhave P, Minthon L, Zetterberg H, Wallin AK, Blennow K, Hansson O. Cerebrospinal fluid levels of beta-amyloid 1-42, but not of tau, are fully changed already 5 to 10 years before the onset of Alzheimer dementia. *Arch Gen Psychiatry* 2012; 69: 98-106.

28. Bateman RJ, Xiong C, Benzinger TL et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. *N Engl J Med* 2012; 367: 795-804.

29. Jansen WJ, Ossenkoppele R, Knol DL et al. Prevalence of cerebral amyloid pathology in persons without dementia: a meta-analysis. *JAMA* 2015; 313: 1924-38.

30. Bonneh-Barkay D, Bissel SJ, Wang G et al. YKL-40, a marker of simian immunodeficiency virus encephalitis, modulates the biological activity of basic fibroblast growth factor. *Am J Pathol* 2008; 173: 130-43.

31. Horbinski C, Wang G, Wiley CA. YKL-40 is directly produced by tumor cells and is inversely linked to EGFR in glioblastomas. *Int J Clin Exp Pathol* 2010; 3: 226-37.

Table 1. Biomarkers analyzed where available in AD vs. CTRL and MCI-AD vs. stable MCI in CSF, serum and plasma.

Neurodegeneration				
T-tau	NFL	NSE	VLP-1	HFABP
APP metabolism				
Aβ42	Aβ40	Aβ38	sAPPα	sAPPβ
Tangle pathology				
P-tau				
Glial activation				
YKL-40	MCP-1	GFAP		
BBB function				
Albumin ratio				

Figure legends

Figure 1: Study selection.

Figure 2: AD/control ratios for CSF core biomarkers.

CSF ratios of T-tau (A), P-tau (B) and Aβ42 (C) between AD and control groups. Individual AD/control ratios and their corresponding 95% CIs are indicated by filled squares. The size of the square indicates the weight of the study. All average ratios and their corresponding 95%

CIs are indicated by clear diamonds. The solid line indicates a ratio of one and the dotted indicates the average ratio.

Figure 3: AD/control ratios for CSF markers of neurodegeneration, glial activation, and blood-brain barrier function.

CSF ratios of NFL (A), NSE (B), VLP-1 (C), HFABP (D), YKL-40 (E), GFAP (F), MCP-1 (G) and CSF/serum albumin ratio between AD and control groups. Individual AD/control ratios and their corresponding 95% CIs are indicated by filled squares. The size of the square indicates the weight of the study. All average ratios and their corresponding 95% CIs are indicated by clear diamonds. The solid line indicates a ratio of one and the dotted indicates the average ratio.

Figure 4: AD/control ratios for CSF markers of APP metabolism.

CSF ratios of Aβ40 (A), Aβ38 (B), sAPPα (C) and sAPPβ (D) between AD and control groups. Individual AD/control ratios and their corresponding 95% CIs are indicated by filled squares. The size of the square indicates the weight of the study. All average ratios and their corresponding 95% CIs are indicated by clear diamonds. The solid line indicates a ratio of one and the dotted indicates the average ratio.

Figure 5: AD/control ratios for plasma biomarkers.

Plasma or serum ratios of Aβ42 (A), Aβ40 (B), T-tau (C), NSE (D), HFABP (E), YKL-40 (F), MCP-1 (G) between AD and control groups. Individual AD/control ratios and their

corresponding 95% CIs are indicated by filled squares. The size of the square indicates the weight of the study. All average ratios and their corresponding 95% CIs are indicated by clear diamonds. The solid line indicates a ratio of one and the dotted indicates the average ratio.

Figure 6: MCI-AD/stable MCI ratios for CSF and plasma markers

CSF ratios of T-tau (A), P-tau (B), Aβ42 (C), Aβ40 (D), sAPPα (G), sAPPβ (H) and plasma ratios of Aβ42 (E) and Aβ40 (F) between MCI-AD and stable MCI groups. Individual MCI-AD/stable MCI ratios and their corresponding 95% CIs are indicated by filled squares. The size of the square indicates the weight of the study. All average ratios and their corresponding 95% CIs are indicated by clear diamonds. The solid line indicates a ratio of one and the dotted indicates the average ratio.

Figure 7: Biomarker performance rating in AD vs. controls.

Head-to-head biomarker performance in CSF (Panel A) and in serum and plasma (Panel B) based on average AD/control ratios. The biomarkers in green circles were significant with good effect size and the ones in blue squares were significant with moderate effect size and the ones in red diamonds were non-significant or significant with minor effect size T-tau 2·54 (95% CI 2·44-2·64, p<0·0001), NFL 2·35 (95% CI 1·90-2·91, p<0·0001), P-tau 1·88 (95% CI 1·79-1·97, p<0·0001), Aβ42 0·56 (95% CI 0·55-0·58, p<0·0001), NSE 1·47 (95% CI 1·08-2·00, p=0·01), VLP-1 1·46 (95% CI 1·31-1·62, p<0·0001), HFABP 1·39 (95% CI 1·24-1·57, p<0·0001), YKL-40 1·28 (95% CI 1·23-1·35, p<0·0001), MCP-1 1·12 (95% CI 1·06-1·18, p<0·0001), GFAP 1·12 (95% CI 0·58-2·15, p=0·74), albumin ratio 1·10 (95% CI 1·01-1·20, p=0·04), Aβ40 0·94 (95% CI 0·90-0·99, p=0·02), Aβ38 0·99 (95% CI 0·88-1·12,

p=0·89), sAPPα 1·03 (95% CI 0·93-1·14, p=0·55), and sAPPβ 1·02 (95% CI 0·95-1·09, p=0·61). Plasma and serum markers are depicted in black circles T-tau 1·95 (95% CI 1·12-3·38, p=0·02), YKL-40 1·95 (95% CI 0·99-3·84, p=0·053), HFABP 1·05 (95% CI 0·83-1·33, p=0·69), Aβ40 1·04 (95% CI 0·98-1·11, p=0·17), Aβ42 1·04 (95% CI 0·96-1·12, p=0·32), MCP-1 1·00 (95% CI 0·89-1·13, p=0·99), and NSE 1·00 (95% CI 0·86-1·17, p=0·99). The AD/control ratios of CSF Aβ42, Aβ40, and Aβ38 were inverted to allow for a clearer comparison with the other biomarkers. The dotted line indicates a ratio of one.