Novel CSF tau biomarkers can be used for disease staging of
2 **sporadic Alzheimer's disease**
3 Gemma Salvadó^{1,#}, Kanta Horie^{2,3,4}, Nicolas R. Barthélemy^{2,3}, Jacob W. Vogel^{1,5}, Alexa
2 Pichet Binette¹, Charles D **sporadic Alzheimer's disease**

3

Gemma Salvadó^{1,#}, Kanta Horie^{2,3,4}, Nicolas R. Barthélemy^{2,3}, Jacob

Pichet Binette¹, Charles D. Chen⁶, Andrew J Aschenbrenner^{3,7}, Bria

6 L.S. Benzinger^{6,7}, David M. Holtzm 456789 Gemma Salvadó^{1,#}, Kanta Horie^{2,3,4}, Nicolas R. Barthélemy^{2,3}, Jacob W. Vogel^{1,5}, Alexa 4 Gemma Salvadó^{1,#}, Kanta Horie^{2,3,4}, Nicolas R. Barthélemy^{2,3}, Jacob W. Vogel^{1,5}, Alexa

5 Pichet Binette¹, Charles D. Chen⁶, Andrew J Aschenbrenner^{3,7}, Brian A. Gordon⁶, Tamm

1. S. Benzinger^{6,7}, David Pichet Binette¹, Charles D. Chen⁶, Andrew J Aschenbrenner^{3,7}, Brian A. Gordon⁶ Fichet Binette', Charles D. Chen^o, Andrew J Aschenbrenner^{3,7}, Brian A. Gordon^o, Tammie

L.S. Benzinger^{6,7}, David M. Holtzman^{3,7}, John C. Morris^{3,7}, Sebastian Palmqvist¹, Erik

Stomrud^{1,8}, Shorena Janelidze L.S. Benzinger^{6,7}, David M. Holtzman^{3,7}, John C. Morris^{3,7}, Sebastian Palmqvist¹, Erik 5 L.S. Benzinger^{6,7}, David M. Holtzman^{3,7}, John C. Morris^{3,7}, Sebastian Palmqvist¹, Erik
5 Stomrud^{1,8}, Shorena Janelidze¹, Rik Ossenkoppele^{1,9,10}, Suzanne E. Schindler^{3,7}, Ran
5 J. Bateman^{2,3,7}, Oskar Ha Stomrud^{1,8}, Shorena Janelidze¹, Rik Ossenkoppele^{1,9,10}, Suzanne E. Schindler^{3,7} Stomrud^{1,8}, Shorena Janelidze¹, Rik Ossenkoppele^{1,9,10}, Suzanne E. Schindler^{3,7}, Randall

J. Bateman^{2,3,7}, Oskar Hansson^{1,8,#}

9

¹ Clinical Memory Research Unit, Department of Clinical Sciences Malmö, Lund U J. Bateman^{2,3,7}, Oskar Hansson^{1,8,#} 901234 0 1 2 3 4 5 1 ¹ Clinical Memory Research Unit, Department of Clinical Sciences Malmö, Lund University,

11 Lund, Sweden

² The Tracy Family SILQ Center, Washington University School of Medicine, St Louis, MO,

United States

³ Dep 11 Lund, Sweden
12 ² The Tracy Fa
13 United States
14 ³ Department o
15 United States
16 ⁴ Eisai Inc., Nut ² The Tracy Family SILQ Center, Washington University School of Medicine, St Louis, MO, ² The Tracy Family SILQ Center, Washington University School of Medicine, St Louis, MO,

13 United States

³ Department of Neurology, Washington University School of Medicine, St. Louis, MO,

³ United States

⁴ Eis 13 United States
14 ³ Department
15 United States
16 ⁴ Eisai Inc., Nu
17 ⁵ Department of
18 ⁶ Department of ³ Department of Neurology, Washington University School of Medicine, St. Louis, MO, ³ Department of Neurology, Washington University School of Medicine, St. Louis, MO,

15 United States

⁴ Eisai Inc., Nutley, NJ, United States

⁵ Department of Clinical Science, Malmö, SciLifeLab, Lund University, Lu 15 United States
16 ⁴ Eisai Inc., Nu
17 ⁵ Department of
⁶ Department of
19 ⁷ Charles F. ar
20 University Sch
21 ⁸ Memory Clin ⁴ Eisai Inc., Nutley, NJ, United States ⁴ Eisai Inc., Nutley, NJ, United States

⁵ Department of Clinical Science, Malm

⁶ Department of Radiology, Washington

⁷ Charles F. and Joanne Knight Alzhei

20 University School of Medicine, St. Lou

⁸ Memory C ⁵ Department of Clinical Science, Malmö, SciLifeLab, Lund University, Lund, Sweden ⁹ Department of Clinical Science, Malmö, SciLifeLab, Lund University, Lund, Sweden

⁶ Department of Radiology, Washington University School of Medicine, St. Louis, MO,

⁷ Charles F. and Joanne Knight Alzheimer Diseas ⁶ Department of Radiology, Washington University School of Medicine, St. Louis, MO, USA ^o Department of Radiology, Washington University School of Medicine, St. Louis, MO, USA

⁷ Charles F. and Joanne Knight Alzheimer Disease Research Center, Washington

⁷ Oniversity School of Medicine, St. Louis, MO, U ⁷ Charles F. and Joanne Knight Alzheimer Disease Research Center, Washington ² Charles F. and Joanne Knight Alzheimer Disease Research Center, Washington

20 University School of Medicine, St. Louis, MO, USA

⁸ Memory Clinic, Skåne University Hospital, Malmö, Sweden

⁹ Alzheimer Center Amster 20 University School of Medicine, St. Louis, MO, USA

⁸ Memory Clinic, Skåne University Hospital, Malmö

⁹ Alzheimer Center Amsterdam, Neurology, Vrije U

location VUmc, Amsterdam, The Netherlands

¹⁰ Amsterdam Neuro ⁸ Memory Clinic, Skåne University Hospital, Malmö, Sweden ⁸ Memory Clinic, Skåne University Hospital, Malmö, Sweden

⁹ Alzheimer Center Amsterdam, Neurology, Vrije Universiteit

¹⁰ Amsterdam Neuroscience, Neurodegeneration, Amsterdam

¹⁰ Amsterdam Neuroscience, Neurodegen ⁹ Alzheimer Center Amsterdam, Neurology, Vrije Universiteit Amsterdam, Amsterdam UMC ⁹ Alzheimer Center Amsterdam, Neurology, Vrije Universiteit Amsterdam, Amsterdam UMC

23 Iocation VUmc, Amsterdam, The Netherlands

¹⁰ Amsterdam Neuroscience, Neurodegeneration, Amsterdam, The Netherlands

25 **# Corres** 23 location VUmc, Amsterdam, The Netherlands

24 ¹⁰ Amsterdam Neuroscience, Neurodegenerat

25 **# Corresponding authors:**

27 Oskar Hansson, MD, PhD

28 Memory Clinic, Skåne University Hospital

29 SE-205 02, Malmö, Swed ¹⁰ Amsterdam Neuroscience, Neurodegeneration, Amsterdam, The Netherlands ¹⁰ Amsterdam Neuroscience, Neurodegeneration, Amsterdam, The Netherlands
25 **# Corresponding authors:**
27 Oskar Hansson, MD, PhD
28 Memory Clinic, Skåne University Hospital
30 E-mail: oskar hansson@med.lu.se: Tel: +46 (0 26
27
28
29
30
31 **# Corresponding authors:**

27 Oskar Hansson, MD, PhD

28 Memory Clinic, Skåne Unive

29 SE-205 02, Malmö, Sweden

30 E-mail: oskar.hansson@med

31 27 Oskar Hansson, MD, PhD
28 Memory Clinic, Skåne Univ
29 SE-205 02, Malmö, Swede
30 E-mail: oskar hansson@m
31
22 Or 28 Memory Clinic, Skåne University Hospital
29 SE-205 02, Malmö, Sweden
30 E-mail: oskar.hansson@med.lu.se; Tel: +4
31 Or 29 SE-205 02, Malmö, Sweden
30 E-mail: oskar.hansson@med
31
32 Or
33 Gemma Salvadó, PhD 30 E-mail: oskar.hansson@med.lu.se; Tel: +46 (0)40 331000
31
32 Or
33 Gemma Salvadó, PhD
35 Clinical Memory Research Unit, Department of Clinical Scie
36 OF 995.99 Make ³ Quester 32
32
33
34
35
36 32 Or

33

34 Ge

35 Clir

36 SE

37 E-n

32 34
35
36
37
38 34 Gemma Salvadó, PhD
35 Clinical Memory Resea
36 SE-205 02, Malmö, Sw
37 E-mail: gemma.salvado Clinical Memory Research Unit, Department of Clinical Sciences
36 SE-205 02, Malmö, Sweden
37 E-mail: gemma.salvado@med.lu.se; Tel: +46 (0)40 331000
38
NOTE: This preprint reports new research that has not been certified b 36 SE-205 02, Malmö, Sweden
37 E-mail: gemma.salvado@me
38
NOTE: This preprint reports new research that 37 E-mail: gemma.salvado@med.lu.se; Tel: +46 (0)40 331000
38
NOTE: This preprint reports new research that has not been certified by peer review and s

39 **Abstract**
40 Biological
41 and progne
42 Here, we e
43 algorithm b
44 426 partic Biological staging of individuals with Alzheimer's disease (AD) may improve diagnostic

and prognostic work-up of dementia in clinical practice and the design of clinical trials.

Here, we created a staging model using th and prognostic work-up of dementia in clinical practice and the design of clinical trials.

42 Here, we created a staging model using the Subtype and Stage Inference (SuStaln)

43 algorithm by evaluating cerebrospinal flu Here, we created a staging model using the Subtype and Stage Inference (SuStaIn)
algorithm by evaluating cerebrospinal fluid (CSF) amyloid-β (Aβ) and tau biomarkers in
426 participants from BioFINDER-2, that represent th algorithm by evaluating cerebrospinal fluid (CSF) amyloid-β (Aβ) and tau biomarkers in
44 426 participants from BioFINDER-2, that represent the entire spectrum of AD. The
45 model composition and main analyses were repli 44 426 participants from BioFINDER-2, that represent the entire spectrum of AD. The
45 model composition and main analyses were replicated in 222 participants from the
46 Knight ADRC cohort. SuStaln revealed in the two co model composition and main analyses were replicated in 222 participants from the

46 Knight ADRC cohort. SuStaln revealed in the two cohorts that the data was best

47 explained by a single biomarker sequence (one subtype Explained by a single biomarker sequence (one subtype), and that five CSF biomarkers
explained by a single biomarker sequence (one subtype), and that five CSF biomarkers
(ordered: A β 42/40, tau phosphorylation occupanci explained by a single biomarker sequence (one subtype), and that five CSF biomarkers

(ordered: Aβ42/40, tau phosphorylation occupancies at the residues 217 and 205

[pT217/T217 and pT205/T205], microtubule-binding regio 48 (ordered: Aβ42/40, tau phosphorylation occupancies at the residues 217 and 205 [pT217/T217 and pT205/T205], microtubule-binding region of tau containing the residue 243 [MTBR-tau243], and total tau) were sufficient to 19 [pT217/T217 and pT205/T205], microtubule-binding region of tau containing the
150 residue 243 [MTBR-tau243], and total tau) were sufficient to create an accurate
151 disease staging model. Increasing CSF stages (0-5) w residue 243 [MTBR-tau243], and total tau) were sufficient to create an accurate
51 disease staging model. Increasing CSF stages (0-5) were associated with increased
52 abnormality in other AD-related biomarkers, such as A disease staging model. Increasing CSF stages (0-5) were associated with increased

abnormality in other AD-related biomarkers, such as A β - and tau-PET, and aligned with

different phases of longitudinal biomarker chang abnormality in other AD-related biomarkers, such as Aβ- and tau-PET, and aligned with
different phases of longitudinal biomarker changes consistent with current models of
AD progression. Higher CSF stages at baseline were different phases of longitudinal biomarker changes consistent with current models of

AD progression. Higher CSF stages at baseline were associated with higher hazard

ratio of clinical decline. Our findings indicate that 54 AD progression. Higher CSF stages at baseline were associated with higher hazard
55 ratio of clinical decline. Our findings indicate that a common pathophysiologic
56 molecular pathway develops across all AD patients, a 55 ratio of clinical decline. Our findings indicate that a common pathophysiologic
56 molecular pathway develops across all AD patients, and that a single CSF collection is
57 sufficient to reliably indicate the presence o 56 molecular pathway develops across all AD patients, and that a single CSF collection is
sufficient to reliably indicate the presence of both AD pathologies and the degree and
stage of disease progression. 57 sufficient to reliably indicate the presence of both AD pathologies and the degree and
58 stage of disease progression.
59 stage of disease progression. 58 stage of disease progression.

59 **Introduction**
60 Currently, more
61 expected to more
62 form of dement
63 amyloid-β (Aβ)
64 threads². Over 60 Currently, more than 50 million people are affected by dementia and this number is
61 expected to more than double by 2050¹. Alzheimer's disease (AD) is the most common
62 form of dementia, characterized by the accum expected to more than double by 2050^{\degree} . Alzheimer's disease (AD) is the most common Eq. expected to more than double by 2050¹. Alzheimer's disease (AD) is the most common
form of dementia, characterized by the accumulation of extracellular plaques containing
amyloid-β (Aβ) and intracellular tau aggreg 62 form of dementia, characterized by the accumulation of extracellular plaques containing
63 amyloid-β (Aβ) and intracellular tau aggregates in the forms of tau tangles and neuropil
64 threads². Over the last two deca 63 amyloid-β (Aβ) and intracellular tau aggregates in the forms of tau tangles and neuropil
64 threads². Over the last two decades, the AD field has moved towards the use of
65 biomarkers to support the diagnostic and threads². Over the last two decades, the AD field has moved towards the use of threads². Over the last two decades, the AD field has moved towards the use of
biomarkers to support the diagnostic and prognostic work-up, rather than relying solely
on clinical symptoms³. This has been made possible 65 biomarkers to support the diagnostic and prognostic work-up, rather than relying solely
66 on clinical symptoms³. This has been made possible by advancements of imaging and
67 fluid biomarkers that accurately track A on clinical symptoms³. This has been made possible by advancements of imaging and 66 on clinical symptoms³. This has been made possible by advancements of imaging and
67 fluid biomarkers that accurately track AD pathology *in vivo*. Given that the accumulation
68 of pathology can take many years to d fluid biomarkers that accurately track AD pathology *in vivo*. Given that the accumulation
of pathology can take many years to decades³, before any clinical symptoms appear,
the use of biomarkers is critical to ensuring of pathology can take many years to decades³, before any clinical symptoms appear. 68 of pathology can take many years to decades³, before any clinical symptoms appear,
69 the use of biomarkers is critical to ensuring an early and reliable detection of AD⁴. Key
69 biomarkers may help to improve pati the use of biomarkers is critical to ensuring an early and reliable detection of $AD⁴$. Key the use of biomarkers is critical to ensuring an early and reliable detection of AD⁴. Key

70 biomarkers may help to improve patient diagnosis, management and prognosis^{5–8}. And

71 the use of AD biomarkers will be eve biomarkers may help to improve patient diagnosis, management and prognosis⁵⁻⁸. And biomarkers may help to improve patient diagnosis, management and prognosis⁵⁻⁸. And

the use of AD biomarkers will be even more important when disease-modifying

treatments become widely available $9-11$. In this context The use of AD biomarkers will be even more important when disease-modifying

The treatments become widely available $9-11$. In this context, a more sophisticated

The personalized medicine approach of AD, based on high pe treatments become widely available $9-11$. In this context, a more sophisticated The treatments become widely available $9-11$. In this context, a more sophisticated
personalized medicine approach of AD, based on high performing AD biomarkers, will
become crucial to select the most optimal participant

personalized medicine approach of AD, based on high performing AD biomarkers, will

14 become crucial to select the most optimal participants for specific treatments and for

15 enrolment in new clinical trials.

16 In rec become crucial to select the most optimal participants for specific treatments and for
enrolment in new clinical trials.
T6 In recent years, multiple cerebrospinal fluid (CSF) biomarkers targeting different
pathophysiolog 75 enrolment in new clinical trials.
76 In recent years, multiple cer
77 pathophysiological mechanism:
78 been an increasing interest i
79 phosphorylated at different resi
80 tau181^{12–17}, p-tau217^{12,13,15,18,} The propertion of the methodographic mechanisms have been developed (see 4 for a review). There has

been an increasing interest in developing biomarkers for measuring tau species

phosphorylated at different residues. pathophysiological mechanisms have been developed (see 4 for a review). There has pathophysiological mechanisms have been developed (see 4 for a review). There has

been an increasing interest in developing biomarkers for measuring tau species

phosphorylated at different residues. Among the phosph been an increasing interest in developing biomarkers for measuring tau species

79 phosphorylated at different residues. Among the phosphorylated tau (p-tau) species, p-

80 tau181^{12–17}, p-tau217^{12,13,15,18,19}, and p-279 phosphorylated at different residues. Among the phosphorylated tau (p-tau) species, p-

270 tau181^{12–17}, p-tau217^{12,13,15,18,19}, and p-tau231^{15,20–22} or the phosphorylation

271 occupancies have been studied in tau181^{12–17}, p-tau217^{12,13,15,18,19}, and p-tau231^{15,20–22} and p-tau231^{15,20-22} or the phosphorylation
81 occupancies have been studied in depth, and have shown strong associations with A_B
82 pathology and moderate associations with tau (as measured by both PET^{18,23} and
83 81 occupancies have been studied in depth, and have shown strong associations with Aβ pathology and moderate associations with tau (as measured by both PET^{18,23} and neuropathology^{24,25}). These biomarkers have shown t pathology and moderate associations with tau (as measured by both $PET^{18,23}$ and and moderate associations with tau (as measured by both $PET^{18,23}$ and

83 neuropathology^{24,25}). These biomarkers have shown their utility on improving the

84 diagnostic work-up of AD and the prediction of disease pro neuropathology^{24,25}). These biomarkers have shown their utility on improving the neuropathology^{24,25}). These biomarkers have shown their utility on improving the
84 diagnostic work-up of AD and the prediction of disease progression^{12,13,19,26,27}. Other
85 biomarkers, such as p-tau205 or the occupa diagnostic work-up of AD and the prediction of disease progression^{12,13,19,26,27} diagnostic work-up of AD and the prediction of disease progression^{12,13,19,26,27}. Other
85 biomarkers, such as p-tau205 or the occupancy (pT205/T205)^{28–30} and microtubule
86 binding region (MTBR) of tau containing the biomarkers, such as p-tau205 or the occupancy (pT205/T205) $^{28-30}$ biomarkers, such as p-tau205 or the occupancy (pT205/T205)²⁸⁻³⁰ and microtubule

86 binding region (MTBR) of tau containing the 243 residue (MTBR-tau243)^{31,32}, have

87 been more closely related to tau tangle patholog binding region (MTBR) of tau containing the 243 residue (MTBR-tau243) 31,32 binding region (MTBR) of tau containing the 243 residue (MTBR-tau243)^{31,32}, have
87 been more closely related to tau tangle pathology. Importantly, some of these CSF
88 biomarkers were shown to become abnormal at differe been more closely related to tau tangle pathology. Importantly, some of these CSF
biomarkers were shown to become abnormal at different phases during the
progression of autosomal dominant AD (ADAD)²⁹, suggesting a sequ biomarkers were shown to become abnormal at different phases during the

89 progression of autosomal dominant AD (ADAD)²⁹, suggesting a sequence of CSF

90 biomarker changes that may serve as a measurable biological in progression of autosomal dominant AD $(ADAD)^{29}$, suggesting a sequence of CSF

89 progression of autosomal dominant AD (ADAD)²⁹, suggesting a sequence of CSF
biomarker changes that may serve as a measurable biological indicator tracking
advancing disease progression.
The progression of Aβ- or tau 90 biomarker changes that may serve as a measurable biological indicator tracking
91 advancing disease progression.
92 The progression of Aβ- or tau- pathology across the brain has been previously used to
93 stage partic 91 advancing disease progression.
92 The progression of Aβ- or tau- p
93 stage participants across the AE
94 one Aβ- or tau-PET scan, which 92 The progression of Aβ- or tau- pathology across the brain has been previously used to
93 stage participants across the AD *continuum*³³⁻³⁸. However, these models need at least
94 one Aβ- or tau-PET scan, which is ex stage participants across the AD *continuum*33–38 93 stage participants across the AD *continuum*³³⁻³⁸. However, these models need at least
 94 one A β - or tau-PET scan, which is expensive, and requires specialized personnel and
 2 94 one Aβ- or tau-PET scan, which is expensive, and requires specialized personnel and
2

facilities. Further, information of only one pathological measure (*e.g.*, Aβ or tau) can be
obtained from these images. On the other hand, CSF biomarkers are less expensive,
more accessible, and multiple pathological mea 96 obtained from these images. On the other hand, CSF biomarkers are less expensive,
97 more accessible, and multiple pathological measures may be obtained from a single
98 sample. Given this, and with the idea that diff more accessible, and multiple pathological measures may be obtained from a single
98 sample. Given this, and with the idea that different CSF biomarkers may become
99 abnormal at different stages of the disease, we aimed 98 sample. Given this, and with the idea that different CSF biomarkers may become
99 abnormal at different stages of the disease, we aimed to generate a data-driven
00 staging scheme for sporadic AD using key CSF tau biom 99 abnormal at different stages of the disease, we aimed to generate a data-driven

00 staging scheme for sporadic AD using key CSF tau biomarkers in combination with

01 CSF Aβ42/40. An unresolved question is whether th 100 staging scheme for sporadic AD using key CSF tau biomarkers in combination with

101 CSF Aβ42/40. An unresolved question is whether there is a single molecular pathway

102 throughout the AD *continuum*, or whether t 101 CSF Aβ42/40. An unresolved question is whether there is a single molecular pathway

102 throughout the AD *continuum*, or whether there are subtypes of AD following different

103 fluid biomarker trajectories, as has tangles $36,39,40$.

throughout the AD *continuum*, or whether there are subtypes of AD following different
103 fluid biomarker trajectories, as has been shown for regional spread of insoluble tau
104 tangles ^{36,39,40}.
105 To test this, we u 103 fluid biomarker trajectories, as has been shown for regional spread of insoluble tau

104 tangles $36,39,40$.

105 To test this, we used Subtype and Stage Inference $(SuStaln)^{41}$ to model the most likely

106 sequence 104 tangles $36,39,40$.

105 To test this, we

106 sequence of C

107 machine-learni

108 a probabilistic

109 of this orderin To test this, we used Subtype and Stage Inference $(SuStaln)^{41}$ to model the most likely To test this, we used Subtype and Stage Inference (SuStaIn)⁴¹ to model the most likely
106 sequence of CSF biomarker abnormalities that occur along the AD time course. This
107 machine-learning method uses cross-sectiona sequence of CSF biomarker abnormalities that occur along the AD time course. This
107 machine-learning method uses cross-sectional data to order biomarker abnormalities in
108 a probabilistic manner and at the same time ad machine-learning method uses cross-sectional data to order biomarker abnormalities in
108 a probabilistic manner and at the same time addresses possible diverging trajectories
109 of this ordering. With this approach, we s 108 a probabilistic manner and at the same time addresses possible diverging trajectories

109 of this ordering. With this approach, we staged 426 participants of the Swedish

110 BioFINDER-2 study, ranging from cognitiv 109 of this ordering. With this approach, we staged 426 participants of the Swedish

110 BioFINDER-2 study, ranging from cognitively unimpaired (CU) participants to patients

111 with mild cognitive impairment (MCI) or d 110 BioFINDER-2 study, ranging from cognitively unimpaired (CU) participants to patients

111 with mild cognitive impairment (MCI) or dementia. This model was used to assign each

112 participant to a fluid biomarker sta 111 with mild cognitive impairment (MCI) or dementia. This model was used to assign each

112 participant to a fluid biomarker stage, which was subsequently correlated with

113 established AD key features such as Aβ- an 112 participant to a fluid biomarker stage, which was subsequently correlated with

113 established AD key features such as Aβ- and tau-PET, cortical thickness, and

114 cognition. Next, we investigated the accuracy of o 113 established AD key features such as Aβ- and tau-PET, cortical thickness, and

114 cognition. Next, we investigated the accuracy of our staging model to predict amyloid

115 (A) and tau (T) status as defined by PET⁴² 114 cognition. Next, we investigated the accuracy of our staging model to predict amyloid (A) and tau (T) status as defined by PET^{42} , as well as its potential as a diagnostic tool for distinguishing diagnostic groups. (A) and tau (T) status as defined by PET^{42} , as well as its potential as a diagnostic tool 115 (A) and tau (T) status as defined by PET^{42} , as well as its potential as a diagnostic tool
116 for distinguishing diagnostic groups. Using longitudinal data, we further determined
117 trajectories of several key AD-116 for distinguishing diagnostic groups. Using longitudinal data, we further determined
117 trajectories of several key AD-biomarkers based on participants baseline fluid
118 biomarker disease stage. We also tested the cl 117 trajectories of several key AD-biomarkers based on participants baseline fluid
118 biomarker disease stage. We also tested the clinical utility of our novel staging system
119 for predicting clinical progression. Final 118 biomarker disease stage. We also tested the clinical utility of our novel staging system
119 for predicting clinical progression. Finally, we performed the main analyses in the
120 independent cohort (the Charles F. an 119 for predicting clinical progression. Finally, we performed the main analyses in the

120 independent cohort (the Charles F. and Joanne Knight Alzheimer Disease Research

121 Center [Knight-ADRC]), which included 222 pa 120 independent cohort (the Charles F. and Joanne Knight Alzheimer Disease Research
121 Center [Knight-ADRC]), which included 222 participants. Altogether our results suggest
122 that participants in the AD continuum progr 121 Center [Knight-ADRC]), which included 222 participants. Altogether our results suggest

122 that participants in the AD continuum progress along a single path and can be

123 biologically staged using a single sample o that participants in the AD *continuum* progress along a single path and can be

biologically staged using a single sample of CSF. This may have important implications

in the clinical practice and in the selection of part

biologically staged using a single sample of CSF. This may have important implications
124 in the clinical practice and in the selection of participants for future clinical trials.
125 **Results**
127 A total of 426 particip 124 in the clinical practice and in the selection of participants for future clinical trials.

125
 Results

127 A total of 426 participants of the Swedish BioFINDER-2 study (NCT03174938

128 complete CSF data were incl --
126
127
128
129
130 126 **Results**

127 A total of

128 complete

129 unimpaire

130 participan A total of 426 participants of the Swedish BioFINDER-2 study (NCT03174938)¹⁹ with 217 A total of 426 participants of the Swedish BioFINDER-2 study (NCT03174938)¹⁹ with
128 complete CSF data were included in this study. From these, 80 were cognitively
129 unimpaired Aβ negative (CU-) and 79 cognitive 128 complete CSF data were included in this study. From these, 80 were cognitively

129 unimpaired Aβ negative (CU-) and 79 cognitively unimpaired Aβ positive (CU+)

130 participants, 88 were diagnosed with MCI and Aβ po 129 unimpaired Aβ negative (CU-) and 79 cognitively unimpaired Aβ positive (CU+)
130 participants, 88 were diagnosed with MCI and Aβ positive, 100 were diagnosed with
3 130 participants, 88 were diagnosed with MCI and Aβ positive, 100 were diagnosed with
3

131 AD dementia and Aβ positive (ADD+), and 79 were assessed as non-AD patients (22
132 were Aβ positive). Demographic information is presented in Table 1. Of these, 220
133 participants had longitudinal CSF data availabl

-
-

132 were Aβ positive). Demographic information is presented in Table 1. Of these, 220

133 participants had longitudinal CSF data available (Supplementary Table 1).

134 **CSF staging model**

136 We initially applied SuSt 133 participants had longitudinal CSF data available (Supplementary Table 1).

134 **CSF staging model**

135 **CSF staging model**

136 We initially applied SuStaIn to the BioFINDER-2 cohort using the fo

137 biomarkers: th 135
136
137
138
139 **CSF staging model**
136 We initially applied
137 biomarkers: the Aβ4
138 pT205/T205, pT181/
139 of MTBR-tau243 and
140 literature. Through a 136 We initially applied SuStaIn to the BioFINDER-2 cohort using the following CSF
137 biomarkers: the A β 42/40 ratio, the phosphorylated to non-phosphorylated tau ratio of
138 pT205/T205, pT181/T181, pT217/T217, and pT 137 biomarkers: the Aβ42/40 ratio, the phosphorylated to non-phosphorylated tau ratio of
138 pT205/T205, pT181/T181, pT217/T217, and pT231/T231, as well as the concentrations
139 of MTBR-tau243 and total tau [the residue 138 pT205/T205, pT181/T181, pT217/T217, and pT231/T231, as well as the concentrations

139 of MTBR-tau243 and total tau [the residue 151-153]) based on availability and previous

140 literature. Through a process of model 139 of MTBR-tau243 and total tau [the residue 151-153]) based on availability and previous

140 literature. Through a process of model optimization (Ext. Data Fig. 1, see Methods for

141 further details), we arrived on a 140 literature. Through a process of model optimization (Ext. Data Fig. 1, see Methods for

141 further details), we arrived on a model that excluded pT181/T181 and pT231/T231 due

142 to information redundancy. SuStaln r 141 further details), we arrived on a model that excluded pT181/T181 and pT231/T231 due
142 to information redundancy. SuStaln revealed that a single biomarker sequence best
143 described the progressive abnormality of th 142 to information redundancy. SuStaIn revealed that a single biomarker sequence best
143 described the progressive abnormality of the selected biomarkers (Ext Data Fig. 1C).
144 The final ordering of the model was the A143 described the progressive abnormality of the selected biomarkers (Ext Data Fig. 1C).

144 The final ordering of the model was the A β 42/40 ratio, pT217/T217, pT205/T205,

145 MTBR-tau243 and total-tau (Fig. 1A), res 144 The final ordering of the model was the Aβ42/40 ratio, pT217/T217, pT205/T205,

145 MTBR-tau243 and total-tau (Fig. 1A), resulting in a five-stage model (plus stage 0 as a

146 negative biomarker stage). All BioFINDE 145 MTBR-tau243 and total-tau (Fig. 1A), resulting in a five-stage model (plus stage 0 as a
146 negative biomarker stage). All BioFINDER-2 participants were then classified into one
147 of these biomarker-based disease sta 146 negative biomarker stage). All BioFINDER-2 participants were then classified into one

147 of these biomarker-based disease stages based on their CSF levels, with 124 (29.1%)

148 being at CSF stage 0, 35 (8.2%) at CS 147 of these biomarker-based disease stages based on their CSF levels, with 124 (29.1%)

148 being at CSF stage 0, 35 (8.2%) at CSF stage 1, 53 (12.4%) at CSF stage 2, 49

149 (11.5%) at CSF stage 3, 87 (20.4%) at CSF sta 148 being at CSF stage 0, 35 (8.2%) at CSF stage 1, 53 (12.4%) at CSF stage 2, 49 (11.5%) at CSF stage 3, 87 (20.4%) at CSF stage 4 and 78 (18.3%) at CSF stage 5.

150 Demographic, genetic, and diagnostic characteristics 149 (11.5%) at CSF stage 3, 87 (20.4%) at CSF stage 4 and 78 (18.3%) at CSF stage 5.

150 Demographic, genetic, and diagnostic characteristics of these participants is shown in

151 Ext Data Fig. 2. In brief, the CSF biom 150 Demographic, genetic, and diagnostic characteristics of these participants is shown in

151 Ext Data Fig. 2. In brief, the CSF biomarker-based was not associated with sex
 $(\chi^2(5)=7.7, p=0.180)$ or years of education 151 Ext Data Fig. 2. In brief, the CSF biomarker-based was not associated with sex $(\chi^2(5)=7.7, p=0.180)$ or years of education $(\chi^2(5)=4.7, p=0.452)$, but higher CSF stage was associated with older age $(\chi^2(5)=16.9, p=0.005$ $(\chi^2(5)=7.7, p=0.180)$ or years of education (χ^2) 152 ($\chi^2(5)=7.7$, p=0.180) or years of education ($\chi^2(5)=4.7$, p=0.452), but higher CSF stage

153 was associated with older age ($\chi^2(5)=16.9$, p=0.005), carriership of an *APOE-ε4* allele

154 ($\chi^2(5)=72.8$, p<0.001) was associated with older age ($\chi^2(5)=16.9$, p=0.005), carriership of an APOE- ϵ 4 allele $(\chi^2(5)=72.8, \, p<0.001)$ and a more advanced clinical disease stage (χ^2)

was associated with older age $(\chi^2(5)=16.9, p=0.005)$, carriership of an *APOE-ε4* allele
 $(\chi^2(5)=72.8, p<0.001)$ and a more advanced clinical disease stage $(\chi^2(5)=478.6,$
 $p<0.001$, Ext Data Fig. 2A-E).

We then examine 154 ($\chi^2(5)=72.8$, p<0.001) and a more advanced clinical disease stage ($\chi^2(5)=478.6$,
155 p<0.001, Ext Data Fig. 2A-E).
156 We then examined the distribution of the CSF biomarkers included in the model by
157 CSF bioma p<0.001, Ext Data Fig. 2A-E).
156 We then examined the distrit
157 CSF biomarker stage. Individent
158 statistics for each biomarker
159 Supplementary Table 2. In sur
160 with higher CSF stages, althe 156 We then examined the distribution of the CSF biomarkers included in the model by CSF biomarker stage. Individual plots by CSF biomarker can be found in Fig. 1B and statistics for each biomarker and their differences b 157 CSF biomarker stage. Individual plots by CSF biomarker can be found in Fig. 1B and statistics for each biomarker and their differences by CSF stage can be found in Supplementary Table 2. In summary, degree of abnorma 158 statistics for each biomarker and their differences by CSF stage can be found in

159 Supplementary Table 2. In summary, degree of abnormality of all biomarkers increased

160 with higher CSF stages, although their tr 159 Supplementary Table 2. In summary, degree of abnormality of all biomarkers increased

160 with higher CSF stages, although their trajectories were different. CSF A β 42/40 and

161 pT205/T205 had a steep increase at 160 with higher CSF stages, although their trajectories were different. CSF Aβ42/40 and pT205/T205 had a steep increase at stage 1 and stage 3, respectively, and then continued increasing but in a lower degree. On the ot 161 pT205/T205 had a steep increase at stage 1 and stage 3, respectively, and then
162 continued increasing but in a lower degree. On the other hand, pT217/T217 (CSF
163 stage 2) and MTBR-tau243 (CSF stage 4) levels contin 162 continued increasing but in a lower degree. On the other hand, pT217/T217 (CSF
163 stage 2) and MTBR-tau243 (CSF stage 4) levels continued to increase at all
164 subsequent CSF stages in a similar degree after crossing 163 stage 2) and MTBR-tau243 (CSF stage 4) levels continued to increase at all
164 subsequent CSF stages in a similar degree after crossing the threshold for positivity.
165 Also, both pT217/T217 and MTBR-tau243 reached ve 164 subsequent CSF stages in a similar degree after crossing the threshold for positivity.

165 Also, both pT217/T217 and MTBR-tau243 reached very high levels compared to the

166 reference control group (z-score>10). Thes 165 Also, both pT217/T217 and MTBR-tau243 reached very high levels compared to the
166 reference control group (z-score>10). These different biomarker trajectories revealed
4 166 reference control group (z-score>10). These different biomarker trajectories revealed
4

167 that the included CSF biomarkers exhibit different behaviours across the disease
168 continuum, aside from the biomarker disease stage at which they become abnormal.
169 This is summarized in figure 1C, in which the sm continuum, aside from the biomarker disease stage at which they become abnormal.
169 This is summarized in figure 1C, in which the smoothed locally estimated scatterplot
170 smoothing (LOESS) regression of all CSF biomarke

169 This is summarized in figure 1C, in which the smoothed locally estimated scatterplot
170 smoothing (LOESS) regression of all CSF biomarkers are plotted.
171 Finally, we assessed the stability of our model using the lon smoothing (LOESS) regression of all CSF biomarkers are plotted.
171 Finally, we assessed the stability of our model using the longitud
172 mean (SD) of 2.1 (0.2) years (n=220, Supplementary Table 1). W
173 participants rem 171 Finally, we assessed the stability of our model using the longitudinal CSF data over a
172 mean (SD) of 2.1 (0.2) years (n=220, Supplementary Table 1). We observed that most
173 participants remained at the same stage mean (SD) of 2.1 (0.2) years (n=220, Supplementary Table 1). We observed that most

participants remained at the same stage (N=183, 83.2%) or progressed (n=29, 13.2%),

while only few regressed (n=8, 2.9%, Ext Fig. 3A-B).

173 participants remained at the same stage (N=183, 83.2%) or progressed (n=29, 13.2%),

174 while only few regressed (n=8, 2.9%, Ext Fig. 3A-B). Of those that progressed, most

175 (n=25, 86.2%) progressed only one CSF s 174 while only few regressed (n=8, 2.9%, Ext Fig. 3A-B). Of those that progressed, most

175 (n=25, 86.2%) progressed only one CSF stage during the two-year follow-up. This

176 indicates a high stability of our model over 175 (n=25, 86.2%) progressed only one CSF stage during the two-year follow-up. This

176 indicates a high stability of our model over time.

177 **CSF biomarker stages are associated with AD pathology, biomarkers and

179**

176 indicates a high stability of our model over time.

177 **CSF biomarker stages are associated wirely**

179 **cognition**

180 Next, we investigated the association betw

181 aggregates (A β -PET), insoluble tau aggregat 178
179
180
181
182 178 **CSF biomarker stages are associated with AD pathology, biomarkers and cognition**
179 **Next, we investigated the association between CSF stages and insoluble Aβ**
181 aggregates (Aβ-PET), insoluble tau aggregates (tau 179 **cognition**
180 Next, we
181 aggregates
182 thickness
183 composite
184 Cognitive 180 Next, we investigated the association between CSF stages and insoluble Aβ aggregates (Aβ-PET), insoluble tau aggregates (tau-PET), neurodegeneration (cortical thickness and CSF neurofilament light [NfL]) and cognitio 181 aggregates (Aβ-PET), insoluble tau aggregates (tau-PET), neurodegeneration (cortical
182 thickness and CSF neurofilament light [NfL]) and cognition, using a global cognitive
183 composite sensitive to early AD change 182 thickness and CSF neurofilament light [NfL]) and cognition, using a global cognitive

183 composite sensitive to early AD changes (modified version of preclinical Alzheimer's

184 Cognitive Composite [mPACC]¹³⁵, Fi 183 composite sensitive to early AD changes (modified version of preclinical Alzheimer's
184 Cognitive Composite [mPACC]¹³⁵, Fig. 2). The degree of biomarker abnormality
185 increased with higher CSF stages, although t Cognitive Composite $[mPACC]^{35}$, Fig. 2). The degree of biomarker abnormality 184 Cognitive Composite [mPACC]^{1,36}, Fig. 2). The degree of biomarker abnormality
185 increased with higher CSF stages, although the trajectories were different. For
186 instance, Aβ-PET was the first to start increasi 185 increased with higher CSF stages, although the trajectories were different. For
186 instance, Aβ-PET was the first to start increasing (already at CSF stage 1 [Aβ42/40
187 stage], Δz -score=1.12, p=0.032 compared t 186 instance, Aβ-PET was the first to start increasing (already at CSF stage 1 [Aβ42/40 stage], Δz-score=1.12, p=0.032 compared to the previous stage) and continued to increase until the CSF stage 5 (total-tau stage), whe 187 stage], Δz-score=1.12, p=0.032 compared to the previous stage) and continued to
188 increase until the CSF stage 5 (total-tau stage), where it reached a plateau. However, it
189 was at CSF stage 2 (pT217/T217) that th 188 increase until the CSF stage 5 (total-tau stage), where it reached a plateau. However, it

189 was at CSF stage 2 (pT217/T217) that the mean of Aβ-PET was above 1.96SD

190 (95%Cl) of CU- levels (mean[SD] z-score=4.01 189 was at CSF stage 2 (pT217/T217) that the mean of Aβ-PET was above 1.96SD
190 (95%Cl) of CU- levels (mean[SD] z-score=4.01[2.71]). Tau-PET was the next
191 biomarker to show a significant increase, emerging at CSF stag 190 (95%Cl) of CU- levels (mean[SD] z-score=4.01[2.71]). Tau-PET was the next

191 biomarker to show a significant increase, emerging at CSF stage 3 (pT205/T205 stage,

192 Δz -score=2.09, p<0.001), in which its mean wa 191 biomarker to show a significant increase, emerging at CSF stage 3 (pT205/T205 stage,

192 Δz -score=2.09, p<0.001), in which its mean was already above 1.96SD of CU-

(mean[SD] z -score=2.02[3.16]). Notably, tau-PE Δz-score=2.09, p<0.001), in which its mean was already above 1.96SD of CU-

(mean[SD] z-score=2.02[3.16]). Notably, tau-PET levels continued to significantly

increase in all subsequent CSF stages. The mPACC followed a si (mean[SD] z-score=2.02[3.16]). Notably, tau-PET levels continued to significantly

194 increase in all subsequent CSF stages. The mPACC followed a similar trajectory,

195 starting to increase also at CSF stage 3 (pT205/T2 194 increase in all subsequent CSF stages. The mPACC followed a similar trajectory,

195 starting to increase also at CSF stage 3 (pT205/T205 stage, Δz -score=0.92, p=0.004),

196 and crossing the 1.96SD threshold at th 195 starting to increase also at CSF stage 3 (pT205/T205 stage, Δz-score=0.92, p=0.004),

196 and crossing the 1.96SD threshold at the following stage (mean[SD] z-

197 score=2.60[1.84]). Finally, both measures of neurode 200 and crossing the 1.96SD threshold at the following stage (mean[SD] z-

200 score=2.60[1.84]). Finally, both measures of neurodegeneration (cortical thickness and

200 CSF NfL) showed a significantly lower degree of ab 197 score=2.60[1.84]). Finally, both measures of neurodegeneration (cortical thickness and
198 CSF NfL) showed a significantly lower degree of abnormality compared to the other
199 biomarkers, even in the most advanced CSF 198 CSF NfL) showed a significantly lower degree of abnormality compared to the other

199 biomarkers, even in the most advanced CSF stages. The abnormality of cortical

100 thickness significantly increased at CSF stage 3 199 biomarkers, even in the most advanced CSF stages. The abnormality of cortical

200 thickness significantly increased at CSF stage 3 (pT205/T205 stage, Δz -score=0.80,

201 p=0.006), but it was above the 1.96SD (95%C 200 thickness significantly increased at CSF stage 3 (pT205/T205 stage, Δz-score=0.80,

201 p=0.006), but it was above the 1.96SD (95%Cl) threshold only at the last CSF stage

202 (total-tau stage, mean[SD] z-score=2.07[1 201 p=0.006), but it was above the 1.96SD (95%CI) threshold only at the last CSF stage
202 (total-tau stage, mean[SD] z-score=2.07[1.91]). CSF NfL, on the other hand, only
5 202 (total-tau stage, mean[SD] z-score=2.07[1.91]). CSF NfL, on the other hand, only

203 showed significant differences between CSF stages 3 and 4 (Δz -score=0.36, p=0.016),

204 but did not cross the 1.96SD threshold at any CSF stage. Statistics of each these AD

205 biomarkers and their differences pe

204 but did not cross the 1.96SD threshold at any CSF stage. Statistics of each these AD
205 biomarkers and their differences per CSF stage can be found in Supplementary Table
206 3.
207 We further studied the associations 205 biomarkers and their differences per CSF stage can be found in Supplementary Table

206 3.

207 We further studied the associations between our CSF-based staging model and other

208 biomarkers as additional analyses. 206 3.

207 We

208 bic

209 bra

210 sta

211 be 207 We further studied the associations between our CSF-based staging model and other

208 biomarkers as additional analyses. For tau-PET, we quantified the signal in different

209 brain regions, using the previously vali 208 biomarkers as additional analyses. For tau-PET, we quantified the signal in different

209 brain regions, using the previously validated ROIs reflecting the different Braak

210 stages⁴³. Early (Braak I) and intermed 209 brain regions, using the previously validated ROIs reflecting the different Braak

210 stages⁴³. Early (Braak I) and intermediate (Braak III-IV) regions of tau deposition

211 become abnormal at CSF stage 3 (pT205/T2 stages⁴³. Early (Braak I) and intermediate (Braak III-IV) regions of tau deposition 210 stages⁴³. Early (Braak I) and intermediate (Braak III-IV) regions of tau deposition
211 become abnormal at CSF stage 3 (pT205/T205 stage, Braak I: mean z-score: 2.26;
212 Braak III-IV: mean z-score: 1.97) and later r 211 become abnormal at CSF stage 3 (pT205/T205 stage, Braak I: mean z-score: 2.26;

212 Braak III-IV: mean z-score: 1.97) and later regions (Braak V-VI) become abnormal at

213 the CSF stage 4 (MTBR-tau243 stage, mean z-sc 212 Braak III-IV: mean z-score: 1.97) and later regions (Braak V-VI) become abnormal at
213 the CSF stage 4 (MTBR-tau243 stage, mean z-score: 4.58; Ext Data. Fig. 4 and
214 Supplementary Table 4). Noteworthy, the number of 213 the CSF stage 4 (MTBR-tau243 stage, mean z-score: 4.58; Ext Data. Fig. 4 and
214 Supplementary Table 4). Noteworthy, the number of participants with Braak I z-scores
215 over 1.96SD at CSF stage 3 was higher than that

214 Supplementary Table 4). Noteworthy, the number of participants with Braak I z-scores

215 over 1.96SD at CSF stage 3 was higher than that with Braak III-IV (n=21, 42.9% vs.

216 n=14, 28.6%).

217 We also examined dif over 1.96SD at CSF stage 3 was higher than that with Braak III-IV (n=21, 42.9% vs.

216 n=14, 28.6%).

217 We also examined different measures of cognitive function including composites for

218 memory, executive, language 216 n=14, 28.6%).
217 We also exam
218 memory, exed
219 composite to c
220 2.18), and kep
221 Table 5). This 217 We also examined different measures of cognitive function including composites for

218 memory, executive, language and visuospatial functions, respectively. The first

219 composite to cross the 1.96SD threshold was m 218 memory, executive, language and visuospatial functions, respectively. The first
219 composite to cross the 1.96SD threshold was memory at CSF stage 4 (mean z-score:
210 2.18), and kept increasing in the next CSF stage 219 composite to cross the 1.96SD threshold was memory at CSF stage 4 (mean z-score:
220 2.18), and kept increasing in the next CSF stage (Ext. Data Fig. 5 and Supplementary
221 Table 5). This was followed by the executive 220 2.18), and kept increasing in the next CSF stage (Ext. Data Fig. 5 and Supplementary

221 Table 5). This was followed by the executive function composite, which was above the

222 the shold at CSF stage 5 (t-tau stage, 221 Table 5). This was followed by the executive function composite, which was above the
222 threshold at CSF stage 5 (t-tau stage, mean z-score: 2.15). Finally, the language and
223 visuospatial cognitive composites showe 222 threshold at CSF stage 5 (t-tau stage, mean z-score: 2.15). Finally, the language and
223 visuospatial cognitive composites showed an increasing degree of abnormality across
224 CSF stages but did not cross the 1.96SD

223 visuospatial cognitive composites showed an increasing degree of abnormality across
224 CSF stages but did not cross the 1.96SD threshold at any CSF stage.
225 CSF biomarker stages can be used for predicting A/T statu 224 CSF stages but did not cross the 1.96SD threshold at any CSF stage.
225 **CSF biomarker stages can be used for predicting A/T status and c**
227 Subsequently, we looked at the accuracy of our CSF staging model
228 (A) a ---
226
227
228
229
230 **CSF biomarker stages can be used for predicting A/T status and cognitive stages**
227 Subsequently, we looked at the accuracy of our CSF staging model for predicting Aβ
228 (A) and tau (T) status, as defined by $PET³⁴$ 227 Subsequently, we looked at the accuracy of our CSF staging model for predicting Aβ

228 (A) and tau (T) status, as defined by PET³⁴. We first looked at each independent

239 pathology dichotomously (*i.e.*, positive (A) and tau (T) status, as defined by $PET³⁴$. We first looked at each independent 228 (A) and tau (T) status, as defined by PET^{34} . We first looked at each independent

229 pathology dichotomously (*i.e.*, positive or negative) and independently and, later, we

230 looked at the ordinal categories me 229 pathology dichotomously (*i.e.*, positive or negative) and independently and, later, we

230 looked at the ordinal categories merging both pathologies (*i.e.,* A-T-, A+T-, and A+T+).

231 The number of positive partic 230 looked at the ordinal categories merging both pathologies (*i.e.,* A-T-, A+T-, and A+T+).

231 The number of positive participants by CSF stage and category are presented in Fig.

232 3A. Using receiver operating char 231 The number of positive participants by CSF stage and category are presented in Fig.

232 3A. Using receiver operating characteristic (ROC) curves analyses, we determined that

233 CSF stage 2 (pT217/T217 stage) was th 232 3A. Using receiver operating characteristic (ROC) curves analyses, we determined that

233 CSF stage 2 (pT217/T217 stage) was the optimal threshold for predicting AB-PET

234 positivity with high accuracy (area under 233 CSF stage 2 (pT217/T217 stage) was the optimal threshold for predicting Aβ-PET

234 positivity with high accuracy (area under the curve and 95% confidence interval

235 [AUC[95%CI]]=0.96[0.93, 0.98], sensitivity=0.93 234 positivity with high accuracy (area under the curve and 95% confidence interval

235 [AUC[95%Cl]]=0.96[0.93, 0.98], sensitivity=0.93 and specificity=0.89, first column Fig.

236 3B and Supplementary Table 6). Tau-PET p 235 [AUC[95%CI]]=0.96[0.93, 0.98], sensitivity=0.93 and specificity=0.89, first column Fig.

236 3B and Supplementary Table 6). Tau-PET positivity was also predicted with high

237 accuracy when using CSF stage 4 (MTBR-tau 236 3B and Supplementary Table 6). Tau-PET positivity was also predicted with high
237 accuracy when using CSF stage 4 (MTBR-tau243 stage) as a threshold
238 (AUC[95%CI]=0.95[0.93, 0.97], sensitivity=0.91 and specificity=0 237 accuracy when using CSF stage 4 (MTBR-tau243 stage) as a threshold

238 (AUC[95%CI]=0.95[0.93, 0.97], sensitivity=0.91 and specificity=0.92, second column

239 Fig. 3B and Supplementary Table 6). 238 (AUC[95%CI]=0.95[0.93, 0.97], sensitivity=0.91 and specificity=0.92, second column
239 Fig. 3B and Supplementary Table 6).
6 239 Fig. 3B and Supplementary Table 6).

The Supplementary Table 6).

The Supplementary Table 6).

240 Ordinal logistic regression was used to assess the utility of CSF stages for predicting

241 A/T status (*i.e.*, A-T-, A+T- or A+T+), and we calculated the c-index (an overall

242 measure of discrimination equivalent 241 A/T status (*i.e.*, A-T-, A+T- or A+T+), and we calculated the c-index (an overall

242 measure of discrimination equivalent to AUC for dichotomic outcomes) as a measure of

243 accuracy. We observed that higher CSF s 242 measure of discrimination equivalent to AUC for dichotomic outcomes) as a measure of

243 accuracy. We observed that higher CSF stages were associated with higher predicted

244 probabilities of being at more advanced 243 accuracy. We observed that higher CSF stages were associated with higher predicted

244 probabilities of being at more advanced A/T status (c-index[95%Cl]=0.95[0.93, 0.97],

245 last column Fig. 3B, and Supplementary 244 probabilities of being at more advanced A/T status (c-index[95%CI]=0.95[0.93, 0.97],

245 last column Fig. 3B, and Supplementary Table 6). More specifically, participants at

246 CSF stages 0 and 1 (negative biomarker 245 last column Fig. 3B, and Supplementary Table 6). More specifically, participants at

246 CSF stages 0 and 1 (negative biomarkers and A β 42/40 stages) had the highest

247 probability of being A-T-, at CSF stages 2 a 246 CSF stages 0 and 1 (negative biomarkers and Aβ42/40 stages) had the highest
247 probability of being A-T-, at CSF stages 2 and 3 (pT217/T217 and pT205/T205 stages)
248 being A+T- and at CSF stages 4 and 5 (MTBR-tau24 probability of being A-T-, at CSF stages 2 and 3 (pT217/T217 and pT205/T205 stages)

248 being A+T- and at CSF stages 4 and 5 (MTBR-tau243 and total-tau stages) of being

249 A+T+. Only one participant was classified as A-

248 being A+T- and at CSF stages 4 and 5 (MTBR-tau243 and total-tau stages) of being
249 A+T+. Only one participant was classified as A-T+, which was excluded from this
250 analysis.
251 Finally, we also aimed at investiga 249 A+T+. Only one participant was classified as A-T+, which was excluded from this
250 analysis.
251 Finally, we also aimed at investigating whether our staging model could be used as a
252 diagnostic tool. In the first a 250 analysis.

251 Finally, w

252 diagnostic

253 cognitive

254 were ass

255 0.91], firs 251 Finally, we also aimed at investigating whether our staging model could be used as a
252 diagnostic tool. In the first analysis, we used the CSF staging model for predicting
253 cognitive stages within the AD continuum 252 diagnostic tool. In the first analysis, we used the CSF staging model for predicting

253 cognitive stages within the AD continuum (*i.e.,* excluding non-AD). Higher CSF stages

254 were associated with more advanced 253 cognitive stages within the AD continuum (*i.e.,* excluding non-AD). Higher CSF stages
254 were associated with more advanced cognitive stages (c-index[95%Cl]=0.88[0.86,
255 0.91], first column Fig. 3C and Supplementa 254 were associated with more advanced cognitive stages (c-index[95%CI]=0.88[0.86,

255 0.91], first column Fig. 3C and Supplementary Table 6). The model predicted that

256 participants at CSF stage 0 (negative biomarker 255 0.91], first column Fig. 3C and Supplementary Table 6). The model predicted that

256 participants at CSF stage 0 (negative biomarkers stage) had the highest probability of

257 being CU-; at CSF stages 1 and 2 (A β 256 participants at CSF stage 0 (negative biomarkers stage) had the highest probability of
257 being CU-; at CSF stages 1 and 2 (A β 42/40 and pT217/T217 stages) were more
258 probably CU+ (as assessed by CSF), at CSF st 257 being CU-; at CSF stages 1 and 2 (Aβ42/40 and pT217/T217 stages) were more

258 probably CU+ (as assessed by CSF), at CSF stage 3 (pT205/T205) MCI+ and, finally,

259 at CSF stages 4 and 5 (MTBR-tau243 and total-tau) 258 probably CU+ (as assessed by CSF), at CSF stage 3 (pT205/T205) MCI+ and, finally, at CSF stages 4 and 5 (MTBR-tau243 and total-tau) ADD+. Lastly, we aimed at differentiating cognitive impairment due to AD or due to oth 259 at CSF stages 4 and 5 (MTBR-tau243 and total-tau) ADD+. Lastly, we aimed at differentiating cognitive impairment due to AD or due to other neurodegenerative diseases. We therefore compared patients with AD to patients 260 differentiating cognitive impairment due to AD or due to other neurodegenerative
261 diseases. We therefore compared patients with AD to patients with non-AD dementia,
262 only including those with objective cognitive 261 diseases. We therefore compared patients with AD to patients with non-AD dementia,

262 only including those with objective cognitive impairment *(i.e., MCI* and dementia

263 patients). Participants at CSF stage 2 (pT 262 only including those with objective cognitive impairment (*i.e.,* MCI and dementia

263 patients). Participants at CSF stage 2 (pT217/T217 stage) or higher with objective

264 cognitive impairment had a high probabilit 263 patients). Participants at CSF stage 2 (pT217/T217 stage) or higher with objective

264 cognitive impairment had a high probability of having AD as the cause of their cognitive

265 impairment (AUC[95%CI]=0.95[0.93, 0. 264 cognitive impairment had a high probability of having AD as the cause of their cognitive

265 impairment (AUC[95%CI]=0.95[0.93, 0.98], sensitivity=0.97 and specificity=0.75, last

266 column Fig. 3C-D and Supplementary

265 impairment (AUC[95%CI]=0.95[0.93, 0.98], sensitivity=0.97 and specificity=0.75, last
266 column Fig. 3C-D and Supplementary Table 6).
267 **Longitudinal rates of change of AD biomarkers differ by CSF stages**
269 Next, w 266 column Fig. 3C-D and Supplementary Table 6).

267 **Longitudinal rates of change of AD biomarke**

269 Next, we used longitudinal imaging and cogniti

270 change over time based on the baseline CS

271 Table 7). The ra 268
269
270
271
272 **Longitudinal rates of change of AD biomarkers differ by CSF stages**
269 Next, we used longitudinal imaging and cognitive data to assess how A
270 change over time based on the baseline CSF stage classification (S
271 Tab 269 Next, we used longitudinal imaging and cognitive data to assess how AD biomarkers

270 change over time based on the baseline CSF stage classification (Supplementary

271 Table 7). The rate of accumulation of A β ag 270 change over time based on the baseline CSF stage classification (Supplementary

271 Table 7). The rate of accumulation of A β aggregates as measured with PET (n=218)

272 increased at early CSF stages reaching the h 271 Table 7). The rate of accumulation of Aβ aggregates as measured with PET (n=218)

272 increased at early CSF stages reaching the highest values at CSF stage 2

273 (pT217/T217 stage) and thereafter the rate decreased 272 increased at early CSF stages reaching the highest values at CSF stage 2

273 (pT217/T217 stage) and thereafter the rate decreased but still remained positive (Fig. 4

274 and Supplementary Table 8). On the other hand, 273 (pT217/T217 stage) and thereafter the rate decreased but still remained positive (Fig. 4
274 and Supplementary Table 8). On the other hand, the tau-PET (n=312), cortical
275 thickness (n=300) and mPACC (n=342) exhibit 274 and Supplementary Table 8). On the other hand, the tau-PET (n=312), cortical

275 thickness (n=300) and mPACC (n=342) exhibited monotonic increases in rates of

276 change over time, with the rates starting to be sign 275 thickness (n=300) and mPACC (n=342) exhibited monotonic increases in rates of
276 change over time, with the rates starting to be significantly different from contiguous
 7 276 change over time, with the rates starting to be significantly different from contiguous
7

277 CSF stages at CSF stage 3 (pT205/T205 stage; Fig. 4). Figure 4B depicts that tau-

278 PET, followed by mPACC had the highest rate of change (z-scored), while Aβ-PET and

279 cortical thickness had lower rates of cha 278 PET, followed by mPACC had the highest rate of change (z-scored), while Aβ-PET and
279 cortical thickness had lower rates of change that were in a similar range.
280 **CSF biomarker stages predict clinical progression** cortical thickness had lower rates of change that were in a similar range.

280
 CSF biomarker stages predict clinical progression

282 In the next set of analyses, we tested whether our CSF staging model

283 predicting 281
282
283
284
285 **CSF biomarker stages predict clinical progression**

282 In the next set of analyses, we tested whether our C:

283 predicting subsequent clinical progression (up to 5-yea

284 visit). First, we tested the ability of our m 282 In the next set of analyses, we tested whether our CSF staging model was useful for predicting subsequent clinical progression (up to 5-years of follow-up after the baseline visit). First, we tested the ability of our predicting subsequent clinical progression (up to 5-years of follow-up after the baseline

284 visit). First, we tested the ability of our model to predict progression to AD dementia

285 from CU or MCI status at baseline 284 visit). First, we tested the ability of our model to predict progression to AD dementia

285 from CU or MCI status at baseline (progressors: n=41). Based on Kaplan-Meier

286 analyses (Fig. 5A), participants at higher 285 from CU or MCI status at baseline (progressors: n=41). Based on Kaplan-Meier

286 analyses (Fig. 5A), participants at higher CSF stages (4-5; MTBR-tau243 and total-tau

287 stages) at baseline had higher probability t 286 analyses (Fig. 5A), participants at higher CSF stages (4-5; MTBR-tau243 and total-tau

287 stages) at baseline had higher probability to progress to AD dementia, than those at

288 positive lower CSF stages (*i.e.*, 1 287 stages) at baseline had higher probability to progress to AD dementia, than those at positive lower CSF stages (*i.e.*, 1-3). When adjusting for age, sex, and clinical status at baseline (*i.e.*, CU or MCl), the hazar 288 positive lower CSF stages (*i.e.*, 1-3). When adjusting for age, sex, and clinical status at
289 baseline (*i.e.*, CU or MCl), the hazard ratio (HR) was 5.2 (95%Cl: [2.2, 12.6], p<0.001)
290 when comparing participant 289 baseline (*i.e.*, CU or MCI), the hazard ratio (HR) was 5.2 (95%CI: [2.2, 12.6], p<0.001)

290 when comparing participants at CSF stages 4 or 5 to participants at lower, but positive,

291 CSF stages (1-3, Aβ42/40 to 290 when comparing participants at CSF stages 4 or 5 to participants at lower, but positive,

291 CSF stages (1-3, A β 42/40 to pT205/T205, reference; Fig. 5B and Supplementary Table

292 9). When including only those wi 291 CSF stages (1-3, Aβ42/40 to pT205/T205, reference; Fig. 5B and Supplementary Table 292 9). When including only those with MCI at baseline (progressors: 38/88), we still found that those at CSF stages 4 or 5 at baseli 292 9). When including only those with MCI at baseline (progressors: 38/88), we still found

293 that those at CSF stages 4 or 5 at baseline had a significantly higher probability to

294 progress to AD dementia (HR[95%CI 293 that those at CSF stages 4 or 5 at baseline had a significantly higher probability to

294 progress to AD dementia (HR[95%Cl]=4.5[1.8, 10.8], p<0.001, Fig. 5C-D and

295 Supplementary Table 9). Finally, we investigate 294 progress to AD dementia (HR[95%CI]=4.5[1.8, 10.8], p<0.001, Fig. 5C-D and
295 Supplementary Table 9). Finally, we investigated the utility of the CSF staging model
296 when predicting progression from CU to MCI status 295 Supplementary Table 9). Finally, we investigated the utility of the CSF staging model
296 when predicting progression from CU to MCI status (progressors: 11/159). Again, those
297 CU participants at higher CSF stages (when predicting progression from CU to MCI status (progressors: 11/159). Again, those

297 CU participants at higher CSF stages (4-5) at baseline, were much more prone to

298 progress to MCI with a HR of 16.0 (95%Cl: 3.2, 297 CU participants at higher CSF stages (4-5) at baseline, were much more prone to progress to MCI with a HR of 16.0 (95%Cl: 3.2, 81.1, p<0.001, Fig. 5E-F and Supplementary Table 9) compared to those in stage 1-3, support 298 progress to MCI with a HR of 16.0 (95%CI: 3.2, 81.1, p<0.001, Fig. 5E-F and
299 Supplementary Table 9) compared to those in stage 1-3, supporting the clinical utility of
300 the proposed staging model. There were no pr 299 Supplementary Table 9) compared to those in stage 1-3, supporting the clinical utility of

200 the proposed staging model. There were no progressors from CSF stage 0 in any case,

201 which prevented us from comparing 300 the proposed staging model. There were no progressors from CSF stage 0 in any case,
301 which prevented us from comparing these participants with the other CSF stages
302 groups. Kaplan-Meier curves for each individual 301 which prevented us from comparing these participants with the other CSF stages
302 groups. Kaplan-Meier curves for each individual CSF stage are depicted in Ext Data
303 Fig. 6.
Replication in an independent cohort
3

302 groups. Kaplan-Meier curves for each individual CSF stage are depicted in Ext Data
303 Fig. 6.
Replication in an independent cohort
305 **Replication in an independent cohort**
306 Finally, we replicated the staging mo 303 Fig. 6.

304

305 **Replic**

305 Finally,

307 Joanne

308 Table : 305
306
307
308
309 **Replication in an independent cohort**
306 Finally, we replicated the staging model
307 Joanne Knight Alzheimer's Disease Re
308 Table 2). SuStaln selected one unique
309 CSF abnormality ordering as the one pri
310 this co 306 Finally, we replicated the staging model and the main analyses in the Charles F. and
307 Joanne Knight Alzheimer's Disease Research Center (Knight-ADRC) cohort (n=222,
308 Table 2). SuStaln selected one unique subtype 307 Joanne Knight Alzheimer's Disease Research Center (Knight-ADRC) cohort (n=222,
308 Table 2). SuStaln selected one unique subtype as the optimal model with the same
309 CSF abnormality ordering as the one previously ob 308 Table 2). SuStaln selected one unique subtype as the optimal model with the same
309 CSF abnormality ordering as the one previously obtained in BioFINDER-2 (Fig. 6A). In
310 this cohort, however, there was slightly hi 309 CSF abnormality ordering as the one previously obtained in BioFINDER-2 (Fig. 6A). In
310 this cohort, however, there was slightly higher uncertainty between the ordering of first
311 two (Aβ42/40 and pT217/T217) and 310 this cohort, however, there was slightly higher uncertainty between the ordering of first
311 two (A β 42/40 and pT217/T217) and the last two (MTBR-tau243 and total-tau) stages.
312 These differences may be mostly du 311 two (A β 42/40 and pT217/T217) and the last two (MTBR-tau243 and total-tau) stages.
312 These differences may be mostly due to the difference in sample size, especially in
313 more advanced AD cases (only 9 mild AD d 312 These differences may be mostly due to the difference in sample size, especially in
313 more advanced AD cases (only 9 mild AD dementia cases). Nonetheless, the overall
8 313 more advanced AD cases (only 9 mild AD dementia cases). Nonetheless, the overall

It is made available under a [CC-BY-NC-ND 4.0 International license](http://creativecommons.org/licenses/by-nc-nd/4.0/) . **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. medRxiv preprint doi: [https://doi.org/10.1101/2023.07.14.23292650;](https://doi.org/10.1101/2023.07.14.23292650) this version posted July 16, 2023. The copyright holder for this preprint

314 behaviour of these CSF biomarkers by the biomarker stages was similar to that in the
315 main cohort (Fig. 6B and Supplementary Table 2). Further, the other AD biomarkers
316 available (not included in the CSF staging 315 main cohort (Fig. 6B and Supplementary Table 2). Further, the other AD biomarkers
316 available (not included in the CSF staging model), showed similar trajectories to those
317 in the discovery (BioFINDER-2) sample. 316 available (not included in the CSF staging model), showed similar trajectories to those
317 in the discovery (BioFINDER-2) sample. Aβ-PET was once again the first modality to
318 cross the 1.96SD threshold at CSF sta 317 in the discovery (BioFINDER-2) sample. Aβ-PET was once again the first modality to
318 cross the 1.96SD threshold at CSF stage 1 (Aβ42/Ab40 stage), followed by tau-PET in
319 CSF stage 3 (pT205/T205 stage) and, finall 318 cross the 1.96SD threshold at CSF stage 1 (A β 42/Ab40 stage), followed by tau-PET in
319 CSF stage 3 (pT205/T205 stage) and, finally, CSF NfL crossing this threshold at the
320 final stage (total-tau stage; Fig. 6C 319 CSF stage 3 (pT205/T205 stage) and, finally, CSF NfL crossing this threshold at the
320 final stage (total-tau stage; Fig. 6C and Supplementary Table 3). Neither the global
321 cognitive composite nor AD cortical thick 320 final stage (total-tau stage; Fig. 6C and Supplementary Table 3). Neither the global
321 cognitive composite nor AD cortical thickness signature crossed this threshold. The
322 main difference compared with BioFINDER-2 321 cognitive composite nor AD cortical thickness signature crossed this threshold. The
322 main difference compared with BioFINDER-2, was the lower degree of abnormality for
323 all markers in the last CSF stages. This mi 322 main difference compared with BioFINDER-2, was the lower degree of abnormality for
323 all markers in the last CSF stages. This might be explained by the lower number of
324 advanced patient cases in this cohort. The i 323 all markers in the last CSF stages. This might be explained by the lower number of advanced patient cases in this cohort. The individual plots for each CSF and imaging biomarker by CSF stages are shown in Ext Data Fig. 324 advanced patient cases in this cohort. The individual plots for each CSF and imaging
325 biomarker by CSF stages are shown in Ext Data Fig. 7. Details of participants
326 characteristics (Ext Data Fig. 2), tau-PET bind 325 biomarker by CSF stages are shown in Ext Data Fig. 7. Details of participants
326 characteristics (Ext Data Fig. 2), tau-PET binding in different regions (Ext Data Figure 4
327 and Supplementary Table 4) and other cog 326 characteristics (Ext Data Fig. 2), tau-PET binding in different regions (Ext Data Figure 4
327 and Supplementary Table 4) and other cognitive measures (Ext Data Figure 5 and
328 Supplementary Table 5) per CSF stage can 327 and Supplementary Table 4) and other cognitive measures (Ext Data Figure 5 and
328 Supplementary Table 5) per CSF stage can be found in the Ext Data. Stability
329 analyses, within participants with available longitud 328 Supplementary Table 5) per CSF stage can be found in the Ext Data. Stability
329 analyses, within participants with available longitudinal CSF measures (n=51,
330 Supplementary Table 10), also showed that most partic

329 analyses, within participants with available longitudinal CSF measures (n=51,
330 Supplementary Table 10), also showed that most participants remained at the same
331 stage (n=46, 90.2%) or progressed (n=4, 7.8%) at 330 Supplementary Table 10), also showed that most participants remained at the same
331 stage (n=46, 90.2%) or progressed (n=4, 7.8%) at follow-up (Ext Data Fig. 3C-D).
332 We also calculated the optimal CSF stages for 331 stage (n=46, 90.2%) or progressed (n=4, 7.8%) at follow-up (Ext Data Fig. 3C-D).
332 We also calculated the optimal CSF stages for predicting Aβ-PET and tau
333 positivity using ROC curves. As in the case of BioFINDE 332 We also calculated the optimal CSF stages for predicting Aβ-PET and tau-PET

positivity using ROC curves. As in the case of BioFINDER-2, CSF stage 2 (pT217/T217

stage) was optimal for predicting Aβ-PET positivity (A 333 positivity using ROC curves. As in the case of BioFINDER-2, CSF stage 2 (pT217/T217
334 stage) was optimal for predicting Aβ-PET positivity (AUC[95%CI]=0.89[0.85, 0.94], Fig.
6D&G and Supplementary Table 6), whereas C 334 stage) was optimal for predicting Aβ-PET positivity (AUC[95%CI]=0.89[0.85, 0.94], Fig.

335 6D&G and Supplementary Table 6), whereas CSF stage 4 (MTBR-tau243 stage) was

336 optimal for predicting tau-PET positivity (335 6D&G and Supplementary Table 6), whereas CSF stage 4 (MTBR-tau243 stage) was
336 optimal for predicting tau-PET positivity (AUC[95%CI]=0.94[0.91, 0.96], Fig. 6E&H).
337 Consistent with findings in BioFINDER-2, higher 336 optimal for predicting tau-PET positivity (AUC[95%CI]=0.94[0.91, 0.96], Fig. 6E&H).
337 Consistent with findings in BioFINDER-2, higher CSF stages were predictive of more
338 advanced A/T stages, as assessed by PET (c 337 Consistent with findings in BioFINDER-2, higher CSF stages were predictive of more
338 advanced A/T stages, as assessed by PET (c-index[95%CI]=0.89[0.86, 0.92], Fig. 6
539 F&I, Supplementary Table 6). Being at CSF sta 338 advanced A/T stages, as assessed by PET (c-index[95%CI]=0.89[0.86, 0.92], Fig. 6
339 F&I, Supplementary Table 6). Being at CSF stages 0 and 1 (biomarker negative and
40 AB42/40 stage) was highly predictive of being A-339 F&I, Supplementary Table 6). Being at CSF stages 0 and 1 (biomarker negative and Aβ42/40 stage) was highly predictive of being A-T-, being at CSF stages 2 and 3 (pT217/T217 and pT205/T205 stages) was predictive of b

340 Aβ42/40 stage) was highly predictive of being A-T-, being at CSF stages 2 and 3 (pT217/T217 and pT205/T205 stages) was predictive of being A+T- and, at CSF stages 4 and 5 (MTBR-tau243 and total-tau stages) of being 341 (pT217/T217 and pT205/T205 stages) was predictive of being A+T- and, at CSF
342 stages 4 and 5 (MTBR-tau243 and total-tau stages) of being A+T+.
343 Finally, we investigated the prognostic capacity of our model for pr stages 4 and 5 (MTBR-tau243 and total-tau stages) of being A+T+.

343 Finally, we investigated the prognostic capacity of our model for pre

344 to CDR≥1 (AD dementia, progressors: 41/218) and CDR≥0.5 (M

345 dementia, pr 343 Finally, we investigated the prognostic capacity of our model for predicting progression
344 to CDR≥1 (AD dementia, progressors: 41/218) and CDR≥0.5 (MCI or very mild AD
345 dementia, progressors: 30/214). We found th 344 to CDR≥1 (AD dementia, progressors: 41/218) and CDR≥0.5 (MCI or very mild AD
345 dementia, progressors: 30/214). We found that CU (CDR=0) and very mild AD
346 (CDR=0.5) participants at the highest CSF stages (4-5; MT 345 dementia, progressors: 30/214). We found that CU (CDR=0) and very mild AD (CDR=0.5) participants at the highest CSF stages (4-5; MTBR-tau243 and total-tau stages) exhibited an increased risk (HR[95%CI]=6.9[3.0, 16.0<mark>]</mark> 346 (CDR=0.5) participants at the highest CSF stages (4-5; MTBR-tau243 and total-tau
347 stages) exhibited an increased risk (HR[95%CI]=6.9[3.0, 16.0<mark>],</mark> p<0.001) of progressing
348 to AD dementia (CDR≥1) at follow-up, ev 347 stages) exhibited an increased risk (HR[95%CI]=6.9[3.0, 16.0<mark>],</mark> p<0.001) of progressing
348 to AD dementia (CDR≥1) at follow-up, even when adjusting for age, sex and clinical
349 status (*i.e.*, CDR=0 or CDR=0.5) at 348 to AD dementia (CDR≥1) at follow-up, even when adjusting for age, sex and clinical
349 status (*i.e.*, CDR=0 or CDR=0.5) at baseline, compared to participants at CSF stages
9 349 status (*i.e.*, CDR=0 or CDR=0.5) at baseline, compared to participants at CSF stages

350 1-3 (Fig. 6J-K and Supplementary Table 9). Similarly, CU participants at higher CSF stages (*i.e.*, 4-5) had higher risk (HR[95%Cl]=4.2[2.0, 8.8], p<0.001) of progressing to very mild AD or more advanced disease stage 351 stages (*i.e.*, 4-5) had higher risk (HR[95%CI]=4.2[2.0, 8.8], p<0.001) of progressing to very mild AD or more advanced disease stages when compared to participants at lower, but positive, CSF stages (1-3, Fig. 6L-M a 352 very mild AD or more advanced disease stages when compared to participants at lower, but positive, CSF stages (1-3, Fig. 6L-M and Supplementary Table 9). In this case, participants at CSF stages 1-3 also showed signif 353 lower, but positive, CSF stages (1-3, Fig. 6L-M and Supplementary Table 9). In this case, participants at CSF stages 1-3 also showed significant higher risk to progress to CDR≥0.5 than those at CSF stage 0 (HR[95%CI]= 354 case, participants at CSF stages 1-3 also showed significant higher risk to progress to

355 CDR≥0.5 than those at CSF stage 0 (HR[95%CI]=5.0[1.6, 15.0], p=0.005). There were

356 no progressors to CDR≥1 at CSF stage 355 CDR≥0.5 than those at CSF stage 0 (HR[95%CI]=5.0[1.6, 15.0], p=0.005). There were

no progressors to CDR≥1 at CSF stage 0, which prevented us to compare this group to

the others. Kaplan-Meier curves for each individu 356 no progressors to CDR≥1 at CSF stage 0, which prevented us to compare this group to
357 the others. Kaplan-Meier curves for each individual CSF stage are depicted in Ext Data
359 **Discussion**
360 **Discussion**
361 In t

357 the others. Kaplan-Meier curves for each individual CSF stage are depicted in Ext Data
358 Fig. 6.
359 **Discussion**
161 In this study, we created and evaluated a staging model for AD using five CSF
362 biomarkers refl 358 Fig. 6.
359
360 **Discu**:
361 In this
362 biomar
363 have o 360
361
362
363
364 360 **Discussion**
361 In this study,
362 biomarkers re
363 have demons
364 participants r
365 biomarker ab 361 In this study, we created and evaluated a staging model for AD using five CSF
362 biomarkers reflecting abnormalities of soluble A β and different soluble tau species. We
363 have demonstrated that a single CSF coll 362 biomarkers reflecting abnormalities of soluble Aβ and different soluble tau species. We
363 have demonstrated that a single CSF collection is sufficient to accurately stage
364 participants representing the entire AD 363 have demonstrated that a single CSF collection is sufficient to accurately stage
364 participants representing the entire AD continuum. This is possible because CSF
365 biomarker abnormalities followed a stereotypical 364 participants representing the entire AD *continuum*. This is possible because CSF
365 biomarker abnormalities followed a stereotypical trajectory in all participants, which
366 enabled a single staging model usable fo 365 biomarker abnormalities followed a stereotypical trajectory in all participants, which
366 enabled a single staging model usable for everyone. Notably, we have been able to
367 relate the CSF stages of our model to ab enabled a single staging model usable for everyone. Notably, we have been able to
367 relate the CSF stages of our model to abnormality in other well-described AD
368 biomarkers, such as Aβ-PET, tau-PET, MRI and in cognit 367 relate the CSF stages of our model to abnormality in other well-described AD
368 biomarkers, such as A β -PET, tau-PET, MRI and in cognitive measures. Further, our
369 CSF staging model was able to accurately predict 368 biomarkers, such as Aβ-PET, tau-PET, MRI and in cognitive measures. Further, our
369 CSF staging model was able to accurately predict positivity of the imaging biomarkers
370 of Aβ and tau, and to predict A/T status, 369 CSF staging model was able to accurately predict positivity of the imaging biomarkers
370 of Aβ and tau, and to predict A/T status, as assessed by PET. The CSF staging model
371 was also related to cognitive stages a 370 of Aβ and tau, and to predict A/T status, as assessed by PET. The CSF staging model
371 was also related to cognitive stages and was able to differentiate cognitive impairment
372 due to AD from other dementias. Impor 371 was also related to cognitive stages and was able to differentiate cognitive impairment
372 due to AD from other dementias. Importantly, we also observed different longitudinal
373 rates of change of AD biomarkers at d 372 due to AD from other dementias. Importantly, we also observed different longitudinal
373 rates of change of AD biomarkers at different CSF stages, which may allow us to
374 determine which participants will progress mo 373 rates of change of AD biomarkers at different CSF stages, which may allow us to
374 determine which participants will progress more in key aspects of the disease. We also
375 showed that participants in the more advanc 374 determine which participants will progress more in key aspects of the disease. We also
375 showed that participants in the more advanced stages of our CSF-based model were at
376 higher risk for clinical decline. And, 375 showed that participants in the more advanced stages of our CSF-based model were at
376 higher risk for clinical decline. And, finally, we were able to replicate the model and
377 main results in an independent cohort. 376 higher risk for clinical decline. And, finally, we were able to replicate the model and
377 main results in an independent cohort. Altogether, these results prove the validity and
378 clinical utility of our CSF stagin clinical practice and in clinical trials 44 .

377 main results in an independent cohort. Altogether, these results prove the validity and
378 clinical utility of our CSF staging model, suggesting that it may hold promise for both
379 clinical practice and in clinical 378 clinical utility of our CSF staging model, suggesting that it may hold promise for both
379 clinical practice and in clinical trials⁴⁴.
380 The first aim of this analysis sought to establish whether there was a ster 379 clinical practice and in clinical trials⁴⁴.

380 The first aim of this analysis sough

381 ordering in when key CSF biomark

382 approach to solve this question, as i

383 existent, using cross-sectional data⁴

38 380 The first aim of this analysis sought to establish whether there was a stereotypic
381 ordering in when key CSF biomarkers become abnormal. SuStaln is an optimal
382 approach to solve this question, as it allows the m 381 ordering in when key CSF biomarkers become abnormal. SuStaln is an optimal
382 approach to solve this question, as it allows the modelling of different trajectories, if
383 existent, using cross-sectional data⁴¹, as approach to solve this question, as it allows the modelling of different trajectories, if
383 existent, using cross-sectional data⁴¹, as has been successfully applied to imaging
384 biomarkers^{35,36,45}. We observed that existent, using cross-sectional data⁴¹, as has been successfully applied to imaging 383 existent, using cross-sectional data⁴¹, as has been successfully applied to imaging
384 biomarkers^{35,36,45}. We observed that the CSF biomarkers investigated in this study
385 became abnormal in a particular sequ biomarkers^{35,36,45} 384 biomarkers^{35,36,45}. We observed that the CSF biomarkers investigated in this study
385 became abnormal in a particular sequence, and more importantly, that this sequence
10 385 became abnormal in a particular sequence, and more importantly, that this sequence
10
10

386 did not vary systematically across participants. This result is important by itself as it
387 tells us that there may be a single cascade of events that leads to sequential
388 ahormality of these soluble proteins in 387 tells us that there may be a single cascade of events that leads to sequential
388 abnormality of these soluble proteins in the brain, common to all AD patients. Previous
389 studies already suggested that changes in 388 abnormality of these soluble proteins in the brain, common to all AD patients. Previous
389 studies already suggested that changes in the levels of tau fragments phosphorylated
390 at different sites may be linked mec 389 studies already suggested that changes in the levels of tau fragments phosphorylated
390 at different sites may be linked mechanistically and could be associated with disease
391 stages^{46–50}. Based on our results, A 390 at different sites may be linked mechanistically and could be associated with disease
391 stages^{46–50}. Based on our results, A β plaques reflected by an imbalance of soluble
392 amyloid species (*i.e.*, low A β 4 stages^{46–50} 391 stages^{46–50}. Based on our results, Aβ plaques reflected by an imbalance of soluble amyloid species (*i.e.*, low Aβ42/40) may drive hyper-phosphorylation of tau in early phosphorylation site (pT217/T217), as previou 392 amyloid species (*i.e.*, low A β 42/40) may drive hyper-phosphorylation of tau in early
393 phosphorylation site ($pT217/T217$), as previously suggested by human and animal
394 data^{51,52}, which would subsequently be 393 phosphorylation site (pT217/T217), as previously suggested by human and animal
394 data^{51,52}, which would subsequently be followed by hyper-phosphorylation in later site
395 (pT205/T205) and eventually increase othe $data^{51,52}$, which would subsequently be followed by hyper-phosphorylation in later site 394 data^{51,52}, which would subsequently be followed by hyper-phosphorylation in later site
395 (pT205/T205) and eventually increase other tau fragments (MTBR-tau243 and total
396 tau) due to tangles formation and neurod 395 (pT205/T205) and eventually increase other tau fragments (MTBR-tau243 and total
396 tau) due to tangles formation and neurodegeneration, respectively. Notably, this
397 sequence of events is in line with previous lite 396 tau) due to tangles formation and neurodegeneration, respectively. Notably, this sequence of events is in line with previous literature^{53,54}, and demonstrates that late onset sporadic AD molecular pathway matches th sequence of events is in line with previous literature^{53,54}, and demonstrates that late 397 sequence of events is in line with previous literature^{53,54}, and demonstrates that late
398 onset sporadic AD molecular pathway matches the same sequence of events as
399 autosomal dominant AD^{29} . Exploring in det 398 onset sporadic AD molecular pathway matches the same sequence of events as
399 autosomal dominant AD^{29} . Exploring in detail this cascade of events may provide
400 mechanistic insights into disease pathology and pro autosomal dominant AD^{29} . Exploring in detail this cascade of events may provide 399 autosomal dominant AD²⁹. Exploring in detail this cascade of events may provide
400 mechanistic insights into disease pathology and progression. And, in turn, could have
401 important consequences in drug developmen mechanistic insights into disease pathology and progression. And, in turn, could have

401 important consequences in drug development, as targeting some of the earliest events

402 of this sequence may stop or reduce subs having a significant effect on tau aggregation^{46,50,55}.

401 important consequences in drug development, as targeting some of the earliest events
402 of this sequence may stop or reduce subsequent events in the cascade and thereby
403 having a significant effect on tau aggregat 402 of this sequence may stop or reduce subsequent events in the cascade and thereby

403 having a significant effect on tau aggregation^{46,50,55}.

404 Nonetheless, perhaps the most important result of our study was prov 403 having a significant effect on tau aggregation^{46,50,55}.
404 Nonetheless, perhaps the most important result of c
405 CSF model as a method to stage AD *in vivo*⁴⁴. In
406 related to main molecular changes and clini 404 Nonetheless, perhaps the most important result of our study was proving the utility of a
405 CSF model as a method to stage AD *in vivo*⁴⁴. In our model, CSF stages could be
406 related to main molecular changes and CSF model as a method to stage AD *in vivo*⁴⁴. In our model, CSF stages could be 405 CSF model as a method to stage AD *in vivo*⁴⁴. In our model, CSF stages could be related to main molecular changes and clinical tipping points in the course of the disease, including abnormal levels of deposited AB 406 related to main molecular changes and clinical tipping points in the course of the
407 disease, including abnormal levels of deposited A β (CSF stage 2:
408 pT217/T217)^{19,20,23,25,26,28,29,32} and tau (CSF stages 3 407 disease, including abnormal levels of deposited Aβ (CSF stage 2:

408 pT217/T217)^{19,20,23,25,26,28,29,32} and tau (CSF stages 3: pT205/T205^{28,29,32}), early cognitive

409 impairment (CSF stage 4: MTBR-tau243³²) $pT217/T217$ ^{19,20,23,25,26,28,29,32} and tau (CSF stages 3: $pT205/T205^{28,29,32}$), early cognitive 408 pT217/T217)^{19,20,23,25,26,28,29,32} and tau (CSF stages 3: pT205/T205^{28,29,32}), early cognitive
409 impairment (CSF stage 4: MTBR-tau243³²) and neurodegeneration (CSF stage 5: total
410 tau⁴²), following the e impairment (CSF stage 4: MTBR-tau243 32) and neurodegeneration (CSF stage 5: total 409 impairment (CSF stage 4: MTBR-tau243³²) and neurodegeneration (CSF stage 5: total
410 tau⁴²), following the expected pattern. With the objective of characterizing the molecular
411 status of the participants usin $tau^{(2)}$, following the expected pattern. With the objective of characterizing the molecular 410 tau⁴²), following the expected pattern. With the objective of characterizing the molecular status of the participants using our model, we observed that participants at CSF stages 2 and 3 (pT217/T217 and pT205/T205 411 status of the participants using our model, we observed that participants at CSF stages
412 and 3 (pT217/T217 and pT205/T205 stages) could be categorized with high accuracy
413 as being Aβ-positive and tau-negative by 412 2 and 3 (pT217/T217 and pT205/T205 stages) could be categorized with high accuracy

413 as being Aβ-positive and tau-negative by PET (A+T-), while participants at CSF stage 4

414 (MTBR-tau243) or higher were Aβ- and 413 as being Aβ-positive and tau-negative by PET (A+T-), while participants at CSF stage 4 (MTBR-tau243) or higher were Aβ- and tau-PET positive (A+T+)⁴². Importantly, these cut-points were reproduced in the Knight-ADR (MTBR-tau243) or higher were Aβ- and tau-PET positive $(A+T+)^{42}$. Importantly, these 414 (MTBR-tau243) or higher were β P and tau-PET positive $(A+T+)^{42}$. Importantly, these
415 cut-points were reproduced in the Knight-ADRC cohort, even using different PET
416 tracers and quantification methods support 415 cut-points were reproduced in the Knight-ADRC cohort, even using different PET
416 tracers and quantification methods supporting the consistency of the model. Being able
417 to accurately assess Aβ and tau status wit 416 tracers and quantification methods supporting the consistency of the model. Being able
417 to accurately assess A β and tau status with a single CSF collection may be very useful
418 to select the optimal participan 417 to accurately assess Aβ and tau status with a single CSF collection may be very useful
418 to select the optimal participants for a clinical trial, such as has been done in the
419 donanemab trial (NCT03367403)¹¹, 418 to select the optimal participants for a clinical trial, such as has been done in the
419 donanemab trial (NCT03367403)¹¹, without the need of acquiring both an Aβ-PET and
420 tau-PET scan to determine if a patient donanemab trial (NCT03367403)¹¹, without the need of acquiring both an A β -PET and 419 donanemab trial (NCT03367403)¹¹, without the need of acquiring both an Aβ-PET and
420 tau-PET scan to determine if a patient is eligible for treatment. In BioFINDER-2, we
421 also observed the diagnostic utility of 420 tau-PET scan to determine if a patient is eligible for treatment. In BioFINDER-2, we
421 also observed the diagnostic utility of this CSF staging model, as it was able to
422 accurately discern AD from non-AD related c 421 also observed the diagnostic utility of this CSF staging model, as it was able to
422 accurately discern AD from non-AD related cognitive impairment, and could
11 422 accurately discern AD from non-AD related cognitive impairment, and could

differentiate cognitive stages. Thus, the use of our model as a diagnostic tool may have

424 important consequences also at the clinical level.

425 Notably, our CSF staging model also showed prognostic utility. First, w 424 important consequences also at the clinical level.
425 Notably, our CSF staging model also showed pro
426 participants at different CSF stages showed c
427 biomarkers. For instance, rates of Aβ accumula
428 previousl Action 425 Notably, our CSF staging model also showed prognostic utility. First, we observed that
426 participants at different CSF stages showed different rates of change in multiple
427 biomarkers. For instance, rates o 426 participants at different CSF stages showed different rates of change in multiple
427 biomarkers. For instance, rates of Aβ accumulation across CSF stages showed the
428 previously reported inverted U shape^{3,56}, wi 427 biomarkers. For instance, rates of Aβ accumulation across CSF stages showed the
428 previously reported inverted U shape^{3,56}, with participants at CSF stage 2 (pT217/T217)
429 exhibiting the highest rates of change. previously reported inverted U shape^{3,56}, with participants at CSF stage 2 ($pT217/T217$) quares are the highest rates of change. On the other hand, the other imaging
exhibiting the highest rates of change. On the other hand, the other imaging
displanant and cognitive scores showed increased rate of change wit 429 exhibiting the highest rates of change. On the other hand, the other imaging
430 biomarkers and cognitive scores showed increased rate of change with increasing CSF
431 stages, only plateauing at the last stage, as ex biomarkers and cognitive scores showed increased rate of change with increasing CSF
431 stages, only plateauing at the last stage, as expected⁵⁷. These results support the use
432 of our staging model as an enrichment te stages, only plateauing at the last stage, as expected 5^7 . These results support the use 431 stages, only plateauing at the last stage, as expected⁹⁷. These results support the use
432 of our staging model as an enrichment technique for clinical trials⁵⁸. But, more
433 importantly, we also observed that th of our staging model as an enrichment technique for clinical trials⁵⁸. But, more 432 of our staging model as an enrichment technique for clinical trials⁵⁸. But, more
433 importantly, we also observed that the CSF staging model was able to predict clinical
434 progression. Being at the later stages of importantly, we also observed that the CSF staging model was able to predict clinical
434 progression. Being at the later stages of our model, increased the risk of progressing to
435 AD dementia, even when accounting for progression. Being at the later stages of our model, increased the risk of progressing to
435 AD dementia, even when accounting for cognitive status at baseline (Fig. 5). Further,
436 we also observed a higher risk of prog 435 AD dementia, even when accounting for cognitive status at baseline (Fig. 5). Further,
436 we also observed a higher risk of progressing to MCI or very mild AD, although this
437 analysis should be replicated in larger 436 we also observed a higher risk of progressing to MCI or very mild AD, although this
437 analysis should be replicated in larger cohorts with longer follow-up. Notably, the
438 prognostic ability of our CSF staging mod analysis should be replicated in larger cohorts with longer follow-up. Notably, the

438 prognostic ability of our CSF staging model was replicated in the Knight-ADRC cohort.

439 These results suggest a clear prognostic prognostic ability of our CSF staging model was replicated in the Knight-ADRC cohort.

439 These results suggest a clear prognostic utility on staging participants based on their

440 CSF profile, which may imply significa to previous staging methods based on $PET^{34,36,37,59}$.

These results suggest a clear prognostic utility on staging participants based on their

440 CSF profile, which may imply significant reductions in costs and complexity compared

441 to previous staging methods based on 440 CSF profile, which may imply significant reductions in costs and complexity compared
441 to previous staging methods based on PET^{34,36,37,59}.
442 Importantly, we view the present model as a first step toward providi 441 to previous staging methods based on PET^{34,36,37,59}.
442 Importantly, we view the present model as a first
443 disease progression staging using a single CSF
444 additional biomarkers will be included to the model
4 Importantly, we view the present model as a first step toward providing meaningful

443 disease progression staging using a single CSF measurement⁴⁴. We expect that

444 additional biomarkers will be included to the mod disease progression staging using a single CSF measurement⁴⁴. We expect that disease progression staging using a single CSF measurement⁴⁴. We expect that
444 additional biomarkers will be included to the model to either gain further granularity in
445 specific disease stages, or to signify to ot additional biomarkers will be included to the model to either gain further granularity in
specific disease stages, or to signify to other pathophysiological events (*e.g.*, microglial
reactivity)⁶⁰. Being able to measur specific disease stages, or to signify to other pathophysiological events (*e.g.*, microglial
446 reactivity)⁶⁰. Being able to measure several pathophysiological abnormalities using one
447 sample is one of the main adva reactivity)⁶⁰. Being able to measure several pathophysiological abnormalities using one reactivity)^{ov}. Being able to measure several pathophysiological abnormalities using one
447 sample is one of the main advantages of using fluid samples instead of PET for
448 staging. Another advantage of this model is t sample is one of the main advantages of using fluid samples instead of PET for
448 staging. Another advantage of this model is that financial and infrastructure cost of CSF
449 is low compared to other measures, such as PE 448 staging. Another advantage of this model is that financial and infrastructure cost of CSF
449 is low compared to other measures, such as PET. Looking toward the future, we hope
450 to be able translate these results in 449 is low compared to other measures, such as PET. Looking toward the future, we hope
450 to be able translate these results into plasma biomarkers, which would facilitate even
451 greater availability and cost-effectiven 450 to be able translate these results into plasma biomarkers, which would facilitate even
451 greater availability and cost-effectiveness. Widespread use of our fluid biomarker
452 staging model in primary care would like quare availability and cost-effectiveness. Widespread use of our fluid biomarker
452 staging model in primary care would likely require replacing CSF measures with
453 plasma measures without greatly sacrificing model perf 452 staging model in primary care would likely require replacing CSF measures with
453 plasma measures without greatly sacrificing model performance. Efforts in this direction
454 are currently underway, but development of

plasma measures without greatly sacrificing model performance. Efforts in this direction
454 are currently underway, but development of reliable plasma assays for pT205/T205
455 and MTBR-tau243 are still ongoing.
456 This are currently underway, but development of reliable plasma assays for pT205/T205
455 and MTBR-tau243 are still ongoing.
456 This study has several strengths but also some limitations. The main strength of this
457 study is 455 and MTBR-tau243 are still ongoing.
456 This study has several strengths bu
457 study is the proven utility of the mod
458 and thereby supports the generali
459 strength is the use of several bu This study has several strengths but also some limitations. The main strength of this
457 study is the proven utility of the model, which was replicated in an independent cohort,
458 and thereby supports the generalizabili 457 study is the proven utility of the model, which was replicated in an independent cohort,
458 and thereby supports the generalizability of our staging model. Another important
459 strength is the use of several biomarke 458 and thereby supports the generalizability of our staging model. Another important
459 strength is the use of several biomarkers measured with very high-performing
12 459 strength is the use of several biomarkers measured with very high-performing
12

assays^{28,61}, which is crucial for the accurate assessment of pathology⁴. However, some assays^{28,61}, which is crucial for the accurate assessment of pathology⁴. However, some

461 limitations must be recognized. Although we included CSF biomarkers with proven

462 utility, we acknowledge that there are s 461 limitations must be recognized. Although we included CSF biomarkers with proven
462 utility, we acknowledge that there are some other interesting markers, such as p-
463 taging model in its current form was still succ 462 utility, we acknowledge that there are some other interesting markers, such as p-

463 tau235⁶², that have not been analyzed in this study. However, we think that our CSF

464 staging model in its current form was s tau235 62 tau 235⁶², that have not been analyzed in this study. However, we think that our CSF
staging model in its current form was still successful at signaling the main inflection
points of the disease. Further, p-tau 231, whi 464 staging model in its current form was still successful at signaling the main inflection
465 points of the disease. Further, p-tau231, which is thought to become abnormal early in
466 the disease^{20–22}, although not a quarely in the disease. Further, p-tau231, which is thought to become abnormal early in the disease^{20–22}, although not always^{25,61}, was excluded from our model as it followed a similar abnormality tendency as pT217/T2 the disease $^{20-22}$, although not always 25,61 466 the disease^{20–22}, although not always^{25,61}, was excluded from our model as it followed a
467 similar abnormality tendency as pT217/T217, without providing better performance for
468 staging than the later. We hypo 467 similar abnormality tendency as $pT217/T217$, without providing better performance for
468 staging than the later. We hypothesize that this may be in part related to difference in
469 analytical performances, as the ma 468 staging than the later. We hypothesize that this may be in part related to difference in
469 analytical performances, as the mass spectrometry platform used in our study provided
470 rather higher coefficient of variat analytical performances, as the mass spectrometry platform used in our study provided

470 rather higher coefficient of variation for pT231/T231 measurements (12-18% compared

471 to 5-7% for pT217/T217). Future studies i 470 rather higher coefficient of variation for $pT231/T231$ measurements (12-18% compared
471 to 5-7% for $pT217/T217$). Future studies in earlier cohorts or with optimized assays for
472 measuring p-tau231 should test whet 471 to 5-7% for pT217/T217). Future studies in earlier cohorts or with optimized assays for
472 measuring p-tau231 should test whether the present model could be improved. Another
473 important issue is that we acknowledge measuring p-tau231 should test whether the present model could be improved. Another
473 important issue is that we acknowledge that CSF collection require trained clinicians,
474 and we plan to move towards a plasma-based 473 important issue is that we acknowledge that CSF collection require trained clinicians,
474 and we plan to move towards a plasma-based staging models when these biomarkers
475 become available. A replication of these re 474 and we plan to move towards a plasma-based staging models when these biomarkers
475 become available. A replication of these results in a more diverse population is also
476 needed to confirm the utility of our model i 475 become available. A replication of these results in a more diverse population is also
476 needed to confirm the utility of our model in a less selected population. We would also
477 like to point out that the CSF stage meeded to confirm the utility of our model in a less selected population. We would also

477 like to point out that the CSF stages proposed here are related to events of the disease

478 and not to time. Thus, it may be po

477 like to point out that the CSF stages proposed here are related to events of the disease
478 and not to time. Thus, it may be possible that the time for progressing from one CSF
479 stage to the next vary significantly 478 and not to time. Thus, it may be possible that the time for progressing from one CSF
479 stage to the next vary significantly depending on the CSF stage at baseline.
480 In conclusion, in this study we have developed a 479 stage to the next vary significantly depending on the CSF stage at baseline.
480 In conclusion, in this study we have developed an accurate staging mo
481 based on only five CSF biomarkers, and we have evaluated it in 480 In conclusion, in this study we have developed an accurate staging model for AD
481 based on only five CSF biomarkers, and we have evaluated it in two large independent
482 cohorts. We have shown that the model is stab 481 based on only five CSF biomarkers, and we have evaluated it in two large independent
482 cohorts. We have shown that the model is stable, and accurately reflects biomarker
483 changes in AD, providing an easier and che cohorts. We have shown that the model is stable, and accurately reflects biomarker

483 changes in AD, providing an easier and cheaper method for characterization of

484 participants for both clinical setting and trials. 483 changes in AD, providing an easier and cheaper method for characterization of
484 participants for both clinical setting and trials. Further, our model has demonstrated its
485 utility for prognosis, being able to iden 484 participants for both clinical setting and trials. Further, our model has demonstrated its
485 utility for prognosis, being able to identify participants with more pronounced
486 longitudinal changes in AD biomarkers a utility for prognosis, being able to identify participants with more pronounced
486 longitudinal changes in AD biomarkers as well as those individuals with higher risk of
487 deteriorating in cognitive status. This CSF sta 186 longitudinal changes in AD biomarkers as well as those individuals with higher risk of deteriorating in cognitive status. This CSF staging model may be a useful, cheap, and accessible method in clinical trials for opt deteriorating in cognitive status. This CSF staging model may be a useful, cheap, and
accessible method in clinical trials for optimal selection of study participants and as a
489 surrogate outcome measure. Further, the st 488 accessible method in clinical trials for optimal selection of study participants and as a
489 surrogate outcome measure. Further, the staging model has great potential for use in
490 clinical practice in the diagnosti 489 surrogate outcome measure. Further, the staging model has great potential for use in
490 clinical practice in the diagnostic and prognostic work-up of patients with cognitive
491 symptoms and potentially also for sele 490 clinical practice in the diagnostic and prognostic work-up of patients with cognitive
491 symptoms and potentially also for selecting optimal candidates for disease-modifying
492 treatments. And we expect it may have 491 symptoms and potentially also for selecting optimal candidates for disease-modifying
492 treatments. And we expect it may have an influence on the update of the $A/T/(N)$
493 criteria.⁴² Altogether, our staging model 492 treatments. And we expect it may have an influence on the update of the $AT/(N)$
493 criteria.⁴² Altogether, our staging model may be an important step towards a more
494 sophisticated personalized medicine approach o criteria.⁴² Altogether, our staging model may be an important step towards a more expanding the may be an important step towards a more
sophisticated personalized medicine approach of AD, which will be key with the
advancement of novel disease-modifying treatments.
496 494 sophisticated personalized medicine approach of AD, which will be key with the
495 advancement of novel disease-modifying treatments.
496 495 advancement of novel disease-modifying treatments.
496
The section of the sease of the section of the sease of the

497 **DATA AVAILABILITY:**
498 The datasets generated
499 authors (O.H and R.J.E
500 approvals, upon reason:
501 For BioFINDER-2 data,
602 academic investigator

The datasets generated and/or analyzed during the current study are available from the

499 authors (O.H and R.J.B). We will share datasets within the restrictions of IRB ethics

500 approvals, upon reasonable request.

50 authors (O.H and R.J.B). We will share datasets within the restrictions of IRB ethics
500 approvals, upon reasonable request.
501 For BioFINDER-2 data, anonymized data will be shared by request from a qualified
502 academi soo approvals, upon reasonable request.
501 For BioFINDER-2 data, anonymized
502 academic investigator for the sole
503 presented in the article and as long a
504 on the general data protection regula
505 Sweden and Region 501 For BioFINDER-2 data, anonymized data will be shared by request from a qualified
502 academic investigator for the sole purpose of replicating procedures and results
503 presented in the article and as long as data tra 502 academic investigator for the sole purpose of replicating procedures and results
503 presented in the article and as long as data transfer is in agreement with EU legislation
504 on the general data protection regulati 503 presented in the article and as long as data transfer is in agreement with EU legislation

504 on the general data protection regulation and decisions by the Ethical Review Board of

505 Sweden and Region Skåne, which 504 on the general data protection regulation and decisions by the Ethical Review Board of
505 Sweden and Region Skåne, which should be regulated in a material transfer
506 agreement.
507 **DISCLOSURES**
510 **DISCLOSURES**

505 Sweden and Region Skåne, which should be regulated in a material transfer
506 agreement.
507
508 **DISCLOSURES**
510 OH has acquired research support (for the institution) from ADx, AVID
511 Radiopharmaceuticals, Biogen, 506 agreement.
507
508 **DISCLOSUI**
510 OH has
511 Radiopharm
512 Roche. In t 508
509
510
511
512
513 509
510
511
512
513
514 509 **DISCLOSURES**
510 OH has acqui
511 Radiopharmaceu
512 Roche. In the p
513 Immune, Amyly;
514 Genentech, Mer
515 supported by the 510 OH has acquired research support (for the institution) from ADx, AVID
511 Radiopharmaceuticals, Biogen, Eli Lilly, Eisai, Fujirebio, GE Healthcare, Pfizer, and
512 Roche. In the past 2 years, he has received consultanc 511 Radiopharmaceuticals, Biogen, Eli Lilly, Eisai, Fujirebio, GE Healthcare, Pfizer, and
512 Roche. In the past 2 years, he has received consultancy/speaker fees from AC
513 Immune, Amylyx, Alzpath, BioArctic, Biogen, Cer 512 Roche. In the past 2 years, he has received consultancy/speaker fees from AC
513 Immune, Amylyx, Alzpath, BioArctic, Biogen, Cerveau, Eisai, Eli Lilly, Fujirebio,
514 Genentech, Merck, Novartis, Novo Nordisk, Roche, Sa 513 Immune, Amylyx, Alzpath, BioArctic, Biogen, Cerveau, Eisai, Eli Lilly, Fujirebio,
514 Genentech, Merck, Novartis, Novo Nordisk, Roche, Sanofi and Siemens. JWV is
515 supported by the SciLifeLab & Wallenberg Data Driven 514 Genentech, Merck, Novartis, Novo Nordisk, Roche, Sanofi and Siemens. JWV is
515 supported by the SciLifeLab & Wallenberg Data Driven Life Science Program (grant:
516 KAW 2020.0239). KH is an Eisai-sponsored voluntary r supported by the SciLifeLab & Wallenberg Data Driven Life Science Program (grant:

516 KAW 2020.0239). KH is an Eisai-sponsored voluntary research associate professor at

517 Washington University and has received salary f 516 KAW 2020.0239). KH is an Eisai-sponsored voluntary research associate professor at
517 Washington University and has received salary from Eisai. Washington University,
518 RJB, and DMH have equity ownership interest in Washington University and has received salary from Eisai. Washington University,

518 RJB, and DMH have equity ownership interest in C2N Diagnostics. RJB. and DMH

520 RKB, and RJB. may receive income based on technology (518 RJB, and DMH have equity ownership interest in C2N Diagnostics. RJB. and DMH
519 receive income from C2N Diagnostics for serving on the scientific advisory board. KH,
520 NRB, and RJB. may receive income based on techn 519 receive income from C2N Diagnostics for serving on the scientific advisory board. KH,
520 NRB, and RJB. may receive income based on technology (METHODS TO DETECT
521 MTBR TAU ISOFORMS AND USE THEREOF) licensed by Washi S20 NRB, and RJB. may receive income based on technology (METHODS TO DETECT
521 MTBR TAU ISOFORMS AND USE THEREOF) licensed by Washington University to
522 C2N Diagnostics. DMH may receive income based on technology (ANTIB 521 MTBR TAU ISOFORMS AND USE THEREOF) licensed by Washington University to
522 C2N Diagnostics. DMH may receive income based on technology (ANTIBODIES TO
523 MID-DOMAIN OF TAU) licensed by Washington University to C2N Dia 522 C2N Diagnostics. DMH may receive income based on technology (ANTIBODIES TO
523 MID-DOMAIN OF TAU) licensed by Washington University to C2N Diagnostics. RJB
524 has received honoraria as a speaker, consultant, or adviso MID-DOMAIN OF TAU) licensed by Washington University to C2N Diagnostics. RJB
524 has received honoraria as a speaker, consultant, or advisory board member from
525 Amgen and Roche. DMH is on the scientific advisory board o 524 has received honoraria as a speaker, consultant, or advisory board member from
525 Amgen and Roche. DMH is on the scientific advisory board of Genentech, Denali, and
526 Cajal Neurosciences, and consults for Alector an 525 Amgen and Roche. DMH is on the scientific advisory board of Genentech, Denali, and
526 Cajal Neurosciences, and consults for Alector and Asteroid. NRB is co-inventor on the
527 following US patent applications: 'Method 526 Cajal Neurosciences, and consults for Alector and Asteroid. NRB is co-inventor on the
527 following US patent applications: 'Methods to detect novel tau species in CSF and use
528 thereof to track tau neuropathology in 527 following US patent applications: 'Methods to detect novel tau species in CSF and use

thereof to track tau neuropathology in Alzheimer's disease and other tauopathies'

(PCT/US2020/046224); 'CSF phosphorylated tau and 528 thereof to track tau neuropathology in Alzheimer's disease and other tauopathies' (PCT/US2020/046224); 'CSF phosphorylated tau and amyloid beta profiles as biomarkers of tauopathies' (PCT/US2022/022906); and 'Methods o 529 (PCT/US2020/046224); 'CSF phosphorylated tau and amyloid beta profiles as
530 biomarkers of tauopathies' (PCT/US2022/022906); and 'Methods of diagnosing and
531 treating based on site-specific tau phosphorylation' (PCT 530 biomarkers of tauopathies' (PCT/US2022/022906); and 'Methods of diagnosing and
531 treating based on site-specific tau phosphorylation' (PCT/US2019/030725). NRB may
532 receive a royalty income based on technology lice 531 treating based on site-specific tau phosphorylation' (PCT/US2019/030725). NRB may
532 receive a royalty income based on technology licensed by Washington University to
533 C2N Diagnostics.
534 532 receive a royalty income based on technology licensed by Washington University to
533 C2N Diagnostics.
534 533 C2N Diagnostics.
534
C2N Diagnostics.

535

536

536
537
538
539
540 537
538
539
540
541 537 **FUNDING**

538 The BioFIN

539 ERA Perl

540 (2017-0383)

541 Parkinson's

542 980907), the State Constant 538 The BioFINDER study was supported by the Swedish Research Council (2022-00775),
539 ERA PerMed (ERAPERMED2021-184), the Knut and Alice Wallenberg foundation
540 (2017-0383), the Strategic Research Area MultiPark (Multi ERA PerMed (ERAPERMED2021-184), the Knut and Alice Wallenberg foundation

540 (2017-0383), the Strategic Research Area MultiPark (Multidisciplinary Research in

541 Parkinson's disease) at Lund University, the Swedish Alzh 540 (2017-0383), the Strategic Research Area MultiPark (Multidisciplinary Research in

541 Parkinson's disease) at Lund University, the Swedish Alzheimer Foundation (AF-

542 980907), the Swedish Brain Foundation (FO2021-541 Parkinson's disease) at Lund University, the Swedish Alzheimer Foundation (AF-
542 980907), the Swedish Brain Foundation (FO2021-0293), The Parkinson foundation of
543 Sweden (1412/22), the Cure Alzheimer's fund, the 542 980907), the Swedish Brain Foundation (FO2021-0293), The Parkinson foundation of
543 Sweden (1412/22), the Cure Alzheimer's fund, the Konung Gustaf V:s och Drottning
544 Victorias Frimurarestiftelse, the Skåne Univers 543 Sweden (1412/22), the Cure Alzheimer's fund, the Konung Gustaf V:s och Drottning

544 Victorias Frimurarestiftelse, the Skåne University Hospital Foundation (2020-O000028),

545 Regionalt Forskningsstöd (2022-1259) an Victorias Frimurarestiftelse, the Skåne University Hospital Foundation (2020-O000028),

545 Regionalt Forskningsstöd (2022-1259) and the Swedish federal government under the

546 ALF agreement (2022-Projekt0080). The precu Example Torskningsstöd (2022-1259) and the Swedish federal government under the

546 ALF agreement (2022-Projekt0080). The precursor of ¹⁸F-flutemetamol was sponsored

547 by GE Healthcare. The precursor of ¹⁸F-RO948 w ALF agreement (2022-Projekt0080). The precursor of 18 F-flutemetamol was sponsored 546 ALF agreement (2022-Projekt0080). The precursor of ¹⁸F-flutemetamol was sponsored
547 by GE Healthcare. The precursor of ¹⁸F-RO948 was provided by Roche. GS received
548 funding from the European Union's Horizon 20 by GE Healthcare. The precursor of ¹⁸F-RO948 was provided by Roche. GS received by GE Healthcare. The precursor of ¹⁸F-RO948 was provided by Roche. GS received
548 funding from the European Union's Horizon 2020 research and innovation program
550 under the Marie Sklodowska-Curie action grant agreeme 548 funding from the European Union's Horizon 2020 research and innovation program
549 under the Marie Sklodowska-Curie action grant agreement No 101061836, from an
550 Alzheimer's Association Research Fellowship (#AARF-22 949 under the Marie Sklodowska-Curie action grant agreement No 101061836, from an
550 Alzheimer's Association Research Fellowship (#AARF-22-972612), Greta och Johan
551 Kocks research grants and, travel grants from the Str

550 Alzheimer's Association Research Fellowship (#AARF-22-972612), Greta och Johan
551 Kocks research grants and, travel grants from the Strategic Research Area MultiPark
552 (Multidisciplinary Research in Parkinson's dise S51 Kocks research grants and, travel grants from the Strategic Research Area MultiPark

(Multidisciplinary Research in Parkinson's disease) at Lund University.

This work was supported by resources and effort provided by (Multidisciplinary Research in Parkinson's disease) at Lund University.

553 This work was supported by resources and effort provided by the Tra

554 Center (PI: RJB) established by the Tracy Family, Richard Frimel at

555 This work was supported by resources and effort provided by the Tracy Family SILQ
554 Center (PI: RJB) established by the Tracy Family, Richard Frimel and Gary Werths,
555 GHR Foundation, David Payne, and the Willman Famil S54 Center (PI: RJB) established by the Tracy Family, Richard Frimel and Gary Werths,
555 GHR Foundation, David Payne, and the Willman Family brought together by The
556 Foundation for Barnes-Jewish Hospital. This work was 555 GHR Foundation, David Payne, and the Willman Family brought together by The
556 Foundation for Barnes-Jewish Hospital. This work was also supported by resources
557 and effort provided by the Hope Center for Neurologic 556 Foundation for Barnes-Jewish Hospital. This work was also supported by resources
557 and effort provided by the Hope Center for Neurological Disorders and the Department
558 of Neurology at the Washington University Sc 557 and effort provided by the Hope Center for Neurological Disorders and the Department
558 of Neurology at the Washington University School of Medicine. This work was also
559 supported by the Clinical, Fluid Biomarker, 558 of Neurology at the Washington University School of Medicine. This work was also

559 supported by the Clinical, Fluid Biomarker, and Imaging Cores of the Knight ADRC

560 (P30 AG066444 [PI: JCM], P01 AG03991 [PI: JCM] 559 supported by the Clinical, Fluid Biomarker, and Imaging Cores of the Knight ADRC (P30 AG066444 [PI: JCM], P01 AG03991 [PI: JCM], and P01 AG026276 [PI: JCM]) at the Washington University School of Medicine for participa 560 (P30 AG066444 [PI: JCM], P01 AG03991 [PI: JCM], and P01 AG026276 [PI: JCM]) at
561 the Washington University School of Medicine for participant evaluation, samples, and
562 data collection. The mass spectrometry analys 561 the Washington University School of Medicine for participant evaluation, samples, and

562 data collection. The mass spectrometry analyses of BioFINDER-2 and Knight ADRC

563 samples were supported by Eisai industry gr 562 data collection. The mass spectrometry analyses of BioFINDER-2 and Knight ADRC
563 samples were supported by Eisai industry grant to Washington University (PI: KH and
564 RJB) and the Knight ADRC developmental project 563 samples were supported by Eisai industry grant to Washington University (PI: KH and
564 RJB) and the Knight ADRC developmental project (PI: NRB) and R01AG070941 (PI:
565 SES).
566 The funding sources had no role in the

564 RJB) and the Knight ADRC developmental project (PI: NRB) and R01AG070941 (PI:
565 SES).
566
567 The funding sources had no role in the design and conduct of the study; in the
569 collection, analysis, interpretation of 565 SES).
566
567
568 The fu
569 collect
570 of the l
571 - 567
568
569
570
571
572 568
569
570
571
572 568 The funding sources had no role in the design and conduct of the study; in the collection, analysis, interpretation of the data; or in the preparation, review, or approval of the manuscript.
570 of the manuscript.
571 569 collection, analysis, interpretation of the data; or in the preparation, review, or approval
570 of the manuscript.
571
572 570 of the manuscript.
571
572

-
- 572
1

573 573 574

783

785 **Online methods**
786 **Participants**
787 *BioFINDER-2*
788 We assessed
789 (NCT03174938)¹⁹,
790 were recruited at S 786 **Participants**

787 *BioFINDER-2*

788 We assess

789 (NCT0317493

790 were recruited

791 These partici 787 *BioFINDER-2*
788 We assesse
789 (NCT0317493
790 were recruited
791 These participate in 788 We assessed 426 participants from the Swedish BioFINDER-2 study

789 (NCT03174938)¹⁹, with the complete set of CSF biomarkers available. Participants

790 were recruited at Skåne University Hospital and the Hospital $(NOT03174938)^{19}$, with the complete set of CSF biomarkers available. Participants 789 (NCT03174938)¹⁹, with the complete set of CSF biomarkers available. Participants
790 were recruited at Skåne University Hospital and the Hospital of Ängelholm in Sweden.
791 These participants also had amyloid-PET (790 were recruited at Skåne University Hospital and the Hospital of Angelholm in Sweden.
791 These participants also had amyloid-PET (n=251), tau-PET (n=417), a magnetic
792 resonance imaging (MRI, n=420) and cognitive as 791 These participants also had amyloid-PET (n=251), tau-PET (n=417), a magnetic
792 resonance imaging (MRI, n=420) and cognitive assessment (n=426). Also, 220
793 participants had available CSF biomarkers at follow-up (m resonance imaging (MRI, n=420) and cognitive assessment (n=426). Also, 220

793 participants had available CSF biomarkers at follow-up (mean time (SD) = 2.05(0.22)

794 years). Inclusion and exclusion criterion for this s 293 participants had available CSF biomarkers at follow-up (mean time (SD) = 2.05(0.22)

294 years). Inclusion and exclusion criterion for this study has been detailed before¹⁹. In

295 summary, CU participants do not f years). Inclusion and exclusion criterion for this study has been detailed before¹⁹. In 794 years). Inclusion and exclusion criterion for this study has been detailed before¹⁹. In

795 summary, CU participants do not fulfil criteria for MCI or dementia according to DSM-

796 5⁶³. Subjective cognitive dec 795 summary, CU participants do not fulfil criteria for MCI or dementia according to DSM-
796 5^{63} . Subjective cognitive decline (SCD) participants were considered as CU, in
797 accordance with the research framework b 5^{63} . Subjective cognitive decline (SCD) participants were considered as CU, in 796 $5^{\circ3}$. Subjective cognitive decline (SCD) participants were considered as CU, in accordance with the research framework by the National Institute on Aging-Alzheimer's Association⁶⁴. MCI participants had a MMSE sc zordance with the research framework by the National Institute on Aging-Alzheimer's

798 Association⁶⁴. MCI participants had a MMSE score above 23, they did not fulfil the

799 criteria for major neurocognitive disorder Association⁶⁴. MCI participants had a MMSE score above 23, they did not fulfil the 798 Association⁶⁴. MCI participants had a MMSE score above 23, they did not fulfil the
799 criteria for major neurocognitive disorder (dementia) according to DSM-5 and
800 performed worse than -1.5 SD in at least one co 799 criteria for major neurocognitive disorder (dementia) according to DSM-5 and
800 performed worse than -1.5 SD in at least one cognitive domain according to age and
801 education stratified test norms. AD dementia was 1800 performed worse than -1.5 SD in at least one cognitive domain according to age and

801 education stratified test norms. AD dementia was diagnosed according to the DSM-5

802 criteria for major neurocognitive disorde 801 education stratified test norms. AD dementia was diagnosed according to the DSM-5

802 criteria for major neurocognitive disorder due to AD and an abnormal biomarker for Aβ

803 pathology was also required. Participan criteria for major neurocognitive disorder due to AD and an abnormal biomarker for Aβ

pathology was also required. Participants fulfilling the criteria for any other dementia

were categorized as non-AD dementias, as pre were categorized as non-AD dementias, as previously described.¹⁹

903 pathology was also required. Participants fulfilling the criteria for any other dementia

904 were categorized as non-AD dementias, as previously described.¹⁹

905 *Knight-ADRC*

906 Knight ADRC cohort consisted of c 805
806
807
808
809 805 *Knight-ADRC*
806 Knight ADRC
807 memory and a
808 underwent a c
819 Rating (CDR[®] 806 Knight ADRC cohort consisted of community-dwelling volunteers enrolled in studies of
807 memory and aging at Washington University in St. Louis. All Knight ADRC participants
808 underwent a comprehensive clinical asse 807 memory and aging at Washington University in St. Louis. All Knight ADRC participants
808 underwent a comprehensive clinical assessment that included a detailed interview of a
810 collateral source, a neurological exam 808 underwent a comprehensive clinical assessment that included a detailed interview of a
809 collateral source, a neurological examination of the participant, the Clinical Dementia
810 Rating (CDR®)⁶⁵ and the MMSE⁶⁶. 809 collateral source, a neurological examination of the participant, the Clinical Dementia
810 Rating (CDR®)⁶⁵ and the MMSE⁶⁶. Individuals with a CDR of 0.5 or greater were
811 considered to have a dementia syndrome Rating (CDR®)⁶⁵ and the MMSE⁶⁶ Rating (CDR®)⁶⁵ and the MMSE⁶⁶. Individuals with a CDR of 0.5 or greater were
811 considered to have a dementia syndrome and the probable aetiology of the dementia
812 syndrome was formulated by clinicians based on cl 811 considered to have a dementia syndrome and the probable aetiology of the dementia
812 syndrome was formulated by clinicians based on clinical features in accordance with
813 standard criteria and methods⁶⁷. In Kni 812 syndrome was formulated by clinicians based on clinical features in accordance with
813 standard criteria and methods⁶⁷. In Knight-ADRC cohort, participants were categorized
814 as CU if they were scored CDR=0, ei standard criteria and methods⁶⁷. In Knight-ADRC cohort, participants were categorized Maximum standard criteria and methods⁶⁷. In Knight-ADRC cohort, participants were categorized

814 as CU if they were scored CDR=0, either Aβ negative or positive (CU- and CU+,

815 respectively); very mild AD patient 814 as CU if they were scored CDR=0, either Aβ negative or positive (CU- and CU+,
815 respectively); very mild AD patients (CDR=0.5); and mild AD dementia patients
816 (CDR≥1) if clinical syndrome was typical of symptoma 815 respectively); very mild AD patients (CDR=0.5); and mild AD dementia patients (CDR≥1) if clinical syndrome was typical of symptomatic AD. Participants with CDR≥0.5 with different aetiology were assessed as being other 816 (CDR≥1) if clinical syndrome was typical of symptomatic AD. Participants with
817 CDR≥0.5 with different aetiology were assessed as being other dementia patients
818 regardless of their amyloid status.
819 All partici

817 CDR≥0.5 with different aetiology were assessed as being other dementia patients
818 regardless of their amyloid status.
819 All participants gave written informed consent and ethical approval was granted by the
820 Re 818 regardless of their amyloid status.
819 All participants gave written inform
820 Regional Ethical Committee in Lu
821 Research Protection Office, respec 819 All participants gave written informed consent and ethical approval was granted by the
820 Regional Ethical Committee in Lund, Sweden and the Washington University Human
821 Research Protection Office, respectively.
22 820 Regional Ethical Committee in Lund, Sweden and the Washington University Human
821 Research Protection Office, respectively.
22 821 Research Protection Office, respectively.

Research Protection Office, respectively.

822

823
824
825
826
827 823 **Fluid biomarkers**
824 Measurement of C
825 tau243 and total-ta
826 both cohorts using
827 BioFINDER-2, CS
828 platform as previo Reasurement of CSF tau species (*i.e.*, pT205/T205, pT217/T217, pT231/T231, MTBR-

825 tau243 and total-tau [the residue 151-153]) was performed at Washington University in

826 both cohorts using the newly developed IP/M 825 tau243 and total-tau [the residue 151-153]) was performed at Washington University in
826 both cohorts using the newly developed IP/MS method, as previously detailed³². In
827 BioFINDER-2, CSF levels of Aβ42/40 and both cohorts using the newly developed IP/MS method, as previously detailed³². In 826 both cohorts using the newly developed IP/MS method, as previously detailed³². In
827 BioFINDER-2, CSF levels of Aβ42/40 and NfL were measured using the Elecsys
828 platform as previously explained¹⁹. Aβ positivi 827 BioFINDER-2, CSF levels of Aβ42/40 and NfL were measured using the Elecsys
828 platform as previously explained¹⁹. Aβ positivity was assessed using CSF Aβ42/40
829 (<0.080), unless otherwise stated, based on Gaussi platform as previously explained¹⁹ 828 platform as previously explained¹⁹. Aβ positivity was assessed using CSF Aβ42/40 (<0.080), unless otherwise stated, based on Gaussian-mixture modeling. In Knight-
ADRC, CSF Aβ42/40 levels were measured as explained 829 (<0.080), unless otherwise stated, based on Gaussian-mixture modeling. In Knight-

830 ADRC, CSF Aβ42/40 levels were measured as explained previously^{28,68}. The CSF Aβ42/40

831 positivity threshold (0.0673) had the ADRC, CSF Aβ42/40 levels were measured as explained previously^{28,68}. The CSF Aβ42/40 positivity threshold (0.0673) had the maximum combined sensitivity and specificity in 830 ADRC, CSF Aβ42/40 levels were measured as explained previously^{26,66}. The CSF Aβ42/40
831 positivity threshold (0.0673) had the maximum combined sensitivity and specificity in
832 distinguishing amyloid-PET status. 931 positivity threshold (0.0673) had the maximum combined sensitivity and specificity in

832 distinguishing amyloid-PET status. CSF NfL was measured with a commercial ELISA kit

933 (UMAN Diagnostics, Umea, Sweden), as (UMAN Diagnostics, Umea, Sweden), as explained previously 69 .

832 distinguishing amyloid-PET status. CSF NfL was measured with a commercial ELISA kit
833 (UMAN Diagnostics, Umea, Sweden), as explained previously⁶⁹.
835 **Image acquisition and processing**
836 Image acquisition and pr 233 (UMAN Diagnostics, Umea, Sweden), as explained previously⁵⁹.

234 **Image acquisition and processing**

236 **Image acquisition and processing details from BioFINDER**

237 reported¹⁹. In brief, amyloid- and tau-PET w 835
836
837
838
839 835 **Image acquisition and processing**
836 **Image acquisition and processing c**
837 reported¹⁹. In brief, amyloid- and tau
838 ^{[18}F]RO948, respectively. Amyloid-F
839 neocortical meta-ROI, and with the ce
840 of the AD 836 Image acquisition and processing details from BioFINDER-2 have been previously
837 reported¹⁹. In brief, amyloid- and tau-PET were acquired using $[18F]$ flutemetamol and
838 $[18F]RO948$, respectively. Amyloid-PET bi reported¹⁹. In brief, amyloid- and tau-PET were acquired using 1^{18} F]flutemetamol and example 1⁹. In brief, amyloid- and tau-PET were acquired using $[{}^{18}F]$ flutemetamol and $[{}^{18}F]$ RO948, respectively. Amyloid-PET binding was measured as SUVR using a neocortical meta-ROI, and with the cerebellar gre \mathfrak{l}^{18} 1^{18} F]RO948, respectively. Amyloid-PET binding was measured as SUVR using a neocortical meta-ROI, and with the cerebellar grey as a reference region. Of note, most of the AD dementia patients did not undergo amyloid-PE 839 neocortical meta-ROI, and with the cerebellar grey as a reference region. Of note, most
840 of the AD dementia patients did not undergo amyloid-PET in BioFINDER-2, due to the
841 study design. For main analyses, tau-P 840 of the AD dementia patients did not undergo amyloid-PET in BioFINDER-2, due to the
841 study design. For main analyses, tau-PET binding was measured in a temporal meta-
842 ROI⁷⁰, which included, entorhinal, amygdal 841 study design. For main analyses, tau-PET binding was measured in a temporal meta-

842 ROI⁷⁰, which included, entorhinal, amygdala, parahippocampal, fusiform, inferior

843 temporal, and middle temporal ROIs, using t ROI⁷⁰, which included, entorhinal, amygdala, parahippocampal, fusiform, inferior 842 $ROI^{(0)}$, which included, entorhinal, amygdala, parahippocampal, fusiform, inferior
843 temporal, and middle temporal ROIs, using the inferior cerebellar cortex as reference
844 region without partial volume correctio 843 temporal, and middle temporal ROIs, using the inferior cerebellar cortex as reference
844 region without partial volume correction. Additionally, tau-PET binding was also
845 measured in regions covering early (Braak I 844 region without partial volume correction. Additionally, tau-PET binding was also
845 measured in regions covering early (Braak I), intermediate (Braak III-IV) and late
846 (Braak V-VI) tau deposition areas⁴³. For ass measured in regions covering early (Braak I), intermediate (Braak III-IV) and late

846 (Braak V-VI) tau deposition areas⁴³. For assessing cortical thickness, T1-weighted

847 anatomical magnetization-prepared rapid grad (Braak V-VI) tau deposition areas⁴³. For assessing cortical thickness, T1-weighted 846 (Braak V-VI) tau deposition areas⁴³. For assessing cortical thickness, T1-weighted
847 anatomical magnetization-prepared rapid gradient echo (MPRAGE) images (1mm
848 isotropic voxels) were used. A cortical thickness 847 anatomical magnetization-prepared rapid gradient echo (MPRAGE) images (1mm
848 isotropic voxels) were used. A cortical thickness meta-ROI was calculated including
849 entorhinal, inferior temporal, middle temporal, an 848 isotropic voxels) were used. A cortical thickness meta-ROI was calculated including
849 entorhinal, inferior temporal, middle temporal, and fusiform using FreeSurfer (v.6.0.
850 https://surfer.nmr.mgh.harvard.edu) par susceptible to AD-related atrophy 1 .

849 entorhinal, inferior temporal, middle temporal, and fusiform using FreeSurfer (v.6.0.
850 https://surfer.nmr.mgh.harvard.edu) parcellation, which are areas known to be
851 susceptible to AD-related atrophy⁷¹.
852 Me 850 https://surfer.nmr.mgh.harvard.edu) parcellation, which are areas known to be
851 susceptible to AD-related atrophy⁷¹.
852 Methodological details for imaging processing and quantification for the Knight-ADRC
853 coh 851 susceptible to AD-related atrophy⁷¹.
852 Methodological details for imaging p
853 cohort have been also previously rep
854 using Freesurfer (v.5.3) to generate
855 either $[1^1C]PIB$ or $[1^8F]$ florbetapir an
856 cer 852 Methodological details for imaging processing and quantification for the Knight-ADRC
853 cohort have been also previously reported^{71,72}. In brief, MPRAGE data were processed
854 using Freesurfer (v.5.3) to generate cohort have been also previously reported 71,72 853 cohort have been also previously reported^{71,72}. In brief, MPRAGE data were processed
854 using Freesurfer (v.5.3) to generate regions of interest. Amyloid-PET was acquired with
855 either [¹¹C]PIB or [¹⁸F]florbe 854 using Freesurfer (v.5.3) to generate regions of interest. Amyloid-PET was acquired with
855 either $[1^1C]PIB$ or $[1^8F]f$ lorbetapir and was quantified in a neocortical meta-ROI using
856 cerebellar grey as a referenc either 1^{11} C]PIB or 1^{18} F]florbetapir and was quantified in a neocortical meta-ROI using 855 either \lbrack ¹¹C]PIB or \lbrack ¹⁸F]florbetapir and was quantified in a neocortical meta-ROI using
856 cerebellar grey as a reference region. SUVR values were transformed to Centiloids⁷³ to
857 allow direct compar cerebellar grey as a reference region. SUVR values were transformed to Centiloids⁷³ to 856 cerebellar grey as a reference region. SUVR values were transformed to Centiloids⁷³ to
 857 allow direct comparison between tracers using previously validated transformations⁷⁴.
 $[18F]$ flortaucipir ($[18F]$ AV allow direct comparison between tracers using previously validated transformations⁷⁴. 857 allow direct comparison between tracers using previously validated transformations⁷⁴.
858 $\int_0^{18} F \cdot \int_0^{18} F \cdot \int_$ [¹⁸F]flortaucipir ([¹⁸ 858 [¹⁸F]flortaucipir ([¹⁸F]AV1451) was used as a tau-PET tracer and images were
23

quantified in the same temporal meta-ROI as in BioFINDER-2 and assessed as
860 positive if SUVR>1.32 based on previous work. The same additional regions as in
861 BioFINDER-2 were also used to quantify tau-PET binding in e 860 positive if SUVR>1.32 based on previous work. The same additional regions as in
861 BioFINDER-2 were also used to quantify tau-PET binding in early, intermediate, and
862 late tau deposition regions. In all cases, cere 861 BioFINDER-2 were also used to quantify tau-PET binding in early, intermediate, and
862 late tau deposition regions. In all cases, cerebellar grey was used as a reference
863 region. T1-weighted images were used to meas 862 late tau deposition regions. In all cases, cerebellar grey was used as a reference

863 region. T1-weighted images were used to measure cortical thickness using the same

864 approach as is in the BioFINDER-2 cohort.

863 region. T1-weighted images were used to measure cortical thickness using the same
864 approach as is in the BioFINDER-2 cohort.
865 **Neuropsychological testing**
867 **NEUROPSYCHOLOGIC SECTA** and Knight-ADRC participants 864 approach as is in the BioFINDER-2 cohort.
865 **Neuropsychological testing**
867 mPACC and a global cognitive composite
868 BioFINDER-2 and Knight-ADRC part
869 participants, the mPACC-5 composite w ---
866
867
868
869
870 866 **Neuropsychological testing**
867 mPACC and a global cognitive
868 BioFINDER-2 and Knight-
969 participants, the mPACC-5 c
870 Alzheimer's disease assessn
871 animal fluency, MMSE⁶⁶, and 867 mPACC and a global cognitive composite were used as the main cognitive outcome in
868 BioFINDER-2 and Knight-ADRC participants, respectively. In BioFINDER-2
869 participants, the mPACC-5 composite was calculated using 868 BioFINDER-2 and Knight-ADRC participants, respectively. In BioFINDER-2
869 participants, the mPACC-5 composite was calculated using mean of z-scores of
870 Alzheimer's disease assessment scale (ADAS) delayed recall (w 869 participants, the mPACC-5 composite was calculated using mean of z-scores of Alzheimer's disease assessment scale (ADAS) delayed recall (weighted double), animal fluency, MMSE⁶⁶, and trail making test-A (TMT-A)⁷⁵, 870 Alzheimer's disease assessment scale (ADAS) delayed recall (weighted double),
871 animal fluency, MMSE⁶⁶, and trail making test-A (TMT-A)⁷⁵, as a sensitive measure of
872 early cognitive impairment⁷⁶. Z-scores w animal fluency, MMSE⁶⁶, and trail making test-A $(TMT-A)^{75}$, as a sensitive measure of 871 animal fluency, MMSE⁶⁶, and trail making test-A (TMT-A)⁷⁵, as a sensitive measure of
872 early cognitive impairment⁷⁶. Z-scores were calculated with a group of CU- as
873 reference. Further, we also calculated s early cognitive impairment⁷⁶. Z-scores were calculated with a group of CU- as 872 early cognitive impairment⁷⁶. Z-scores were calculated with a group of CU- as
873 reference. Further, we also calculated several cognitive composites averaging z-scores
874 of different cognitive tests. For the memo 873 reference. Further, we also calculated several cognitive composites averaging z-scores
874 of different cognitive tests. For the memory composite we used ADAS delayed and
875 immediate word recall; for the executive f 874 of different cognitive tests. For the memory composite we used ADAS delayed and
875 immediate word recall; for the executive function composite, we used TMT-A, TMT-B
876 and symbol digit test; for the language composi 875 immediate word recall; for the executive function composite, we used TMT-A, TMT-B
876 and symbol digit test; for the language composite we used the animal fluency test and
877 the Boston naming test total score (BNT) the Boston naming test total score $(BNT)^{77}$; and, finally for the visuo-spatial composite

876 and symbol digit test; for the language composite we used the animal fluency test and
877 the Boston naming test total score (BNT)⁷⁷; and, finally for the visuo-spatial composite
878 we used the visual object and sp 877 the Boston naming test total score (BNT)"; and, finally for the visuo-spatial composite
878 we used the visual object and space perception (VOSP) cube and letters tests.
879 In Knight-ADRC, the global cognitive compos we used the visual object and space perception (VOSP) cube and letters tests.

879 In Knight-ADRC, the global cognitive composite was created using mean of z-s

880 free and cued selective reminding test (FCSRT) free recal 879 In Knight-ADRC, the global cognitive composite was created using mean of z-scores of
880 free and cued selective reminding test (FCSRT) free recall⁷⁸, animal fluency, TMT-A
881 and TMT-B. Z-scores were also calculate free and cued selective reminding test (FCSRT) free recall⁷⁸, animal fluency, TMT-A Free and cued selective reminding test (FCSRT) free recall⁷⁸, animal fluency, TMT-A
and TMT-B. Z-scores were also calculated from a CU- group as a reference. For the
executive function composite, we used TMT-A and TMT-B. 881 and TMT-B. Z-scores were also calculated from a CU- group as a reference. For the
882 executive function composite, we used TMT-A and TMT-B. We could not obtain any
883 other cognitive composite similar to those derive executive function composite, we used TMT-A and TMT-B. We could not obtain any

883 other cognitive composite similar to those derived in BioFINDER-2, due to lack of

884 similar tests. However, we selected individual test other cognitive composite similar to those derived in BioFINDER-2, due to lack of
884 similar tests. However, we selected individual tests to try to recapitulate similar
885 cognitive measures. For memory we used FCSRT, an 884 similar tests. However, we selected individual tests to try to recapitulate similar
885 cognitive measures. For memory we used FCSRT, and for language we used animal
886 fluency. No tests were available related visuo-s

885 cognitive measures. For memory we used FCSRT, and for language we used animal
886 fluency. No tests were available related visuo-spatial capacity.
887 **Model creation**
889 Model development was done with SuStaln⁴¹ Model creation

887 **Model creation**

888 **Model development** was done with SuStaln⁴¹ using cross-se

990 positive participants based on CSF A β 42/40 levels. We sel

891 because we wanted to create a staging model foc 888
889
890
891
892 888 **Model creation**
889 Model developm
890 positive participa
891 because we war
892 learning techniq
893 multiple difference Model development was done with SuStaIn 41 using cross-sectional data of amyloid-Model development was done with SuStaln⁴¹ using cross-sectional data of amyloid-

890 positive participants based on CSF A β 42/40 levels. We selected these participants

891 because we wanted to create a staging model 890 positive participants based on CSF Aβ42/40 levels. We selected these participants
891 because we wanted to create a staging model focused on AD. SuStaln is a machine-
892 learning technique that unravels temporal pro Because we wanted to create a staging model focused on AD. SuStaIn is a machine-

Bearning technique that unravels temporal progression patterns (stages) allowing for

multiple different trajectories (subtypes). For our p 892 learning technique that unravels temporal progression patterns (stages) allowing for
893 multiple different trajectories (subtypes). For our purpose, we used the event-based
894 model⁷⁹ (or mixture SuStaln⁸⁰), in 893 multiple different trajectories (subtypes). For our purpose, we used the event-based
894 model⁷⁹ (or mixture SuStaln⁸⁰), in which the input data is the probability of each
895 biomarker of being abnormal for each model⁷⁹ (or mixture SuStaln⁸⁰), in which the input data is the probability of each 894 model^{ry} (or mixture SuStaln⁸⁰), in which the input data is the probability of each
 895 biomarker of being abnormal for each participant. In our case, we used a Gaussian-
 24 895 biomarker of being abnormal for each participant. In our case, we used a Gaussian-
24

896 mixture modelling approach (with 2 Gaussians) to obtain these probabilities. With this
897 information, SuStaln provides the maximum likelihood sequence by which biomarkers
898 become abnormal, and gives a probability 897 information, SuStaIn provides the maximum likelihood sequence by which biomarkers
898 become abnormal, and gives a probability for this ordering, for all subtypes. The
899 number of SuStaIn stages is defined by the num 898 become abnormal, and gives a probability for this ordering, for all subtypes. The
899 number of SuStaln stages is defined by the number of biomarkers provided to the
900 model (*i.e.*, one per biomarker plus a biomarke 899 number of SuStaIn stages is defined by the number of biomarkers provided to the
900 model (*i.e.*, one per biomarker plus a biomarker negative stage). The selection of the
901 optimal number was determined using cross 900 model (*i.e.*, one per biomarker plus a biomarker negative stage). The selection of the
901 optimal number was determined using cross-validation, optimizing on cross-validation
902 based information criterion (CVIC) an 901 optimal number was determined using cross-validation, optimizing on cross-validation
902 based information criterion (CVIC) and out-of-sample log-likelihood was calculated.
903 The optimal number of subtypes was then 902 based information criterion (CVIC) and out-of-sample log-likelihood was calculated.
903 The optimal number of subtypes was then selected based on these criteria, using the
904 minimal number of subtypes that had the l 903 The optimal number of subtypes was then selected based on these criteria, using the
904 minimal number of subtypes that had the lowest CVIC and higher log-likelihood⁴¹. In
905 this study, we used pySuStaln⁸¹, a Py minimal number of subtypes that had the lowest CVIC and higher log-likelihood 41 . In this study, we used $pySuStaln⁸¹$, a Python implementation of the original method

904 minimal number of subtypes that had the lowest CVIC and higher log-likelihood⁴¹. In
905 this study, we used pySuStaln⁸¹, a Python implementation of the original method
906 (downloaded 08/2022).
907 In our initial 905 this study, we used pySuStaln⁸¹, a Python implementation of the original method
906 (downloaded 08/2022).
907 In our initial model with BioFINDER-2 participants, we included all biomarkers available
908 and performe 906 (downloaded 08/2022).

907 In our initial model with

908 and performed the cro

909 Although CVIC measur

910 likelihood in all three model

911 to its lower complexity⁸² 907 In our initial model with BioFINDER-2 participants, we included all biomarkers available
908 and performed the cross-correlation in models with one, two and three subtypes.
909 Although CVIC measures were lower in the 908 and performed the cross-correlation in models with one, two and three subtypes.
909 Although CVIC measures were lower in the three subtypes model, the similar log-
910 likelihood in all three models, supported the one 909 Although CVIC measures were lower in the three subtypes model, the similar log-
910 likelihood in all three models, supported the one subtype model as the optimal one due
911 to its lower complexity⁸². Upon examinin 910 likelihood in all three models, supported the one subtype model as the optimal one due
911 to its lower complexity⁸². Upon examining the outcome of this model, we observed that
912 pT217/T217, pT231/T231, and pT181/ to its lower complexity 8^2 . Upon examining the outcome of this model, we observed that 911 to its lower complexity⁸². Upon examining the outcome of this model, we observed that
912 pT217/T217, pT231/T231, and pT181/T181 position certainty was low, as they seemed
913 to compete for the second position (Ext 912 pT217/T217, pT231/T231, and pT181/T181 position certainty was low, as they seemed

913 to compete for the second position (Ext Data Fig. 1A). To avoid stages with low

914 certainty, we decided to try to optimize this 913 to compete for the second position (Ext Data Fig. 1A). To avoid stages with low
914 certainty, we decided to try to optimize this model through iterative removal of these
915 biomarkers. All the possible combinations 914 certainty, we decided to try to optimize this model through iterative removal of these
915 biomarkers. All the possible combinations were created (*i.e.*, removing pT217/T217
916 and/or pT231/T231 and/or pT181/T181) a 915 biomarkers. All the possible combinations were created (*i.e.*, removing pT217/T217 and/or pT231/T231 and/or pT181/T181) and compared using the CVIC (Ext Data Fig. 917 1B). We observed that models including only one of 916 and/or pT231/T231 and/or pT181/T181) and compared using the CVIC (Ext Data Fig.
917 1B). We observed that models including only one of these biomarkers (models 5-7)
918 were better than those including two (models 2-4 917 1B). We observed that models including only one of these biomarkers (models 5-7)
918 were better than those including two (models 2-4) or all three (model 1). Further,
919 models including pT181/T181 performed worse, 918 were better than those including two (models 2-4) or all three (model 1). Further,
919 models including pT181/T181 performed worse, and those including pT217/T217
920 performed better. Thus, the optimal model was sele 919 models including pT181/T181 performed worse, and those including pT217/T217
920 performed better. Thus, the optimal model was selected as that including only
921 pT217/T217 (model 7). Once the biomarkers to be included 920 performed better. Thus, the optimal model was selected as that including only
921 pT217/T217 (model 7). Once the biomarkers to be included in the model were
922 selected, we repeated the cross-validation with models up 921 pT217/T217 (model 7). Once the biomarkers to be included in the model were
922 selected, we repeated the cross-validation with models up to three subtypes.
923 Comparing CVIC and log-likelihood values, we once again se selected, we repeated the cross-validation with models up to three subtypes.

923 Comparing CVIC and log-likelihood values, we once again selected the one subtype

924 model as the optimal (Ext Data Fig. 1C). Based on this

923 Comparing CVIC and log-likelihood values, we once again selected the one subtype
924 model as the optimal (Ext Data Fig. 1C). Based on this cross-validated model, we then
925 staged all BioFINDER-2 participants.
926 Fo 924 model as the optimal (Ext Data Fig. 1C). Based on this cross-validated model, we then
925 staged all BioFINDER-2 participants.
926 For Knight-ADRC, we created the model from the biomarkers selected in the
927 BioFINDER 925 staged all BioFINDER-2 participants.
926 For Knight-ADRC, we created the
927 BioFINDER-2 cohort. To investigate v
928 we run SuStaln and created cross-va
929 based on CVIC (subtypes: 1=578.9;
930 29.4; 2=-30.4; 3=-31.0 926 For Knight-ADRC, we created the model from the biomarkers selected in the
927 BioFINDER-2 cohort. To investigate whether one subtype was also the selected model,
928 we run SuStaln and created cross-validated models fo 927 BioFINDER-2 cohort. To investigate whether one subtype was also the selected model,
928 we run SuStaln and created cross-validated models for one to three subtypes. Again,
929 based on CVIC (subtypes: 1=578.9; 2=598.1; 928 we run SuStaIn and created cross-validated models for one to three subtypes. Again,
929 based on CVIC (subtypes: 1=578.9; 2=598.1; 3=611.0) and log-likelihood (mean: 1=-
930 29.4; 2=-30.4; 3=-31.0), the less complex m 929 based on CVIC (subtypes: 1=578.9; 2=598.1; 3=611.0) and log-likelihood (mean: 1=-
930 29.4; 2=-30.4; 3=-31.0), the less complex model (*i.e.*, one subtype) was selected as the
931 optimal. We then staged all Knight-ADR 930 29.4; 2=-30.4; 3=-31.0), the less complex model (*i.e.*, one subtype) was selected as the
931 optimal. We then staged all Knight-ADRC participants based on this cross-validated
932 model. 931 optimal. We then staged all Knight-ADRC participants based on this cross-validated
932 model.
25 932 model.

933 As a sensitivity analysis, we compared the levels of the two biomarkers excluded (pT231/T231 and pT181/T181) with those of pT217/T217 in the optimal model. In summary, we observed that these biomarkers followed a simil 934 (pT231/T231 and pT181/T181) with those of pT217/T217 in the optimal model. In
935 summary, we observed that these biomarkers followed a similar trajectory across CSF
936 stages as pT217/T217, although with lower incre 935 summary, we observed that these biomarkers followed a similar trajectory across CSF
936 stages as pT217/T217, although with lower increases in the two cohorts (Ext Data Fig.
937 8A-B), which supports our decision of re 936 stages as pT217/T217, although with lower increases in the two cohorts (Ext Data Fig.
937 8A-B), which supports our decision of removing them from the model to have more
938 stable and independent stages.
940 Statistic

937 8A-B), which supports our decision of removing them from the model to have more
938 stable and independent stages.
940 **Statistical analyses**
941 All biomarkers were z-scored using participants older than 60 from the C 938 stable and independent stages.
939 **Statistical analyses**
941 All biomarkers were z-scored us
942 reference (BioFINDER-2: n=63
943 data was inverted such that his 940
941
942
943
944 940 **Statistical analyses**
941 All biomarkers were z
942 reference (BioFINDEI
943 data was inverted su
944 biomarkers. Differenc
945 Wallis test. Wilcoxon 941 All biomarkers were z-scored using participants older than 60 from the CU- group as a
942 reference (BioFINDER-2: n=63 and Knight-ADRC: n=71). When necessary, biomarker
943 data was inverted such that higher z-scores 942 reference (BioFINDER-2: n=63 and Knight-ADRC: n=71). When necessary, biomarker
943 data was inverted such that higher z-scores related to higher abnormality across all
944 biomarkers. Differences in biomarkers by CSF 943 data was inverted such that higher z-scores related to higher abnormality across all
944 biomarkers. Differences in biomarkers by CSF stages were assessed using Kruskal-
945 Wallis test. Wilcoxon-test was used for *pos* 944 biomarkers. Differences in biomarkers by CSF stages were assessed using Kruskal-
945 Wallis test. Wilcoxon-test was used for *post-hoc* comparisons adjusted for multiple
946 comparisons with false discovery rate (FDR) Wallis test. Wilcoxon-test was used for *post-hoc* comparisons adjusted for multiple

946 comparisons with false discovery rate (FDR) correction (only differences in consecutive

957 CSF stages are shown in figures). For c 946 comparisons with false discovery rate (FDR) correction (only differences in consecutive
947 CSF stages are shown in figures). For categorical data (*i.e.*, sex, *APOE* carriership and
948 diagnosis), we used chi-square 947 CSF stages are shown in figures). For categorical data (*i.e.*, sex, *APOE* carriership and
948 diagnosis), we used chi-squared tests. LOESS regressions were used to fit the
950 progression of biomarkers abnormalities 948 diagnosis), we used chi-squared tests. LOESS regressions were used to fit the
949 progression of biomarkers abnormalities across the CSF stages. ROC curves were
950 used to assess the utility of CSF stages for predicti progression of biomarkers abnormalities across the CSF stages. ROC curves were
950 used to assess the utility of CSF stages for predicting amyloid-PET, tau-PET positivity
951 and to compare AD to non-AD objective cognitive 950 used to assess the utility of CSF stages for predicting amyloid-PET, tau-PET positivity
951 and to compare AD to non-AD objective cognitive impairment (MCI or dementia states).
952 Maximization of Youden's index was us 951 and to compare AD to non-AD objective cognitive impairment (MCI or dementia states).
952 Maximization of Youden's index was used to select the optimal CSF stage cut-off in
953 each case (*pROC* and *cutpointr* package 952 Maximization of Youden's index was used to select the optimal CSF stage cut-off in

953 each case (pROC and *cutpointr* packages were used). For ordinal categories (*i.e.*, A/T

954 status and diagnosis), ordinal logi each case (*pROC* and *cutpointr* packages were used). For ordinal categories (*i.e.*, A/T
status and diagnosis), ordinal logistic regression models were used (*MASS* and *Imr*
packages). An equivalent measure to AUC, the 954 status and diagnosis), ordinal logistic regression models were used (*MASS* and *lmr* packages). An equivalent measure to AUC, the c-index, was used to assess the performance of the CSF staging⁸³. Confidence interval packages). An equivalent measure to AUC, the c-index, was used to assess the
956 performance of the CSF staging⁸³. Confidence intervals were calculated using
957 bootstrapping. Predicted probabilities of the outcome grou performance of the CSF staging⁸³ 956 performance of the CSF staging³⁵. Confidence intervals were calculated using
957 bootstrapping. Predicted probabilities of the outcome groups per each CSF stages
958 were calculated using the *predict* function, Long 957 bootstrapping. Predicted probabilities of the outcome groups per each CSF stages
958 were calculated using the *predict* function, Longitudinal rates of changes were
959 calculated for every participant using linear re 958 were calculated using the *predict* function, Longitudinal rates of changes were
959 calculated for every participant using linear regression models. One participant with a
960 very negative rate of change in amyloid-P 959 calculated for every participant using linear regression models. One participant with a
960 very negative rate of change in amyloid-PET (z-score< -1.8) was considered an outlier
961 by visual inspection and excluded f 960 very negative rate of change in amyloid-PET (z -score \lt -1.8) was considered an outlier
961 by visual inspection and excluded from the analysis. Kaplan-Meier curves were used to
962 assess clinical progression usin 961 by visual inspection and excluded from the analysis. Kaplan-Meier curves were used to
962 assess clinical progression using *survival* and *survminer* packages. Cox-proportional
963 hazards model were used to calculate 962 assess clinical progression using *survival* and *survminer* packages. Cox-proportional
963 hazards model were used to calculate the risk of clinical progression adjusting for age
964 and sex in all cases, and further 963 hazards model were used to calculate the risk of clinical progression adjusting for age
964 and sex in all cases, and further baseline clinical status if necessary.
965 All analyses were performed with R (v.4.1.0). A t

964 and sex in all cases, and further baseline clinical status if necessary.
965 All analyses were performed with R (v.4.1.0). A two-sided p v
966 considered statistically significant.
967 965 All analyses were performed with R (v.4.1.0). A two-sided p value < 0.05 was

966 considered statistically significant.

967

968 966 considered statistically significant.
967
968

968
|
|

970

971
972
973
974
975 971 **Table 1: BioFINDER-2 participants' characteristics**
972 Data is shown as mean(SD) unless otherwise stat
973 clinical diagnosis and amyloid status based on their CS
974 Only participants who progressed to MCI or dement 972 Data is shown as mean(SD) unless otherwise stated. Participants are divided by

973 clinical diagnosis and amyloid status based on their CSF Aβ42/40 levels (Aβ+: <0.080).

974 Only participants who progressed to MCI 973 clinical diagnosis and amyloid status based on their CSF Aβ42/40 levels (Aβ+: <0.080).

974 Only participants who progressed to MCI or dementia patients due to AD etiology were

975 considered to progress.

976 ^a,

974 Only participants who progressed to MCI or dementia patients due to AD etiology were

975 considered to progress.

976 ^a, 1 participant missing; ^b, 4 participants missing; ^c, 175 participants missing; ^d, 9

97 975 considered to progress.

976 ^a, 1 participant missing;

977 participants missing;

978 participants missing.

979 Abbreviations: Αβ, amy

980 dementia amyloid pos ^a, 1 participant missing; ^b, 4 participants missing; \degree , 175 participants missing; ^d participants missing; ^e, 6 participants missing; ^f, 4 participants missing; ^g

976 ^a, 1 participant missing; ^b, 4 participants missing; ^c, 175 participants missing; ^d, 9
977 participants missing; ^e, 6 participants missing; ^f, 4 participants missing; ⁹, 36
978 participants missing.
979 977 participants missing; ^e, 6 participants missing; ¹, 4 participants missing; ⁹, 36
978 participants missing.
979 Abbreviations: Aβ, amyloid-β; AD, Alzheimer's disease; ADD+, Alzheimer's disease
980 dementia amylo 978 participants missing.

979 Abbreviations: A β , a

980 dementia amyloid p

981 cognitively unimpaire

982 impairment amyloid

983 emission tomograph 979 Abbreviations: Aβ, amyloid-β; AD, Alzheimer's disease; ADD+, Alzheimer's disease
980 dementia amyloid positive; CU-, cognitively unimpaired amyloid negative; CU+,
981 cognitively unimpaired amyloid positive; CSF, cere 980 dementia amyloid positive; CU-, cognitively unimpaired amyloid negative; CU+,
981 cognitively unimpaired amyloid positive; CSF, cerebrospinal fluid; MCI, mild cognitive
982 impairment amyloid positive; nonAD, non-Alzhe 981 cognitively unimpaired amyloid positive; CSF, cerebrospinal fluid; MCI, mild cognitive
982 impairment amyloid positive; nonAD, non-Alzheimer's related disease; PET, positron
983 emission tomography; ROI, region of inte 982 impairment amyloid positive; nonAD, non-Alzheimer's related disease; PET, positron
983 emission tomography; ROI, region of interest; SD, standard deviation; SUVR,
984 standardized uptake value ratio.
985 983 emission tomography; ROI, region of interest; SD, standard deviation; SUVR,
984 standardized uptake value ratio.
985 984 standardized uptake value ratio.
985
.

986

0 (0%) $\left| \frac{14 (14.3\%)}{16 (66.7\%)} \right|$ -

articipants' characteristics

(SD) unless otherwise stated. Participants are divided

myloid status based on their CSF Aβ42/40 levels (A

dementia patients had a CDR=0.5 and mild 987
988
989
990
991
991 **Table 2: Knight-ADRC participants' characteristics**

988 Data is shown as mean(SD) unless otherwise state

989 clinical diagnosis and amyloid status based on th

990 <0.0673). Very mild AD dementia patients had a C

991 988 Data is shown as mean(SD) unless otherwise stated. Participants are divided by
989 clinical diagnosis and amyloid status based on their CSF Aβ42/40 levels (Aβ+:
990 <0.0673). Very mild AD dementia patients had a CDR= 989 clinical diagnosis and amyloid status based on their CSF Aβ42/40 levels (Aβ+:
990 <0.0673). Very mild AD dementia patients had a CDR=0.5 and mild AD dementia
991 patients had a CDR≥1, both with AD as etiology. Other 990 <0.0673). Very mild AD dementia patients had a CDR=0.5 and mild AD dementia
991 patients had a CDR≥1, both with AD as etiology. Other dementias group includes
992 participants with CDR>0 with non-AD etiology. Only par 991 patients had a CDR≥1, both with AD as etiology. Other dementias group includes
992 participants with CDR>0 with non-AD etiology. Only participants who progressed to
993 CDR≥0.5 or CDR≥1 due to AD etiology were consid

^a, 1 participant missing; ^b, 3 participants missing; \degree , 5 participants missing; ^d participants missing; $^{\rm e}$, 4 participants missing; $^{\rm f}$

992 participants with CDR>0 with non-AD etiology. Only participants who progressed to
993 CDR≥0.5 or CDR≥1 due to AD etiology were considered to progress.
994 ^a, 1 participant missing; ^b, 3 participants missing; ^c, 993 CDR≥0.5 or CDR≥1 due to AD etiology were considered to progress.

994 ^a, 1 participant missing; ^b, 3 participants missing; ^c, 5 participants participants missing; ^e, 4 participants missing; ^f, 8 participan 994 ^a, 1 participant missing; ^b, 3 participants missing; ⁶, 5 participants missing; ^a, 2
995 participants missing; ⁶, 4 participants missing; ^f, 8 participants missing.
996 Abbreviations: A β , amyloid- β ; participants missing; ^e, 4 participants missing; ^r, 8 participants missing.

996 Abbreviations: Aβ, amyloid-β; AD, Alzheimer's disease; ADD+, Alzh

997 dementia amyloid positive; CU-, cognitively unimpaired amyloid

9 996 Abbreviations: Aβ, amyloid-β; AD, Alzheimer's disease; ADD+, Alzheimer's disease
997 dementia amyloid positive; CU-, cognitively unimpaired amyloid negative; CU+,
698 cognitively unimpaired amyloid positive; CSF, cere 997 dementia amyloid positive; CU-, cognitively unimpaired amyloid negative; CU+,
998 cognitively unimpaired amyloid positive; CSF, cerebrospinal fluid; MCI, mild cognitive
999 impairment amyloid positive; mPACC, modified 998 cognitively unimpaired amyloid positive; CSF, cerebrospinal fluid; MCI, mild cognitive
999 impairment amyloid positive; mPACC, modified preclinical Alzheimer's cognitive
000 composite; nonAD, non-Alzheimer's related di 999 impairment amyloid positive; mPACC, modified preclinical Alzheimer's cognitive
000 composite; nonAD, non-Alzheimer's related disease; PET, positron emission
001 tomography; ROI, region of interest; SD, standard deviati 1000 composite; nonAD, non-Alzheimer's related disease; PET, positron emission
1001 tomography; ROI, region of interest; SD, standard deviation.
1002 1001 tomography; ROI, region of interest; SD, standard deviation.
1002
The standard deviation.

1003 **Figure captions**
1004 **Fig 1: CSF staging**
1005 **Description of the 4**
1006 model by CSF stag
1007 model is shown in
1008 on the results of 3 1004 **Fig 1: CSF staging model**
1005 Description of the CSF stag
1006 model by CSF stage. Cros
1007 model is shown in A. Bioma
1008 on the results of SuStaln.
1009 becoming abnormal at tha Description of the CSF staging model and the levels of the biomarkers included in the
1006 model by CSF stage. Cross-validated confusion matrix of the CSF biomarkers of the
1007 model is shown in A. Biomarkers are sorted b model by CSF stage. Cross-validated confusion matrix of the CSF biomarkers of the
1007 model is shown in A. Biomarkers are sorted by the time they become abnormal based
1008 on the results of SuStaln. Darkness represents t model is shown in A. Biomarkers are sorted by the time they become abnormal based
1008 on the results of SuStaln. Darkness represents the probability of that biomarker of
1009 becoming abnormal at that position, with black 1008 on the results of SuStaIn. Darkness represents the probability of that biomarker of
1009 becoming abnormal at that position, with black being 100%. Only amyloid-positive
1010 participants are included in this analysis becoming abnormal at that position, with black being 100%. Only amyloid-positive
1010 participants are included in this analysis. Individual biomarker levels by CSF stage in all
1011 BioFINDER-2 participants are shown in B participants are included in this analysis. Individual biomarker levels by CSF stage in all
1011 BioFINDER-2 participants are shown in B. CSF levels are z-scored based on a group
1012 of CU- participants (n=63) and all inc 1011 BioFINDER-2 participants are shown in B. CSF levels are z-scored based on a group
1012 of CU- participants (n=63) and all increases represent increase in abnormality.
1013 Significant differences in contiguous CSF sta 1012 of CU- participants (n=63) and all increases represent increase in abnormality.

1013 Significant differences in contiguous CSF stages are shown with asterisks. Horizontal

1014 line is drawn at z-score=1.96 which rep 1013 Significant differences in contiguous CSF stages are shown with asterisks. Horizontal

1014 line is drawn at z-score=1.96 which represents 95%Cl of the reference group (CU-).

1015 Smoothed LOESS lines of all CSF bi 1014 line is drawn at z-score=1.96 which represents 95%CI of the reference group (CU-).

1015 Smoothed LOESS lines of all CSF biomarkers are shown in C for comparison. CSF

1016 stage 0 represent being classified as norma Smoothed LOESS lines of all CSF biomarkers are shown in C for comparison. CSF
1016 stage 0 represent being classified as normal by the model. Black dots and vertical lines
1017 represent mean and SD by CSF stage, respecti Stage 0 represent being classified as normal by the model. Black dots and vertical lines

1017 represent mean and SD by CSF stage, respectively. *: p<0.05; **: p<0.01; ***: p<0.001.

1018 Abbreviations: Aβ, amyloid-β; Cl, 1017 represent mean and SD by CSF stage, respectively. *: p<0.05; **: p<0.01; ***: p<0.001.

1018 Abbreviations: Aβ, amyloid-β; CI, confidence interval; CU-, cognitively unimpaired

1019 amyloid negative; CSF, cerebrospi 1018 Abbreviations: Aβ, amyloid-β; CI, confidence interval; CU-, cognitively unimpaired
1019 amyloid negative; CSF, cerebrospinal fluid; LOESS, locally estimated scatterplot
1020 smoothing; MTBR, microtubule binding regio 2019 amyloid negative; CSF, cerebrospinal fluid; LOESS, locally estimated scatterplot

1020 smoothing; MTBR, microtubule binding region; pT, phosphorylated tau; SuStaln,

1021 subtype and stage inference.

1022 **Fig 2: AD**

smoothing; MTBR, microtubule binding region; pT, phosphorylated tau; SuStaIn,
1021 subtype and stage inference.
1022 **Fig 2: AD pathology, biomarkers and cognition by CSF stages**
1024 Depiction of individual biomarker leve 1021 subtype and stage inference.
1022 **Fig 2: AD pathology, biomai**
1024 Depiction of individual biomar
1025 stage in BioFINDER-2 particip
1026 PET) and tau (tau-PET in t 1022
1023
1024
1025
1026
1027
1028 **Fig 2: AD pathology, biomarkers and cognition by CSF stages**
1024 Depiction of individual biomarker levels, not used in the creation of
1025 stage in BioFINDER-2 participants (A). These include biomarkers
1026 PET) and ta Depiction of individual biomarker levels, not used in the creation of the model, by CSF
1025 stage in BioFINDER-2 participants (A). These include biomarkers of amyloid (amyloid-
1026 PET) and tau (tau-PET in the meta-tempo stage in BioFINDER-2 participants (A). These include biomarkers of amyloid (amyloid-
1026 PET) and tau (tau-PET in the meta-temporal ROI) pathologies, neurodegeneration
1027 (cortical thickness in the AD signature areas an 1026 PET) and tau (tau-PET in the meta-temporal ROI) pathologies, neurodegeneration
1027 (cortical thickness in the AD signature areas and CSF NfL) and cognition (mPACC).
1028 Biomarkers are z-scored based on a group of CU (cortical thickness in the AD signature areas and CSF NfL) and cognition (mPACC).
1028 Biomarkers are z-scored based on a group of CU- participants (n=63) and all increases
1029 represent increase in abnormality. Significa Biomarkers are z-scored based on a group of CU- participants (n=63) and all increases
1029 represent increase in abnormality. Significant differences in contiguous CSF stages are
1030 shown with asterisks. Horizontal line 1029 represent increase in abnormality. Significant differences in contiguous CSF stages are

1030 shown with asterisks. Horizontal line is drawn at z-score=1.96 which represents 95%Cl

1031 of the reference group (CU-). S shown with asterisks. Horizontal line is drawn at z-score=1.96 which represents 95%CI
1031 of the reference group (CU-). Smoothed LOESS lines of all AD biomarkers are shown
1032 in B for comparison. All participants with a 1031 of the reference group (CU-). Smoothed LOESS lines of all AD biomarkers are shown
1032 in B for comparison. All participants with available data were included in amyloid- and
1033 tau-PET analyses. For neurodegenerati in B for comparison. All participants with available data were included in amyloid- and

1033 tau-PET analyses. For neurodegeneration (cortical thickness and NfL) and cognitive

1034 (mPACC) measures, we excluded non-AD de tau-PET analyses. For neurodegeneration (cortical thickness and NfL) and cognitive
1034 (mPACC) measures, we excluded non-AD dementia patients to avoid bias. Of note,
1035 only few AD dementia cases had amyloid-PET availab 1034 (mPACC) measures, we excluded non-AD dementia patients to avoid bias. Of note,
1035 only few AD dementia cases had amyloid-PET available due to study design. CSF
1036 stage 0 represent being classified as normal by th only few AD dementia cases had amyloid-PET available due to study design. CSF
1036 stage 0 represent being classified as normal by the model. Black dots and vertical lines
1037 represent mean and SD per CSF stage, respecti stage 0 represent being classified as normal by the model. Black dots and vertical lines
1037 represent mean and SD per CSF stage, respectively. *: p<0.05; **: p<0.01; ***:
1038 p<0.001. 1037 represent mean and SD per CSF stage, respectively. *: p<0.05; **: p<0.01; ***:

1038 p<0.001. 1038 p<0.001.

Abbreviations: Aβ, amyloid-β; AD, Alzheimer's disease; CI, confidence interval; CU-, cognitively unimpaired amyloid negative; CSF, cerebrospinal fluid; LOESS, locally estimated scatterplot smoothing; mPACC, modified precl cognitively unimpaired amyloid negative; CSF, cerebrospinal fluid; LOESS, locally

1041 estimated scatterplot smoothing; mPACC, modified preclinical Alzheimer's cognitive

1042 composite; NfL, neurofilament light; PET, pos 1041 estimated scatterplot smoothing; mPACC, modified preclinical Alzheimer's cognitive

1042 composite; NfL, neurofilament light; PET, positron emission tomography; ROI, region of

1043 **Fig 3: CSF stages for predicting A**

2042 composite; NfL, neurofilament light; PET, positron emission tomography; ROI, region of

1043 interest.

1044 **Fig 3: CSF stages for predicting A/T status and cognitive stages**

1046 CSF stages for predicting pathologi 1043 interest.

1044

1045 **Fig 3: C:**

1046 CSF stag

1047 and for p

1048 the numb $\begin{array}{c} 1044 \\ 1045 \\ 1046 \\ 1047 \\ 1048 \\ 1049 \\ 1050 \end{array}$ **Fig 3: CSF stages for predicting A/T status and cognitive stages**
1046 CSF stages for predicting pathological status as measured with PET
1047 and for predicting cognitive stages and diagnostic groups in C-D. E
1048 the n CSF stages for predicting pathological status as measured with PET is shown in A-B,

1047 and for predicting cognitive stages and diagnostic groups in C-D. Barplots represent

1048 the number of participants in each categ and for predicting cognitive stages and diagnostic groups in C-D. Barplots represent

1048 the number of participants in each category per CSF stage. Numbers of participants in

1049 each category per CSF stage are shown the number of participants in each category per CSF stage. Numbers of participants in

1049 each category per CSF stage are shown within the barplots (A and C). In B and D,

1050 ROC curves were used to assess the classif each category per CSF stage are shown within the barplots (A and C). In B and D,
1050 ROC curves were used to assess the classification into dichotomic categories (Aβ-
1051 PET, tau-PET and AD vs non-AD cognitive impairme ROC curves were used to assess the classification into dichotomic categories (Aβ-
1051 PET, tau-PET and AD vs non-AD cognitive impairment), whereas ordinal logistic
1052 regressions were used for ordinal categories (A/T s 1051 PET, tau-PET and AD vs non-AD cognitive impairment), whereas ordinal logistic
1052 regressions were used for ordinal categories (A/T status and diagnosis). Heatmaps
1053 represent the predicted percentage of participa 1052 regressions were used for ordinal categories (A/T status and diagnosis). Heatmaps
1053 represent the predicted percentage of participants in each outcome category (A/T or
1054 diagnosis) by CSF stage. The most probab 1053 represent the predicted percentage of participants in each outcome category (A/T or
1054 diagnosis) by CSF stage. The most probable (highest percentage) category by CSF
1055 stage is framed in black. For ROC analyses, diagnosis) by CSF stage. The most probable (highest percentage) category by CSF
1055 stage is framed in black. For ROC analyses, AUCs, sensitivity and specificity measures
1056 from these analyses are shown in the plot. Th tage is framed in black. For ROC analyses, AUCs, sensitivity and specificity measures
1056 from these analyses are shown in the plot. The optimal cut-off in each case is shown
1057 as a vertical dashed line in A or C. An 1056 from these analyses are shown in the plot. The optimal cut-off in each case is shown
1057 as a vertical dashed line in A or C. An A-T+ participant (n=1) was excluded from the
1058 A/T status analysis. Non-AD dementi 2057 as a vertical dashed line in A or C. An A-T+ participant (n=1) was excluded from the

2058 A/T status analysis. Non-AD dementia cases were excluded from the cognitive stages

2059 analysis. And, only patients with o A/T status analysis. Non-AD dementia cases were excluded from the cognitive stages

analysis. And, only patients with objective impairment (MCI or dementia) were included

in the analyses of AD vs. non-AD. Aβ- and tau-PET

2059 analysis. And, only patients with objective impairment (MCI or dementia) were included

1060 in the analyses of AD vs. non-AD. Aβ- and tau-PET were assessed as positive based

1061 on previously validated cut-offs (1060 in the analyses of AD vs. non-AD. Aβ- and tau-PET were assessed as positive based

1061 on previously validated cut-offs (Aβ: SUVR>1.03, tau: SUVR>1.36).

2062 Abbreviations: Aβ, amyloid-β; AD, Alzheimer's disease; A 1061 on previously validated cut-offs (Aβ: SUVR>1.03, tau: SUVR>1.36).

1062 Abbreviations: Aβ, amyloid-β; AD, Alzheimer's disease; ADD+, A

1063 dementia amyloid-positive; A-T-, amyloid-negative tau-negative

1064 negat Abbreviations: Aβ, amyloid-β; AD, Alzheimer's disease; ADD+, Alzheimer's disease
1063 dementia amyloid-positive; A-T-, amyloid-negative tau-negative; A-T+, amyloid-
1064 negative tau-positive; A+T-, amyloid-positive tau-n dementia amyloid-positive; A-T-, amyloid-negative tau-negative; A-T+, amyloid-negative tau-positive; A+T-, amyloid-positive tau-negative; A+T+, amyloid-positive tau-

1065 positive; CSF, cerebrospinal fluid; CU-, cognitive negative tau-positive; A+T-, amyloid-positive tau-negative; A+T+, amyloid-positive tau-

1065 positive; CSF, cerebrospinal fluid; CU-, cognitively unimpaired amyloid-negative; CU+,

1066 cognitively unimpaired amyloid-posi positive; CSF, cerebrospinal fluid; CU-, cognitively unimpaired amyloid-negative; CU+,
1066 cognitively unimpaired amyloid-positive; MCI+, mild cognitive impairment amyloid-
1067 positive; PET, positron emission tomography cognitively unimpaired amyloid-positive; MCI+, mild cognitive impairment amyloid-
1067 positive; PET, positron emission tomography; ROC, receiver operating characteristic;
1068 ROI, region of interest; SUVR, standardized u

positive; PET, positron emission tomography; ROC, receiver operating characteristic;
1068 ROI, region of interest; SUVR, standardized uptake value ratio.
1070 **Fig 4: Longitudinal rate of change of AD biomarkers by CSF sta** ROI, region of interest; SUVR, standardized uptake value ratio.

1069
 Fig 4: Longitudinal rate of change of AD biomarkers by CSF

1071 Depiction of individual biomarker longitudinal rates of cha

1072 BioFINDER-2 partic 1069
1070
1071
1072
1073
1074 1070 **Fig 4: Longitudinal rate of change of AD biomarkers by CSF stages**
1071 Depiction of individual biomarker longitudinal rates of change by
1072 BioFINDER-2 participants (A). These include biomarkers of amyloid (am:
10 1071 Depiction of individual biomarker longitudinal rates of change by CSF stage in
1072 BioFINDER-2 participants (A). These include biomarkers of amyloid (amyloid-PET) and
1073 tau (tau-PET in the meta-temporal ROI) patho 1072 BioFINDER-2 participants (A). These include biomarkers of amyloid (amyloid-PET) and
1073 tau (tau-PET in the meta-temporal ROI) pathologies, neurodegeneration (cortical
1074 thickness in the AD signature) and cognitio 1073 tau (tau-PET in the meta-temporal ROI) pathologies, neurodegeneration (cortical
1074 thickness in the AD signature) and cognition (mPACC). Biomarkers are z-scored based
30 1074 thickness in the AD signature) and cognition (mPACC). Biomarkers are z-scored based

on a group of CU- participants (n=63) and all increases represent increase in
1076 abnormality. Rates of change were calculated with individual linear regression models.
1077 Significant differences in contiguous CSF stage abnormality. Rates of change were calculated with individual linear regression models.
1077 Significant differences in contiguous CSF stages are shown with asterisks. Smoothed
1078 LOESS lines of all AD biomarkers are show Significant differences in contiguous CSF stages are shown with asterisks. Smoothed
1078 LOESS lines of all AD biomarkers are shown in B for comparison. All participants were
1079 included in amyloid- and tau-PET analyses. LOESS lines of all AD biomarkers are shown in B for comparison. All participants were

1079 included in amyloid- and tau-PET analyses. For neurodegeneration (cortical thickness)

1080 and cognitive (MMSE) measures, we excl 1079 included in amyloid- and tau-PET analyses. For neurodegeneration (cortical thickness)

1080 and cognitive (MMSE) measures, we excluded non-AD dementia patients to avoid bias.

1081 CSF stage 0 represent being classif 2080 and cognitive (MMSE) measures, we excluded non-AD dementia patients to avoid bias.

1081 CSF stage 0 represent being classified as normal by the model. Black dots and vertical

1082 lines represent mean and SD per CS

1081 CSF stage 0 represent being classified as normal by the model. Black dots and vertical

1082 lines represent mean and SD per CSF stage, respectively. *: p<0.05; **: p<0.01; ***:

1083 p<0.001.

4084 Abbreviations: Aβ 2082 lines represent mean and SD per CSF stage, respectively. *: p<0.05; **: p<0.01; ***:

1083 p<0.001.

4084 Abbreviations: Aβ, amyloid-β; AD, Alzheimer's disease; Cl, confidence interval; CU-,

1085 cognitively unimpai 1083 p<0.001.
1084 Abbreviat
1085 cognitivel
1086 estimated
1087 composite Abbreviations: Aβ, amyloid-β; AD, Alzheimer's disease; CI, confidence interval; CU-, cognitively unimpaired amyloid negative; CSF, cerebrospinal fluid; LOESS, locally estimated scatterplot smoothing; mPAC, modified precli cognitively unimpaired amyloid negative; CSF, cerebrospinal fluid; LOESS, locally

1086 estimated scatterplot smoothing; mPAC, modified preclinical Alzheimer's cognitive

1087 composite; PET, positron emission tomography;

estimated scatterplot smoothing; mPAC, modified preclinical Alzheimer's cognitive
1087 composite; PET, positron emission tomography; ROI, region of interest.
1088 **Fig 5: CSF stages for predicting clinical progression**
109 composite; PET, positron emission tomography; ROI, region of interest.

1088
 Fig 5: CSF stages for predicting clinical progression

1090 Higher CSF stages groups (4-5) show higher HR of clinical progressic

1091 Iower p 1088
1089
1090
1091
1092
1093
1094 Fig 5: CSF stages for predicting clinical progression
1090 Higher CSF stages groups (4-5) show higher HR of clin
1091 Iower positive stages (reference: 1-3). Progression from
1092 dementia is shown in A-B. Progression from Higher CSF stages groups (4-5) show higher HR of clinical progression compared to
1091 lower positive stages (reference: 1-3). Progression from CU or MCI at baseline to AD
1092 dementia is shown in A-B. Progression from MC 1091 lower positive stages (reference: 1-3). Progression from CU or MCI at baseline to AD

1092 dementia is shown in A-B. Progression from MCI at baseline to AD dementia is shown

1093 in C-D. Progression from CU at baseli dementia is shown in A-B. Progression from MCI at baseline to AD dementia is shown
1093 in C-D. Progression from CU at baseline to MCI is shown in E-F. Kaplan-Meier curves,
1094 as well as the number of participants per gr in C-D. Progression from CU at baseline to MCI is shown in E-F. Kaplan-Meier curves,
1094 as well as the number of participants per group and timepoint are shown in A, C and E,
1095 respectively. Cox-proportional hazards m as well as the number of participants per group and timepoint are shown in A, C and E,
1095 respectively. Cox-proportional hazards models were used to calculate HR[95%CI] of
1096 higher CSF stages (4-5) compared to the ref respectively. Cox-proportional hazards models were used to calculate HR[95%CI] of
1096 higher CSF stages (4-5) compared to the reference (1-3, B, D and F). These analyses
were adjusted for age, sex in all cases, and additi 1096 higher CSF stages (4-5) compared to the reference (1-3, B, D and F). These analyses

1097 were adjusted for age, sex in all cases, and additionally for clinical status at baseline

1098 (CU or MCI) if appropriate.

41

were adjusted for age, sex in all cases, and additionally for clinical status at baseline

1098 (CU or MCI) if appropriate.

1099 Abbreviations: AD, Alzheimer's disease; CI, confidence interval; CSF, cerebrospinal

1100 fl 1098 (CU or MCI) if appropriate.
1099 Abbreviations: AD, Alzheir
1100 fluid; HR, hazard ratios; MC
1101 **Fig 6: Replication of main**
1103 Description of the model is

Abbreviations: AD, Alzheimer's disease; CI, confidence interval; CSF, cerebrospinal
1100 fluid; HR, hazard ratios; MCI, mild cognitive impairment.
1101 **Fig 6: Replication of main analyses in Knight-ADRC participants**
1103 1100 fluid; HR, hazard ratios; MCI, mild cognitive impairment.
1101 **Fig 6: Replication of main analyses in Knight-ADRC**
1103 Description of the model is shown in A-B. Cross-validate
1104 biomarkers of the model is shown i 1101
1102
1103
1104
1105
1106
1107 **Fig 6: Replication of main analyses in Knight-ADRC participants**
1103 Description of the model is shown in A-B. Cross-validated confusion
1104 biomarkers of the model is shown in A. Biomarkers are sorted by the
1105 abnor Description of the model is shown in A-B. Cross-validated confusion matrix of the CSF
1104 biomarkers of the model is shown in A. Biomarkers are sorted by the time they become
1105 abnormal based on the results of SuStaln. biomarkers of the model is shown in A. Biomarkers are sorted by the time they become
1105 abnormal based on the results of SuStaln. Darkness represents the probability of that
1106 biomarker of becoming abnormal at that po abnormal based on the results of SuStaIn. Darkness represents the probability of that
1106 biomarker of becoming abnormal at that position, with black being 100%. Only amyloid-
1107 positive participants are included in th biomarker of becoming abnormal at that position, with black being 100%. Only amyloid-

1107 positive participants are included in this analysis. Description of the CSF levels of the

1108 biomarkers included in the model b positive participants are included in this analysis. Description of the CSF levels of the
1108 biomarkers included in the model by CSF stage are shown in B for all Knight-ADRC
1109 participants. Depiction of individual bio biomarkers included in the model by CSF stage are shown in B for all Knight-ADRC
1109 participants. Depiction of individual biomarker levels, not used in the creation of the
1110 model, by CSF stage are shown in C. These i 1109 participants. Depiction of individual biomarker levels, not used in the creation of the
1110 model, by CSF stage are shown in C. These include biomarkers of amyloid (amyloid-
1111 PET) and tau (tau-PET in the meta-tem 1110 model, by CSF stage are shown in C. These include biomarkers of amyloid (amyloid-
1111 PET) and tau (tau-PET in the meta-temporal ROI) pathologies, neurodegeneration
31 1111 PET) and tau (tau-PET in the meta-temporal ROI) pathologies, neurodegeneration
31

(cortical thickness in the AD signature areas and CSF NfL) and cognition (global
1113 cognitive composite). CSF and AD biomarker levels are z-scored based on a group of
1114 CU- participants (n=71) and all increases repres cognitive composite). CSF and AD biomarker levels are z-scored based on a group of
1114 CU- participants (n=71) and all increases represent increase in abnormality. Horizontal
1115 line is drawn at z-score=1.96 which repre 1114 CU- participants (n=71) and all increases represent increase in abnormality. Horizontal
1115 line is drawn at z-score=1.96 which represents 95%CI of the reference group (CU-).
1116 CSF stage 0 represent being classifi 1115 line is drawn at z-score=1.96 which represents 95%CI of the reference group (CU-).
1116 CSF stage 0 represent being classified as normal by the model. Prediction of amyloid-
1117 PET (D-G), tau-PET (E-H) and A/T statu 1116 CSF stage 0 represent being classified as normal by the model. Prediction of amyloid-
1117 PET (D-G), tau-PET (E-H) and A/T status (by PET, F-I) are shown next. Number of
1118 participants in each category are colored 1117 PET (D-G), tau-PET (E-H) and A/T status (by PET, F-I) are shown next. Number of
1118 participants in each category are colored in D, E and F. Numbers of participants in
1119 each category per CSF stage are shown withi participants in each category are colored in D, E and F. Numbers of participants in

1119 each category per CSF stage are shown within the barplots. In G and H, ROC curves

1120 were used to determine the CSF stage to opti each category per CSF stage are shown within the barplots. In G and H, ROC curves
1120 were used to determine the CSF stage to optimally classify participants into
1121 positive/negative in each case. AUCs, sensitivity an were used to determine the CSF stage to optimally classify participants into
1121 positive/negative in each case. AUCs, sensitivity and specificity measures from these
1122 analyses are shown in the plot. The optimal cut-o positive/negative in each case. AUCs, sensitivity and specificity measures from these
1122 analyses are shown in the plot. The optimal cut-off in each case is shown as a vertical
1123 dashed line in D and E, respectively. analyses are shown in the plot. The optimal cut-off in each case is shown as a vertical
1123 dashed line in D and E, respectively. Ordinal logistic regression was used for assessing
1124 A/T status (I). The heatmap represe dashed line in D and E, respectively. Ordinal logistic regression was used for assessing
1124 A/T status (I). The heatmap represent the predicted percentage of participants in each
1125 A/T group per CSF stage. The most pr A/T status (I). The heatmap represent the predicted percentage of participants in each

1125 A/T group per CSF stage. The most probable (highest percentage) group per CSF

1126 stage is framed in black. Amyloid-PET was con A/T group per CSF stage. The most probable (highest percentage) group per CSF
1126 stage is framed in black. Amyloid-PET was considered positive if Centiloids>20, tau-
1127 PET was considered positive if SUVR at meta-tempo 1126 stage is framed in black. Amyloid-PET was considered positive if Centiloids>20, tau-
1127 PET was considered positive if SUVR at meta-temporal ROI was higher than 1.32. An
1128 A-T+ participant (n=1) was excluded fro 1127 PET was considered positive if SUVR at meta-temporal ROI was higher than 1.32. An A-T+ participant (n=1) was excluded from the A/T status analysis. Higher CSF stages groups (4-5) show higher HR of clinical progressio 1128 A-T+ participant (n=1) was excluded from the A/T status analysis. Higher CSF stages

1129 groups (4-5) show higher HR of clinical progression compared to lower stages

1130 (reference: 1-3, J-M). Progression from CDR 1129 groups (4-5) show higher HR of clinical progression compared to lower stages

1130 (reference: 1-3, J-M). Progression from CDR=0 or CDR=0.5 at baseline to CDR≥1 is

1131 shown in J-K. Progression from CDR=0 at baseli 1130 (reference: 1-3, J-M). Progression from CDR=0 or CDR=0.5 at baseline to CDR≥1 is

1131 shown in J-K. Progression from CDR=0 at baseline to CDR≥0.5 is shown in L-M.

1132 Kaplan-Meier curves, as well as the number of 1131 shown in J-K. Progression from CDR=0 at baseline to CDR≥0.5 is shown in L-M.

1132 Kaplan-Meier curves, as well as the number of participants per group and timepoint are

1133 shown in J and L. Cox-proportional hazar Example 1132 Kaplan-Meier curves, as well as the number of participants per group and timepoint are

1133 shown in J and L. Cox-proportional hazards models were used to calculate HR[95%Cl]

1134 of higher CSF stages (4-5) Shown in J and L. Cox-proportional hazards models were used to calculate HR[95%CI]
1134 of higher CSF stages (4-5) compared to the reference (1-3, K and M). These analyses
1135 were adjusted for age, sex in all cases, and

1134 of higher CSF stages (4-5) compared to the reference (1-3, K and M). These analyses

1135 were adjusted for age, sex in all cases, and additionally for clinical status at baseline

1136 (CDR=0 or CDR=1) if appropriat were adjusted for age, sex in all cases, and additionally for clinical status at baseline

1136 (CDR=0 or CDR=1) if appropriate.

1137 Abbreviations: Αβ, amyloid-β; AD, Alzheimer's disease; AUC, area under the curve;

113 1136 (CDR=0 or CDR=1) if appropriate.

1137 Abbreviations: Aβ, amyloid-β; AD,

1138 CDR, clinical dementia rating; C

1139 amyloid negative; CSF, cerebrospi

1140 scatterplot smoothing; MMSE, M

1141 binding region; NfL, Abbreviations: Aβ, amyloid-β; AD, Alzheimer's disease; AUC, area under the curve;

CDR, clinical dementia rating; Cl, confidence interval; CU-, cognitively unimpaired

amyloid negative; CSF, cerebrospinal fluid; HR, hazar 1138 CDR, clinical dementia rating; CI, confidence interval; CU-, cognitively unimpaired

1139 amyloid negative; CSF, cerebrospinal fluid; HR, hazard ratio; LOESS, locally estimated

1140 scatterplot smoothing; MMSE, Mini-1139 amyloid negative; CSF, cerebrospinal fluid; HR, hazard ratio; LOESS, locally estimated

1140 scatterplot smoothing; MMSE, Mini-mental state examination; MTBR, microtubule

1141 binding region; NfL, neurofilament light scatterplot smoothing; MMSE, Mini-mental state examination; MTBR, microtubule

1141 binding region; NfL, neurofilament light; PET, positron emission tomography; ROC,

1142 receiver operating characteristic; SuStaln, subtyp inding region; NfL, neurofilament light; PET, positron emission tomography; ROC,
1142 receiver operating characteristic; SuStaln, subtype and stage inference; SUVR,
1143 standardized uptake value ratio.
1145 **Extended data** 1142 receiver operating characteristic; SuStaIn, subtype and stage inference; SUVR,
1143 standardized uptake value ratio.
1144 **Extended data**
1146 **Ext Data Fig 1: Creation and optimization of the model**
1147 Initial mod

1143 standardized uptake value ratio.
1144 **Extended data**
1145 **Ext Data Fig 1: Creation and of**
1147 Initial model with all CSF biomarl
1148 pT205/T205, MTBR-tau243 and 1145
1146
1147
1148 1145 **Extended data**

1146 **Ext Data Fig 1: C**

1147 **Initial model with interpediate process**

1148 **pT205/T205**, MTE

1146 **Ext Data Fig 1: Creation and optimization of the model**
1147 Initial model with all CSF biomarkers (Aβ42/40, pT217/T21
1148 pT205/T205, MTBR-tau243 and total-tau) is shown in A.

1147 Initial model with all CSF biomarkers (Aβ42/40, pT217/T217, pT231/T231, pT181/T181,
1148 pT205/T205, MTBR-tau243 and total-tau) is shown in A. First two columns represent
32 1148 pT205/T205, MTBR-tau243 and total-tau) is shown in A. First two columns represent

the statistics, CVIC and log-likelihood, of this model for one, two and three subtypes.
1150 Each dot in log-likelihood plot represents one of the ten cross-validation sets of data.
1151 Lower CVIC and higher log-likelihoo Each dot in log-likelihood plot represents one of the ten cross-validation sets of data.
1151 Lower CVIC and higher log-likelihood values represent better performance of the
1152 model. Although higher number of subtypes h Lower CVIC and higher log-likelihood values represent better performance of the
1152 model. Although higher number of subtypes had higher CVIC, the comparable log-
1153 likelihood across subtypes suggests that one subtype model. Although higher number of subtypes had higher CVIC, the comparable log-
1153 likelihood across subtypes suggests that one subtype is complex enough to explain the
1154 variability observed in the data. Cross-valida later 1153 likelihood across subtypes suggests that one subtype is complex enough to explain the
1154 variability observed in the data. Cross-validated confusion matrix of the one subtype
1155 model is shown in the last co variability observed in the data. Cross-validated confusion matrix of the one subtype
1155 model is shown in the last column. Here, biomarkers are sorted by the time they
1156 become abnormal based on the results of SuSta model is shown in the last column. Here, biomarkers are sorted by the time they
1156 become abnormal based on the results of SuStaln. Darkness represents the probability
1157 of that biomarker of becoming abnormal at that 1156 become abnormal based on the results of SuStaIn. Darkness represents the probability
1157 of that biomarker of becoming abnormal at that position, with black being 100%. Given
1158 that some biomarkers (pT217/T217, p 1157 of that biomarker of becoming abnormal at that position, with black being 100%. Given
1158 that some biomarkers (pT217/T217, pT231/T231 and pT181/T181) show high overlap
1159 on the ordering, we optimized the model b 1158 that some biomarkers (pT217/T217, pT231/T231 and pT181/T181) show high overlap
1159 on the ordering, we optimized the model by removing these biomarkers systematically
1160 (B). All models without one or two of these on the ordering, we optimized the model by removing these biomarkers systematically
1160 (B). All models without one or two of these biomarkers were tested (models 2 to 7).
1161 CVIC (left) and cross-validated confusion ma 1160 (B). All models without one or two of these biomarkers were tested (models 2 to 7).

1161 CVIC (left) and cross-validated confusion matrixes (right) for each of these models are

1162 shown in B, respectively. CVIC sh 1161 CVIC (left) and cross-validated confusion matrixes (right) for each of these models are

1162 shown in B, respectively. CVIC shows that the optimal model was that excluding both

1163 pT231/T231 and pT181/T181 (mode Shown in B, respectively. CVIC shows that the optimal model was that excluding both
1163 pT231/T231 and pT181/T181 (model 7, shown in C). Both CVIC and log-likelihood
1164 measures show that one subtype was the optimal mo pT231/T231 and pT181/T181 (model 7, shown in C). Both CVIC and log-likelihood
1164 measures show that one subtype was the optimal model when using this set of
1165 biomarkers.
1166 Abbreviations: Aβ, amyloid-β; CVIC, cros

measures show that one subtype was the optimal model when using this set of
1165 biomarkers.
1166 Abbreviations: Aβ, amyloid-β; CVIC, cross-validation information criterion; MTBR,
1167 microtubule binding region; pT, phos 1165 biomarkers.
1166 Abbreviation
1167 microtubule
1168 inference.
1169
1170 Abbreviations: Aβ, amyloid-β; CVIC, cross-validation information criterion; MTBR,
1167 microtubule binding region; pT, phosphorylated tau; SuStaln, subtype and stage
1168 inference.
1170 Ext Data Fig 2: Demographic, genet microtubule binding region; pT, phosphorylated tau; SuStaIn, subtype and stage
1168 inference.
1169
1171 **Ext Data Fig 2: Demographic, genetic and clinical characteristics by CSF stage**
1172 Depiction of basic characterist

1168 inference.
1169
1170
1171 **Ext Data F**
1172 Depiction
1173 CSF stage ----
1170
1171
1172
1173
1174 1171
1172
1173
1174
1175 **Ext Data Fig 2: Demographic, genetic and clinical characteristics by CSF stage**
1172 Depiction of basic characteristics of BioFINDER-2 (A-E) and Knight-ADRC (F-J)
1173 CSF stage. Kruskal-Wallis or chi-square tests were us Depiction of basic characteristics of BioFINDER-2 (A-E) and Knight-ADRC (F-J) by
1173 CSF stage. Kruskal-Wallis or chi-square tests were used to investigate the association
1174 between each of these characteristics and CS CSF stage. Kruskal-Wallis or chi-square tests were used to investigate the association
1174 between each of these characteristics and CSF stages. P-values of these tests are
1175 shown at the top right of each subplot. Num

between each of these characteristics and CSF stages. P-values of these tests are
1175 shown at the top right of each subplot. Number of individuals in each category are
1176 shown inside the barplots.
1177 Abbreviations: shown at the top right of each subplot. Number of individuals in each category are

1176 shown inside the barplots.

1177 Abbreviations: AD, Alzheimer's disease; ADD+, Alzheimer's disease dementia amyloid

1178 positive; C 1176 shown inside the barplots.

1177 Abbreviations: AD, Alzhein

1178 positive; CU-, cognitively

1179 amyloid positive; CSF, ce

1180 positive; nonAD, non-Alzh

1181 dementia. 1177 Abbreviations: AD, Alzheimer's disease; ADD+, Alzheimer's disease dementia amyloid

1178 positive; CU-, cognitively unimpaired amyloid negative; CU+, cognitively unimpaired

1179 amyloid positive; CSF, cerebrospinal f 1178 positive; CU-, cognitively unimpaired amyloid negative; CU+, cognitively unimpaired

1179 amyloid positive; CSF, cerebrospinal fluid; MCI+, mild cognitive impairment amyloid

1180 positive; nonAD, non-Alzheimer's rela 1179 amyloid positive; CSF, cerebrospinal fluid; MCI+, mild cognitive impairment amyloid
1180 positive; nonAD, non-Alzheimer's related disease; other Dem, non-Alzheimer's type
1181 dementia.
1182
1183 **Ext Data Fig 3: Mode** 1180 positive; nonAD, non-Alzheimer's related disease; other Dem, non-Alzheimer's type
1181 dementia.
1182
1183 **Ext Data Fig 3: Model stability** 1181 dementia.
1182
1183
1184 **Ext Data F**

1183
1184 1183 1184 **Ext Data Fig 3: Model stability**

1185 Depiction of the evolution of CSF stages in BioFINDER-2 (n=220, A-B) and Knight-

1186 ADRC participants (n=51, C-D) with longitudinal CSF available. As longitudinal CSF

1187 Aβ42/40 levels were not available for a ADRC participants (n=51, C-D) with longitudinal CSF available. As longitudinal CSF $\widehat{A}\beta42/40$ levels were not available for any BioFINDER-2 participant, we imputed this data with their baseline levels. We show the nu 1187 Aβ42/40 levels were not available for any BioFINDER-2 participant, we imputed this

1188 data with their baseline levels. We show the number of progressors, regressors and

1189 stable participants in A and C, for e 1188 data with their baseline levels. We show the number of progressors, regressors and

1189 stable participants in A and C, for each cohort respectively. In B and D, we further

1190 show the CSF stages at follow-up. Fo 1189 stable participants in A and C, for each cohort respectively. In B and D, we further

1190 show the CSF stages at follow-up. For those Knight-ADRC with more than one

1191 longitudinal visit we took the one more dist 1190 show the CSF stages at follow-up. For those Knight-ADRC with more than one

1191 longitudinal visit we took the one more distant from the baseline.

1192 Abbreviations: A β , amyloid- β ; CSF, cerebrospinal fluid.

1191 longitudinal visit we took the one more distant from the baseline.

1192 Abbreviations: Aβ, amyloid-β; CSF, cerebrospinal fluid.

1193 Ext Data Fig 4: Tau-PET binding in different Braak regions by

1196 Depiction of 1192 Abbreviations: Aβ, amyloid-β; CSF, cerebrospinal fluid.
1193 **Ext Data Fig 4: Tau-PET binding in different Braak re**
1196 Depiction of tau-PET binding in different areas of tau
1197 BioFINDER-2 (A) and Knight-ADRC p 1194
1195
1196
1197
1198 1195
1196
1197
1198
1199 **Ext Data Fig 4: Tau-PET binding in different Braak regions by CSF stages**
1196 Depiction of tau-PET binding in different areas of tau deposition, by CSF sta
1197 BioFINDER-2 (A) and Knight-ADRC participants (B). These are Depiction of tau-PET binding in different areas of tau deposition, by CSF stage in all
1197 BioFINDER-2 (A) and Knight-ADRC participants (B). These areas include regions of
1198 early (Braak I-II), intermediate (Braak II-I BioFINDER-2 (A) and Knight-ADRC participants (B). These areas include regions of early (Braak I-II), intermediate (Braak II-IV) and late (Braak V-VI) tau deposition. Tau-
PET levels are z-scored based on a group of CU- par early (Braak I-II), intermediate (Braak II-IV) and late (Braak V-VI) tau deposition. Tau-
1199 PET levels are z-scored based on a group of CU- participants (BioFINDER-2: n=63
1200 and Knight-ADRC: n=71) and all increases r PET levels are z-scored based on a group of CU- participants (BioFINDER-2: n=63
1200 and Knight-ADRC: n=71) and all increases represent increase in abnormality.
1201 Significant differences in contiguous CSF stages are sho and Knight-ADRC: n=71) and all increases represent increase in abnormality.

1201 Significant differences in contiguous CSF stages are shown with asterisks. Horizontal

1202 line is drawn at z-score=1.96 which represents 9 Significant differences in contiguous CSF stages are shown with asterisks. Horizontal

1202 line is drawn at z-score=1.96 which represents 95%Cl of the reference group (CU-).

1203 Smoothed LOESS lines of all AD biomarker line is drawn at z-score=1.96 which represents 95%CI of the reference group (CU-).

1203 Smoothed LOESS lines of all AD biomarkers are shown in B for comparison. CSF

1204 stage 0 represent being classified as normal by t

Smoothed LOESS lines of all AD biomarkers are shown in B for comparison. CSF
1204 stage 0 represent being classified as normal by the model. *: p<0.05; **: p<0.01; ***:
1205 p<0.001.
Abbreviations: Aβ, amyloid-β; AD, Alzh the model. *: p<0.05; **: p<0.01; ***:

1205 p<0.001.

Abbreviations: Aβ, amyloid-β; AD, Alzheimer's disease; Cl, confidence interval; CU-,

cognitively unimpaired amyloid negative; CSF, cerebrospinal fluid; LOESS, locall 1205 p<0.001.
1206 Abbreviat
1207 cognitivel
1208 estimated
1209 interest.
1210 Abbreviations: Aβ, amyloid-β; AD, Alzheimer's disease; CI, confidence interval; CU-, cognitively unimpaired amyloid negative; CSF, cerebrospinal fluid; LOESS, locally estimated scatterplot smoothing; PET, positron emissio cognitively unimpaired amyloid negative; CSF, cerebrospinal fluid; LOESS, locally

1208 estimated scatterplot smoothing; PET, positron emission tomography; ROI, region of

1210

1211 Ext Data Fig 5: Cognitive composites by estimated scatterplot smoothing; PET, positron emission tomography; ROI, region of
1209 interest.
1210
1211 **Ext Data Fig 5: Cognitive composites by CSF stages**
1213 Depiction of different cognitive measures, by CSF stage

1209 interest.
1210
1211
1212 **Ext Data**
1213 Depiction
1214 Knight-A ----
1211
1212
1213
1215
1216 1212
1213
1214
1215
1216 **Ext Data Fig 5: Cognitive composites by CSF stages**
1213 Depiction of different cognitive measures, by CSF state
1214 Knight-ADRC participants (B). These measures inclu
1215 animal fluency, MMSE and TMT-A), memory (ADAS-d 1213 Depiction of different cognitive measures, by CSF stage in BioFINDER-2 (A) and
1214 Knight-ADRC participants (B). These measures include: mPACC (ADAS-delayed,
1215 animal fluency, MMSE and TMT-A), memory (ADAS-delayed 1214 Knight-ADRC participants (B). These measures include: mPACC (ADAS-delayed, animal fluency, MMSE and TMT-A), memory (ADAS-delayed and ADAS-immediate), executive function (TMT-A, TMT-B and symbols digit), language (anim 215 animal fluency, MMSE and TMT-A), memory (ADAS-delayed and ADAS-immediate),
1216 executive function (TMT-A, TMT-B and symbols digit), language (animal fluency and
1217 BNT-15) and visuospatial (VOSP-cube and VOSP-incomp executive function (TMT-A, TMT-B and symbols digit), language (animal fluency and
1217 BNT-15) and visuospatial (VOSP-cube and VOSP-incomplete) for BioFINDER-2. For
1218 Knight-ADRC we had a global cognitive composite (FCS 1217 BNT-15) and visuospatial (VOSP-cube and VOSP-incomplete) for BioFINDER-2. For
1218 Knight-ADRC we had a global cognitive composite (FCSRT, animals, TMT-A and TMT-
1219 B), an executive function composite (TMT-A and TM 1218 Knight-ADRC we had a global cognitive composite (FCSRT, animals, TMT-A and TMT-
1219 B), an executive function composite (TMT-A and TMT-B), a memory (FCSRT) and
1220 language (animal fluency) tests. Cognitive scores a 1219 B), an executive function composite (TMT-A and TMT-B), a memory (FCSRT) and
1220 Ianguage (animal fluency) tests. Cognitive scores are z-scored based on a group of
1221 CU- participants (BioFINDER-2: n=60 and Knight-A 1220 language (animal fluency) tests. Cognitive scores are z-scored based on a group of
1221 CU- participants (BioFINDER-2: n=60 and Knight-ADRC: n=71) and all increases
34 1221 CU- participants (BioFINDER-2: n=60 and Knight-ADRC: n=71) and all increases

1222 represent increase in abnormality. Significant differences in contiguous CSF stages are

1223 shown with asterisks. Horizontal line is drawn at z-score=1.96 which represents 95%Cl

1224 of the reference group (CU-). S shown with asterisks. Horizontal line is drawn at z-score=1.96 which represents 95%CI
1224 of the reference group (CU-). Smoothed LOESS lines of all AD biomarkers are shown
1225 in B for comparison. We excluded non-AD deme 1224 of the reference group (CU-). Smoothed LOESS lines of all AD biomarkers are shown

1225 in B for comparison. We excluded non-AD dementia patients to avoid bias in these

1226 analyses. CSF stage 0 represents being cla

in B for comparison. We excluded non-AD dementia patients to avoid bias in these
1226 analyses. CSF stage 0 represents being classified as normal by the model. *: p<0.05;
1227 **: p<0.01; ***: p<0.001.
1228 Abbreviations: analyses. CSF stage 0 represents being classified as normal by the model. *: p<0.05;

1227 **: p<0.01; ***: p<0.001.

1228 Abbreviations: AD, Alzheimer's disease; ADAS, Alzheimer's disease assessment

1229 scale; BNT, Bost 1227 **: p<0.01; ***: p<0.001.
1228 Abbreviations: AD, Alzl
1229 scale; BNT, Boston nan
1230 amyloid negative; CSF, i
1231 test; LOESS, locally e
1232 examination; mPACC, r Abbreviations: AD, Alzheimer's disease; ADAS, Alzheimer's disease assessment
1229 scale; BNT, Boston naming test; CI, confidence interval; CU-, cognitively unimpaired
1230 amyloid negative; CSF, cerebrospinal fluid; FCSRT, scale; BNT, Boston naming test; CI, confidence interval; CU-, cognitively unimpaired

1230 amyloid negative; CSF, cerebrospinal fluid; FCSRT, free and cued selective reminding

1231 test; LOESS, locally estimated scatterpl amyloid negative; CSF, cerebrospinal fluid; FCSRT, free and cued selective reminding
1231 test; LOESS, locally estimated scatterplot smoothing; MMSE, Mini-Mental state
1232 examination; mPACC, modified version of preclinic test; LOESS, locally estimated scatterplot smoothing; MMSE, Mini-Mental state
1232 examination; mPACC, modified version of preclinical Alzheimer's disease cognitive
1233 composite; TMT, trial making test; VOSP, visual obje examination; mPACC, modified version of preclinical Alzheimer's disease cognitive

1233 composite; TMT, trial making test; VOSP, visual object and space perception battery.

1235 **Ext Data Fig 6: Individual CSF stages for**

-
-

composite; TMT, trial making test; VOSP, visual object and space perception battery.
1234
1235 **Ext Data Fig 6: Individual CSF stages for predicting clinical progression**
1237 Kaplan-Meier curves for all individual CSF sta 1235
1236
1237
1238
1239 ----
1236
1237
1238
1239
1240 **Ext Data Fig 6: Individual CSF stages for predicting clinical progression**
1237 Kaplan-Meier curves for all individual CSF stages in BioFINDER-2 (A-C) ar
1238 ADRC (D-E) participants. For BioFINDER-2, progression from CU Kaplan-Meier curves for all individual CSF stages in BioFINDER-2 (A-C) and Knight-

1238 ADRC (D-E) participants. For BioFINDER-2, progression from CU or MCI at baseline to

1239 AD dementia is shown in A; progression from ADRC (D-E) participants. For BioFINDER-2, progression from CU or MCI at baseline to
1239 AD dementia is shown in A; progression from MCI at baseline to AD dementia is shown
1240 in B and; progression from CU at baseline t AD dementia is shown in A; progression from MCI at baseline to AD dementia is shown

1240 in B and; progression from CU at baseline to MCI is shown in C. For Knight-ADRC,

1241 progression from CDR=0 or CDR=0.5 at baselin 1240 in B and; progression from CU at baseline to MCI is shown in C. For Knight-ADRC, progression from CDR=0 or CDR=0.5 at baseline to CDR≥1 is shown in D and; progression from CDR=0 at baseline to CDR≥0.5 is shown in E.
 1241 progression from CDR=0 or CDR=0.5 at baseline to CDR≥1 is shown in D and;
1242 progression from CDR=0 at baseline to CDR≥0.5 is shown in E.
1243 Abbreviations: AD, Alzheimer's disease; CDR, clinical dementia rating; 1242 progression from CDR=0 at baseline to CDR≥0.5 is shown in E.
1243 Abbreviations: AD, Alzheimer's disease; CDR, clinical de
1244 cerebrospinal fluid; CU, cognitively unimpaired; MCI, mild cogniti
1245 **Ext Data Fig 7:**

Abbreviations: AD, Alzheimer's disease; CDR, clinical dementia rating; CSF,
1244 cerebrospinal fluid; CU, cognitively unimpaired; MCI, mild cognitive impairment.
1245 **Ext Data Fig 7: Individual biomarker levels by CSF sta**

cerebrospinal fluid; CU, cognitively unimpaired; MCI, mild cognitive impairment.
1245
1246 **Ext Data Fig 7: Individual biomarker levels by CSF stage in Knight-ADRC**
1248 **participants** Individual CSF biomarker levels, incl 1246
1247
1248
1250
1250
1251 1247
1248
1249
1250
1251 **Ext Data Fig 7: Individual biomarker levels by CSF stage in Knight-ADRC**
1248 **participants**
1249 Individual CSF biomarker levels, included in the model, by CSF stage particip
1250 shown in A including all Knight-ADRC par 1248 **participants**
1249 Individual CS
1250 shown in A
1251 biomarker lev
1252 All biomarker
1253 increases rep Individual CSF biomarker levels, included in the model, by CSF stage participants are
1250 shown in A including all Knight-ADRC participants. Depiction of individual AD-
1251 biomarker levels, not used in the creation of t shown in A including all Knight-ADRC participants. Depiction of individual AD-
1251 biomarker levels, not used in the creation of the model, per CSF stage are shown in B.
1252 All biomarker levels are z-scored based on a g biomarker levels, not used in the creation of the model, per CSF stage are shown in B.
1252 All biomarker levels are z-scored based on a group of CU- participants (n=71) and all
1253 increases represent increase in abnorma All biomarker levels are z-scored based on a group of CU- participants (n=71) and all
1253 increases represent increase in abnormality. Significant differences in contiguous CSF
1254 stages are shown with asterisks. Horizo increases represent increase in abnormality. Significant differences in contiguous CSF
1254 stages are shown with asterisks. Horizontal line is drawn at z-score=1.96 which
1255 represents 95%Cl of the reference group (CU-) 1254 stages are shown with asterisks. Horizontal line is drawn at z-score=1.96 which

1255 represents 95%Cl of the reference group (CU-). CSF stage 0 represent being classified

1256 as normal by the model. Black dots and represents 95%CI of the reference group (CU-). CSF stage 0 represent being classified
1256 as normal by the model. Black dots and vertical lines represent mean and SD per CSF
1257 stage. *: p<0.05; **: p<0.01; ***: p<0.001 1256 as normal by the model. Black dots and vertical lines represent mean and SD per CSF
1257 stage. *: p<0.05; **: p<0.01; ***: p<0.001.
35 1257 stage. *: p<0.05; **: p<0.01; ***: p<0.001.

Stage. *: p<0.05; **: p<0.01; ***: p<0.001.

Abbreviations: Aβ, amyloid-β; CI, confidence interval; CU-, cognitively unimpaired
1259 amyloid negative; CSF, cerebrospinal fluid; MMSE, Mini-Mental state examination;
1260 MTBR, microtubule binding region; NfL, neurofil amyloid negative; CSF, cerebrospinal fluid; MMSE, Mini-Mental state examination;
1260 MTBR, microtubule binding region; NfL, neurofilament light; PET, positron emission
1261 tomography; pT, phosphorylated tau; SuStaln, sub

MTBR, microtubule binding region; NfL, neurofilament light; PET, positron emission
1261 tomography; pT, phosphorylated tau; SuStaln, subtype and stage inference.
1262 **Ext Data Fig 8: Excluded CSF biomarkers by CSF stage**
 tomography; pT, phosphorylated tau; SuStaIn, subtype and stage inference.

1262 **Ext Data Fig 8: Excluded CSF biomarkers by CSF stage**

1264 Depiction of the CSF biomarkers excluded in the optimal model (pT231

1265 pT181/ 1262
1263
1264
1265
1266
1267
1268 **Ext Data Fig 8: Excluded CSF biomarkers by CSF stage**
1264 Depiction of the CSF biomarkers excluded in the optima
1265 pT181/T181) by CSF stage in BioFINDER-2 (A-B)
1266 participants. CSF pT217/T217 is also shown for comp 1264 Depiction of the CSF biomarkers excluded in the optimal model (pT231/T231 and
1265 pT181/T181) by CSF stage in BioFINDER-2 (A-B) and Knight-ADRC (C-D)
1266 participants. CSF pT217/T217 is also shown for comparison. C pT181/T181) by CSF stage in BioFINDER-2 (A-B) and Knight-ADRC (C-D)
1266 participants. CSF pT217/T217 is also shown for comparison. CSF levels are z-scored
1267 based on a group of CU- participants (BioFINDER-2: n=63, Knig participants. CSF pT217/T217 is also shown for comparison. CSF levels are z-scored
1267 based on a group of CU- participants (BioFINDER-2: n=63, Knight-ADRC: n=71) and
1268 all increases represent increase in abnormality. based on a group of CU- participants (BioFINDER-2: n=63, Knight-ADRC: n=71) and
1268 all increases represent increase in abnormality. Significant differences in contiguous
1269 CSF stages are shown with asterisks. Horizont all increases represent increase in abnormality. Significant differences in contiguous
1269 CSF stages are shown with asterisks. Horizontal line is drawn at z-score=1.96 which
1270 represents 95%Cl of the reference group (CSF stages are shown with asterisks. Horizontal line is drawn at z-score=1.96 which
1270 represents 95%Cl of the reference group (CU-). Smoothed LOESS lines of all CSF
1271 biomarkers are shown in B (BioFIDNER-2) and D (Kn 1270 represents 95%CI of the reference group (CU-). Smoothed LOESS lines of all CSF

1271 biomarkers are shown in B (BioFIDNER-2) and D (Knight-ADRC) for comparison. CSF

1272 stage 0 represent being classified as normal 1271 biomarkers are shown in B (BioFIDNER-2) and D (Knight-ADRC) for comparison. CSF

1272 stage 0 represent being classified as normal by the model. Black dots and vertical lines

1273 represent mean and SD per CSF stag

Stage 0 represent being classified as normal by the model. Black dots and vertical lines

1273 represent mean and SD per CSF stage, respectively. *: p<0.05; **: p<0.01; ***:

1274 p<0.001.

1275 Abbreviations: Aβ, amyloid 1273 represent mean and SD per CSF stage, respectively. *: p<0.05; **: p<0.01; ***:

1274 p<0.001.

Abbreviations: Aβ, amyloid-β; Cl, confidence interval; CU-, cognitively unimpaired

1275 amyloid negative; CSF, cerebrosp 1274 p<0.001.

1275 Abbreviat

1276 amyloid

1277 smoothing

1278 subtype a

1279 1275 Abbreviations: Aβ, amyloid-β; CI, confidence interval; CU-, cognitively unimpaired
1276 amyloid negative; CSF, cerebrospinal fluid; LOESS, locally estimated scatterplot
1277 smoothing; MTBR, microtubule binding regio 1276 amyloid negative; CSF, cerebrospinal fluid; LOESS, locally estimated scatterplot
1277 smoothing; MTBR, microtubule binding region; pT, phosphorylated tau; SuStaln,
1278 subtype and stage inference. 1277 smoothing; MTBR, microtubule binding region; pT, phosphorylated tau; SuStaIn,
1278 subtype and stage inference.
1279 1278 subtype and stage inference.
1279

A Progression from CU/MCI to AD dementia

C Progression from MCI to AD dementia

E Progression from CU to MCI

J Progression from CDR ≤ 0.5 to CDR ≥ 1

L Progression from CDR = 0 to CDR ≥ 0.5

M

A Initial model

B Optimization

C Final model

A

B Progression from MCI to AD dementia

A Progression from CU/MCI to AD dementia

+

 \mathfrak{g}

0

1 CSF stages:

4 5

2 3

 Ω 1

2

5 4

25

50

AD dementia survival probability (%)

AD dementia survival probability (%)

75

100

AD dementia survival probability (%) **AD dementia survival probability (%)** + ++++++++++++++ + + +++++++++++++++ + ⁺ + + ++ +++++++++++++++ + + +++ +++++ + ++++++ +++++++ + + ++ + +++ +++++++++ + + + + ++ + $+$ $+$ \Box^* +++ + + + + ++ + + + + $+ +$ + +++ +++++ +++ + + + + $+$ 0 1 2 3 4 5 4 74 74 40 33 12 0 1 25 23 12 6 3 0 46 46 34 21 6 0 3 32 31 19 12 3 0 43 40 30 20 5 0 23 20 15 6 1 0 0 1 2 3 4 5 **Years**

Knight-ADRC

D Progression from CDR ≤ 0.5 to CDR ≥ 1

E Progression from CDR = 0 to CDR ≥ 0.5

A

