

LETTER TO THE EDITOR

Open Access



CSF proteins of inflammation, proteolysis and lipid transport define preclinical AD and progression to AD dementia in cognitively unimpaired individuals

Marta del Campo^{1,2,3,4*} , Carlos Quesada⁵, Lisa Vermunt^{1,6}, Carel F. W. Peeters⁷, Yanaika S. Hok-A-Hin¹, Calvin Trieu^{1,6}, Anouk den Braber^{6,8}, Inge M. W. Verberk^{1,6}, Pieter J. Visser⁶, Betty M. Tijms⁶, Wiesje M. van der Flier^{6,9} and Charlotte E. Teunissen¹

Abstract

This preclinical AD CSF proteome study identified a panel of 12-CSF markers detecting amyloid positivity and clinical progression to AD with high accuracy; some of these CSF proteins related to immune function, neurotrophic processes, energy metabolism and endolysosomal functioning (e.g., ITGB2, CLEC5A, IGFBP-1, CST3) changed before amyloid positivity is established.

Main text

Pathological features of Alzheimer's disease (AD) start to develop decades before the appearance of clinical signs, providing a unique opportunity to define and intervene

AD biologically before the irreversible brain damage occurs. AD is a multifactorial disorder in which multiple factors and pathways beyond amyloid and tau pathologies are involved (e.g., immunity, lipid metabolism, vascular dysfunction, endocytic pathway) [1]. Cerebrospinal fluid (CSF) reflects the ante-mortem biochemical alterations occurring in the brain and can thus provide the pathological fingerprint of AD *in vivo* [2–4]. CSF proteome is dynamic and protein levels change over the AD disease stages [2, 5]. This is not trivial as a detailed analysis of the *in vivo* CSF proteome of cognitively unimpaired individuals in the AD preclinical stage may unveil proteins and biological pathways especially relevant for the etiology and progression of AD pathophysiology. These could be useful as biomarkers to improve the biological prognosis of AD, potential therapeutic targets and surrogate endpoints for clinical trials conducted in pre-dementia stages targeting different disease mechanisms [6, 7].

We here analyzed >900 CSF proteins reflecting a wide range of mechanisms in the presymptomatic phase of AD (297 cognitively unimpaired individuals (CU): 232 amyloid negative and 65 amyloid positive), with 72%

*Correspondence:

Marta del Campo
mcampo@barcelonabeta.org

¹ Neurochemistry Laboratory and Biobank, Department of Laboratory Medicine, Amsterdam Neuroscience, Amsterdam University Medical Center, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

² BarcelonaBeta Brain Research Center (BBRC), Pasqual Maragall Foundation, Barcelona, Spain

³ Hospital del Mar Research Institute (IMIM), Barcelona, Spain

⁴ Departamento de Ciencias Farmacéuticas y de La Salud, Facultad de Farmacia, Universidad San Pablo-CEU, CEU Universities, Madrid, Spain

⁵ Departamento de Matemática Aplicada a Las TIC, Polytechnical University of Madrid, Madrid, Spain

⁶ Alzheimer Center Amsterdam, Department of Neurology, Amsterdam Neuroscience, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

⁷ Mathematical & Statistical Methods Group (Biometris), Wageningen University & Research, Wageningen, The Netherlands

⁸ Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

⁹ Department of Epidemiology & Data Science, VU University Medical Center, Amsterdam, Netherlands



followed-up clinically [8, 9]. Proximity extension assay (PEA) proteome data from 122 CU participants (103 amyloid positive and 19 amyloid negative; Supplementary Material: Fig. 1) from the EMIF-AD preclinical cohort [10] was used to dismiss proteins in downstream analyses with highly divergent results across CU cohorts and to validate the biomarker panel (see Supplementary Material 1). Demographic characteristics, core AD CSF biomarker concentrations and follow-up information of the cohorts used are described in Supplementary Material 2 Table 1. Figure 1a presents an overview of the study design including the following aims: (i) define the biological changes that characterize the preclinical stage of AD; (ii) identify and validate the panel of markers needed to identify cases who are in the preclinical stage of AD and evaluate if these markers can predict clinical progression to cognitive impairment; and (iii) model the levels of these markers along the levels of CSF A β 42 as a proxy of pathophysiological progression, and define if biomarker changes occurs before or after significant amyloidosis in the brain is established (structural break, see Supplementary Material 1).

CSF proteome profiling revealed a total of 100 unique proteins differentially regulated in amyloid positive CU compared to those with negative amyloid status (FDR < 0.05; Fig. 1b, Extended data Table 1 (ED Table 1)) after excluding those proteins that showed opposite effects in the independent EMIF-AD data set ($n=43$ proteins, Supplementary Material 1, Supplementary

Material: Table 2 and ED Table 1). The top 5 differentially regulated proteins are involved in immune function (ITGB2, CCL11), protein glycosylation and folding (ENTPD5), insulin growth factor signaling pathway (IGFBP-1) and protein phosphorylation (ABL1). CHIT1, involved in reactive gliosis and increased in different neurodegenerative dementias including AD [2], showed the strongest effects in correlation with amyloid pathology. This was followed by ITGB2, IGFBP-1, PRCP, LGMN (Fig. 1b and ED Table 1), the last two proteins are involved in lysosomal proteolytic function. The different proteins identified support the multifactorial biology of AD from the very early stages. Interestingly, 79 of the 100 unique proteins dysregulated in preclinical AD were also dysregulated in symptomatic stages of AD [2] (Fig. 1c) [2–4, 11, 12], supporting their association to AD. The fact that ITGB2 and APOL1 were the only two proteins among the top candidates previously identified at the AD dementia stage [2], underpins the importance of analyzing preclinical phases to identify relevant proteins at these earlier stages. Functional enrichment analysis showed that the 100 CSF markers dysregulated in preclinical AD were mainly associated to proteolysis and immune response (Supplementary Material: Fig. 2), pathways involved in AD pathophysiology associated with the development of amyloid and tau misfolding [1, 2]. This is in line with a recent targeted CSF proteomic study in autosomal AD showing that proteins associated with immune function were dysregulated 6 years before disease onset [5].

(See figure on next page.)

Fig. 1 CSF proteins of inflammation, proteolysis and lipid transport define preclinical AD and predict progression to AD dementia in cognitively unimpaired individuals **a** Protein levels in CSF from cognitively unimpaired individuals with and without amyloid pathology (based on CSF A β 42 concentration; CU(A β +) = 65; CU(A β -) = 232) were measured using antibody-based PEA technology. We defined which proteins differed across groups and a classification model to identify preclinical AD cases, which was validated in an independent cohort of cognitively unimpaired volunteers (EMIF-AD cohort, CU(A β -):103; CU(A β +) : 19). The performance of the model to predict clinical progression to cognitive impairment was also evaluated in a subset of cases from the discovery cohort (progressors to symptomatic stages: 39). We also evaluated the association of the proteins within the panel with proxies of progression of AD pathophysiology (e.g., CSF A β 42). **b** Volcano plot shows the CSF proteins that are differentially regulated in CU(A β +) vs. CU(A β -). Each dot represents a protein. The beta coefficients (log₂ fold-change) are plotted versus q values (-log₁₀-transformed FDR corrected p-value). Proteins significantly dysregulated after adjusting for false discovery rate (FDR, $q < 0.05$) are coloured in light green. The name of the top ten significant dysregulated CSF proteins and the top five with the strongest effect sizes are annotated. The total number of proteins that are down-regulated (left) or up-regulated (right) in CU(A β +) is indicated. Horizontal dotted line indicates the significance threshold. Adjusted p values ($q < 0.05$) were calculated using a two-sided nested linear model adjusting for FDR. **c** UpSet plot indicates the overlap across the proteins regulated in the preclinical phase of AD (CU(A β +) vs. CU(A β -)) and those at the prodromal or dementia AD stage (MCI(A β +) or AD (positive AD CSF profile) vs. controls) based on the results from our previous study [2]. **d** Receiver operating characteristic (ROC) curves depict the performance of 12-CSF protein panel to discriminate amyloid positive from amyloid negative cognitively unimpaired individuals in the ADC discovery and the EMIF-AD validation cohorts. In the discovery ADC cohort, black line is the mean Area Under the Curve (AUC) over all re-samplings (1000 repeats of fivefold cross-validation, grey lines). Inserts outline corresponding AUC and 95% CI. In the EMIF-AD validation cohort, insert outline the resulting AUC after directly applying the model developed with the discovery cohort. **e** ROC analysis depicts the performance of the CSF preAD panel to predict cases that progressed to MCI or dementia stage (progressors: 39; non progressors: 258). **f** CSF proteins within the preAD panel modelled along CSF A β 42 as an early proxy of AD pathology progression. Each NPX protein value was transformed to z-scores based on the distribution in the actual dataset to allow visual comparison across proteins. Bold line indicates the mean trajectory and shadows the 95% CI. Dotted line depicts CSF A β 42 positivity threshold (< 813 pg/mL). CN, cognitively unimpaired controls; preAD: preclinical AD (CU individuals with amyloid pathology); MCI, Mild cognitive impairment; AD, Alzheimer's disease; MMSE, Mini Mental Score Examination. Some images within Fig. 1a are courtesy of Olink® Proteomics AB

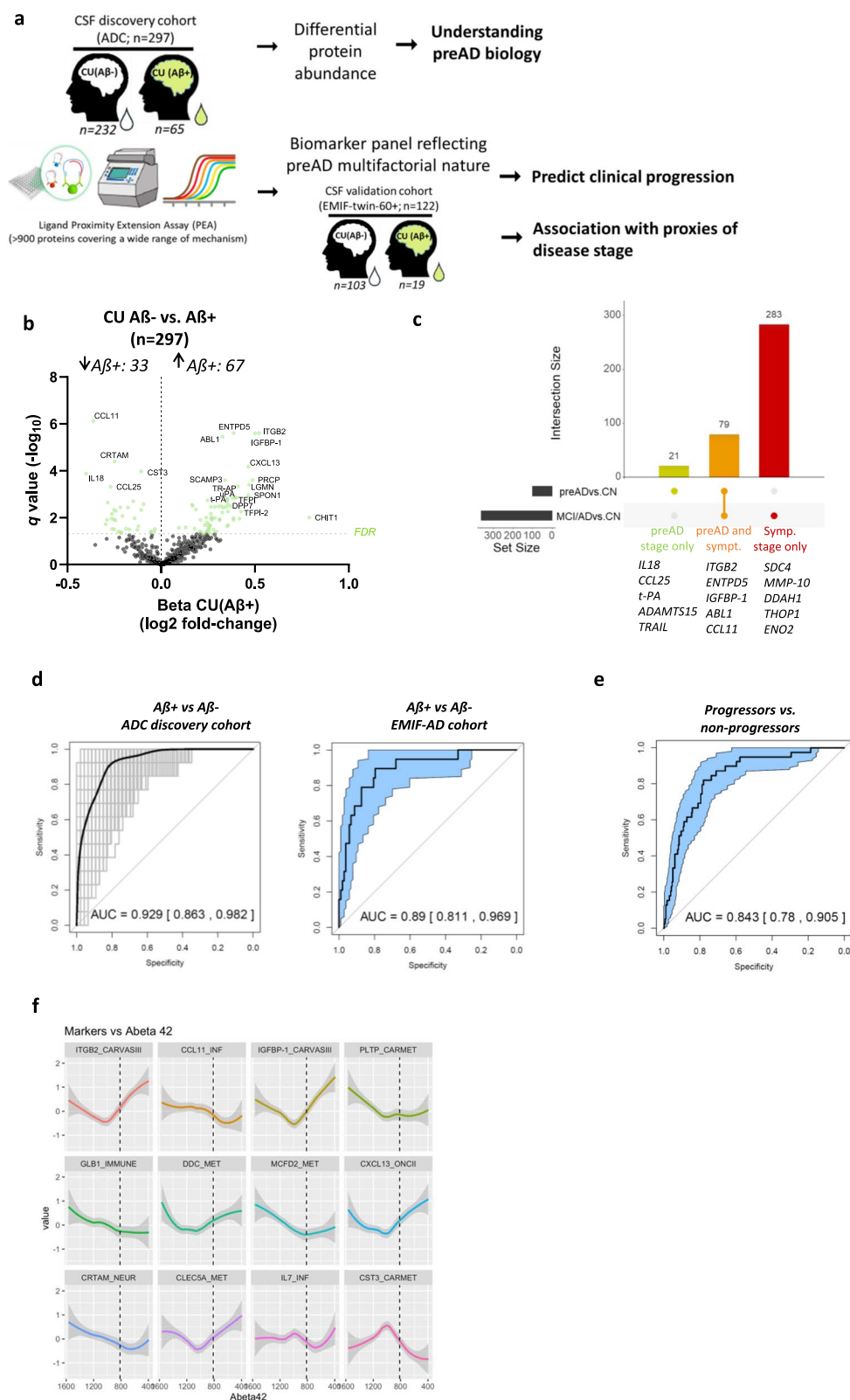


Fig. 1 (See legend on previous page.)

We next aimed to condense the proteomics data into practical biomarker signatures (minimal combination of proteins leading to the highest performance) using generalized linear modeling (GLM) with an elastic net penalty [2, 13]. Classification modelling revealed a panel of 12-CSF proteins to detect individuals in the preclinical phase of AD with an area Under the ROC curve (AUC) of 0.93, which was validated in the EMIF-AD cohort (0.89 AUC, Fig. 1d and Supplementary Material: Table 3). In line with the pathway enrichment analysis for the total regulated protein dataset, most of the proteins within the panel were mainly related to immune function (ITGB2, CXCL13, CLEC5A, CCL11, MCFD2, CRTAM, IL7). The panel also contained proteins related to dopamine biosynthesis (DDC), lysosome activity (GLB1), protease inhibition (CST3 or so-called Cystatin-C) or lipoprotein metabolism and lipid transport (IGFBP-1, PLTP). Interestingly, both CLEC5A and ITGB2 can regulate the expression or activation of TYROBP/DAP12 protein, the strongest microglia network regulator associated with sporadic late onset AD pathophysiology [14], which supports the potential role of these proteins in the earliest stages of AD pathophysiology. We know that many of these proteins (e.g., ITGB2, IGFBP-1, CLEC5A) were also specifically associated with AD when compared to a group with non-AD dementia (Supplementary Material: Fig. 3) [2], and that all markers except IL7 correlated with CSF A β_{42} or (p)Tau levels (Supplementary Material: Fig. 4), overall supporting their association to AD pathophysiology. We observed that the 12-CSF panel predicted progression to mild cognitive impairment (MCI) or dementia stage with 84% performance in the subset of cases followed clinically ($n=39$ out of 213; Fig. 1e). This performance increased to 90% when predicting the group that progressed to MCI or dementia due to AD (i.e. amyloid positive only, Supplementary Material: Figure 5a), underpinning the relevance of the proteins' panel and their mechanisms for progression to AD dementia. In line with those results, panel positivity at baseline was associated with increased risk of clinical dementia (Hazard ratio=8.37; $p<0.0001$; Supplementary Material: Fig. 5b). Linear mixed model analysis further showed that panel positivity was associated with a steeper cognitive decline over time as measured by mini mental score examination (MMSE, $p=0.01$, Supplementary Material: Fig. 6), supporting the clinical relevance of the panel.

We next modelled biomarker levels along the AD pathology using CSF A β_{42} as the proxy of early pathological changes [6] (Fig. 1f), and evaluated at what point of amyloid load the levels of these novel biomarkers start to change (structural break; Supplementary Material 1, Supplementary Material: Fig. 7). We observed that all 12 panel proteins had only one structural break in the slope

along the CSF A β_{42} values. Of note, all proteins except GLB1 and MCFD2 had a structural break before CSF A β_{42} positivity (i.e., CSF-A $\beta_{42}<813$ pg/mL; Supplementary Material: Fig. 7). These results provide additional support of the relation of these proteins to the earliest detectable stages of AD and suggest that processes related to immune function, energy metabolism, neurotrophic and endolysosomal functioning start to change before significant amyloidosis is present in the brain and might thus be relevant in disease etiology. Further experimental studies are needed to understand the connection between these proteins and progression of AD pathogenesis.

Some limitations should be considered. The number of cases that converted to MCI or dementia due to AD was limited, and thus the prognostic capabilities of the panel should be validated in larger and independent cohorts with substantial follow-up clinical data. As for most CSF studies, the study was cross-sectional. Despite our results identify CSF changes associated to the earliest AD stage defined to date, future analysis with longitudinal samples covering the full AD continuum would provide a more precise picture of the temporal evolution of these different processes in sporadic AD considering also interindividual differences (e.g., basal amyloid levels, accumulation rates, resilience factors). Noteworthy, PEA proteomic technology was optimized for blood analysis and thus, different effect sizes might be obtained when using single immunoassays. Still we and others have previously validated different proteins in independent AD cohorts using custom PEA assays (e.g., those relevant for the current study such as ITGB2, DDC, CLEC5A, PARK7) [2], as well as different single immunoassays (e.g., DDC, THOP1, sTREM1, MIF, GFAP, NfL) [11–13], supporting the validity of the results obtained with PEA proteomics.

Overall, this CSF proteome profiling shows that proteins involved in immune function, proteolysis and lipoprotein metabolisms are changed before the appearance of clinical signs. A selection of this proteins can be used not only to identify preclinical AD with >89% performance, but also to predict progression to AD dementia with good accuracy and likelihood (>84% and HR 8.37), supporting the association of these proteins to early AD pathophysiology. Some of these proteins were abnormal even before amyloidosis is detected in CSF, suggesting that these proteins and mechanisms might be useful targets for prevention of amyloidosis and clinical symptoms due to AD. The panel identified can easily be translated into custom assays [4, 11] containing the 12 markers of interest, especially those with the strongest effect sizes (e.g., ITGB2, CCL11). This panel can be used to track in vivo processes associated with AD beyond amyloid and tau pathology before the appearance of cognitive symptoms. These results depict the multifactorial

pathophysiology of AD in the earliest stages before amyloid pathology is established in vivo, providing new leads for the development of new therapeutics to counteract the development of AD and complementary biomarker tools for clinical settings and trials.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13024-024-00767-z>.

Supplementary Material 1.
Supplementary Material 2.
Supplementary Material 3.
Supplementary Material 4.

Authors' contributions

M.C, WvF and C.T conceived and designed the study. M.C, C.Q and C.P performed the statistical analysis. M.C, L.V, Y.H, A. B, I. V, P.J.V, B.T, W.M.F and C.T recruited participants and collected clinical data and samples. M.C and Y.H arranged and prepared samples for proteomics analysis. M.C and C.T drafted de manuscript. All authors contributed to revision and editing of the manuscript.

Authors' information

Not applicable.

Funding

This research is part of the neurodegeneration research program of Amsterdam Neuroscience. Alzheimer Center Amsterdam is supported by Stichting Alzheimer Nederland and Stichting VUmc funds. The chair Wiesje van der Flier is supported by the Pasman stichting. This study was supported by Alzheimer Nederland (CT, MC), ZonMW (#73305095007), Health~Holland, Topsector Life Sciences & Health (PPP-allowance; #LSHM20106). MC is supported by the attraction talent fellowship of Comunidad de Madrid (2018-T2/BMD-11885) and "PROYECTOS I+D+I – 2020"- Retos de investigación from the Ministerio Español de Ciencia e innovación (PID2020-115613RA-I00). The SCIENCe study has received funding from Gieskes-Strijbis Fonds and Stichting Dioraphte. WF and CT are recipient of TAP-Dementia, a ZonMW funded project (#10510032120003) under the Dutch National Dementia Strategy. Alzheimer Nederland, WE.03-2018-05, Charlotte Teunissen, Selfridges Group Foundation, NR170065, Charlotte Teunissen

Data availability

The source data generated in this study are available within this study (Extended data Table 1) or will be deposited in the synapse database under accession code https://www.synapse.org/PRIDE_SCD. The codes and scripts used in this study will be deposited in the synapse database under accession code https://www.synapse.org/PRIDE_SCD. All models were built using publicly available packages and functions in R.

Declarations

Ethics approval and consent to participate

The study was approved by the Institutional Ethical Review Boards (AD CSF biobank METC number. 00–211) and informed consent was obtained from all subjects or their authorized representatives.

Consent for publication

Not applicable.

Competing interests

MC has been an invited speaker at Eisai and Novonordisk. She is an associate editor at Alzheimer's Research & Therapy and has been an invited writer for Springer Healthcare. LV received a grant for CORAL consortium by Olink, and

has been invited speaker for Eli Lilly, and consultant for Roche. IV received a TKI Health~Holland grant for a collaborative project with OLINK and Quanterix. IV has been an invited speaker by Quanterix, funding is paid to her institution. WvF has been an invited speaker at Biogen MA Inc, Danone, Eisai, WebMD Neurology (Medscape), NovoNordisk, Springer Healthcare, European Brain Council. WF is consultant to Oxford Health Policy Forum CIC, Roche, Biogen MA Inc, and Eisai. WF is member of steering committee of NovoNordisk evoke/evoke+. WvF participated in advisory boards of Biogen MA Inc, Roche, and Eli Lilly. All funding is paid to her institution. WvF is member of the steering committee of PAVE, and Think Brain Health. WvF was associate editor of Alzheimer, Research & Therapy in 2020/2021. WvF is associate editor at Brain. CET has a collaboration contract with ADx Neurosciences, Quanterix and Eli Lilly, performed contract research or received grants from AC-Immune, Axon Neurosciences, Bioconnect, Bioorchestra, Brainstorm Therapeutics, Celgene, EIP Pharma, Eisai, Grifols, Novo Nordisk, PeopleBio, Roche, Toyama, Vivoryon. She serves on editorial boards of Medidact Neurologie/Springer, Alzheimer Research and Therapy, Neurology: Neuroimmunology & Neuroinflammation, and is editor of a Neuromethods book Springer. She had speaker contracts for Roche, Grifols, Novo Nordisk. The rest of the authors declare no competing interest.

Received: 20 April 2024 Accepted: 11 October 2024

Published online: 11 November 2024

References

- Jansen IE, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet.* 2019;12:43.
- del Campo M, et al. CSF proteome profiling across the Alzheimer's disease spectrum reflects the multifactorial nature of the disease and identifies specific biomarker panels. *Nat Aging.* 2022;2:1040–53. <https://doi.org/10.1038/s43587-022-00300-1>.
- Higginbotham, L, et al. Integrated Proteomics Reveals Brain-Based Cerebrospinal Fluid Biomarkers in Asymptomatic and Symptomatic Alzheimer's Disease. *Sci Adv.* 2020;6:eaaz9360, 806752. <https://doi.org/10.1126/sciadv.aaz9360>.
- Bader JM, et al. Proteome profiling in cerebrospinal fluid reveals novel biomarkers of Alzheimer's disease. *Mol Syst Biol.* 2020;16(6):e9356 1–17. <https://doi.org/10.15252/msb.20199356>.
- Johnson ECB, et al. Cerebrospinal fluid proteomics define the natural history of autosomal dominant Alzheimer's disease. *Nat Med.* 2023;29:1979–88.
- Jack CR. Advances in Alzheimer's disease research over the past two decades. *Lancet Neurol.* 2022;21:866–9.
- van Bokhoven P, et al. The Alzheimer's disease drug development landscape. *Alzheimers Res Ther.* 2021;13(1):186. <https://doi.org/10.1186/s13195-021-00927-z>.
- Slot RER, et al. Subjective Cognitive Impairment Cohort (SCIENCe): study design and first results. *Alzheimers Res Ther.* 2018;10:76.
- Verberk IMW, et al. Plasma Amyloid as Prescreener for the Earliest Alzheimer Pathological Changes. *Ann Neurol.* 2018;84:648–58.
- Konijnenberg E, et al. The EMIF-AD PreclinAD study: study design and baseline cohort overview. *Alzheimers Res Ther.* 2018;10:75.
- Yanaika S, Hok-A-Hin et al. Thimet Oligopeptidase as a potential CSF biomarker for Alzheimer's Disease. A cross-platform validation study. *Alzheimers Dement Diagn Assess Dis Monit Press.* 2023;15(3):e12456. <https://doi.org/10.1002/dad2.12456>.
- Hok-A-Hin YS, et al. Neuroinflammatory CSF biomarkers MIF, sTREM1, and sTREM2 show dynamic expression profiles in Alzheimer's disease. *J Neuroinflamm.* 2023;20(1):107. <https://doi.org/10.1186/s12974-023-02796-9>.
- del Campo M, et al. CSF proteome profiling reveals biomarkers to discriminate dementia with Lewy bodies from Alzheimer's disease. *Nat Commun.* 2023;14:5635.
- Zhang B, et al. Integrated Systems Approach Identifies Genetic Nodes and Networks in Late-Onset Alzheimer's Disease. *Cell.* 2013;153:707–20.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.