

Table of contents

The MULNIAD study group at NCGG	2
The AIBL study group	3
Study background.....	4
Pilot study	5
Supplementary Discussion.....	6
Differences in discriminating power between PIB-PET and F-18 amyloid radiotracers.....	6
Possible explanations for the high performances of the plasma A β biomarkers.....	6
Validity of our IP-MS method for quantitative biomarker measurements.....	7
Possible causes of the site differences	9
Assessments of clinical utility of the plasma biomarker for clinical trials	9
Plasma biomarker performance in AD and other types of dementing disorders.....	10
Optimal generation of the composite biomarker	12
Additional experiment 1: Secretion of A β -related peptides from neuronal cells	14
Additional experiment 2: Determination of self-assembly tendency of APP ₆₆₉₋₇₁₁	16
References.....	18

The MULNIAD study group at NCGG

Contributors for the Multimodal NeuroImaging for Alzheimer's disease Diagnosis (MULNIAD) study

Hospital/Memory Clinic

Yutaka Arahata
Masahiko Bundo
Hidetoshi Endo
Koji Fukuda
Hideyuki Hattori
Kentaro Horibe

Yoko Konagaya
Takahiro Noguchi
Takashi Sakurai
Keisuke Suzuki
Akinori Takeda
Yukihiko Washimi

Imaging/Research

Junichiro Abe
Yoko Arai
Naohiko Fukaya
Ai Honda
Hiroshi Ikenuma
Yoshitaka Inui
Kengo Ito
Kaori Iwata
Takashi Kato
Kimiko Kato
Yasuhiro Kawasumi
Yumi Kimura
Chiharu Kitamura

Go Kizawa
Izumi Kuratsubo
Akinori Nakamura
Tomoko Nakazawa
Mayumi Oguri
Yuki Sakai
Takanori Sakata
Hitomi Shimizu
Yukari Suganuma
Michiyo Sugawara
Misako Yamagishi
Miyuki Yamauchi
Nozomi Yamawaki

The AIBL study group

AIBL Biomarker Stream

Scott Ayton
Samantha Burnham
Ashley Bush
Lesley Cheng
James Doecke (Convenor)
Christopher Fowler
Veer Gupta
Andy Hill
Eugene Hone

Simon Laws
Qiao-Xin Li
Florence Lim
Ralph Martins
Colin Masters
Steve Pedrini
Tenielle Porter
Stephanie Rainey-Smith
Blaine Roberts

AIBL Imaging Stream

Pierrick Bourgeat
Sveltana Bozinovski (nee Pejoska)
Patricia Desmond
Vincent Doré
Denise El-Sheikh
Jurgen Fripp
Shaun Frost

Gareth Jones
Neil Killeen
Christopher Rowe (Convenor)
Olivier Salvado
Victor Villemagne
Rob Williams

Study background

Alzheimer's disease (AD) is the most common cause of dementia. Its relentlessly progressive course, spanning a period of three decades, covers a continuum of preclinical, prodromal and clinical stages. Recent studies implementing biomarker information have strengthened and refined this concept^{1,2}. Although outcomes of potential disease-modifying trials for AD have been mixed, there is a general consensus that interventions in the earliest/mildest stages of the disease have the best chance of proving efficacy^{3,4}.

Identification of individuals with preclinical/prodromal AD for primary and secondary prevention trials, as envisioned by the Global Alzheimer Platform (GAP)⁵, the European Prevention of Alzheimer's Dementia (EPAD)⁶, Australian Imaging, Biomarker and Lifestyle Study of Aging (AIBL)⁷, and the ORANGE registry in Japan⁸ has become crucial. For the identification of preclinical/prodromal AD individuals, accurate AD-specific biomarker information is necessary. The accumulation of amyloid- β ($A\beta$) is the earliest pathognomonic signature of AD, however, currently available validated methods to assess $A\beta$ deposition ($A\beta$ -PET imaging or CSF- $A\beta$ levels) are invasive and/or costly^{9,10}. Therefore, blood-based biomarkers, which are simple, minimally invasive and cost-effective, are needed^{11,12}. Such biomarkers, especially for plasma $A\beta$ levels, have been extensively investigated, but most have failed to demonstrate clinical utility^{11,12}, some even showing changes in opposite directions¹³. There have been only a few studies that assessed biomarker performances of plasma $A\beta$ at an individual level; for example, Fei et al. (2011)¹⁴ demonstrated that the ratio of plasma $A\beta_{1-42}/A\beta_{1-40}$ could identify incipient AD in individuals with mild cognitive impairment (MCI) with sensitivity 85.7% and specificity 69.7%. However, other large-scale cohort studies could not replicate such high performance. The AIBL group reported¹⁵ that the plasma $A\beta_{1-42}/A\beta_{1-40}$ ratio was slightly reduced in AD by group comparisons among AD, MCI and cognitively normal (CN) subjects, and it was negatively correlated with $A\beta$ -PET burden. In ADNI (the Alzheimer's Disease Neuroimaging Initiative, <http://adni.loni.usc.edu/>) study¹⁶, the ratio of plasma $A\beta_{1-40}/A\beta_{1-42}$ (the ratio is inverted to that in the AIBL study) showed a weak positive correlation with $A\beta$ -PET burden in only *APOE4* negative individuals, and group differences between clinical stages, or between $A\beta$ -positive and $A\beta$ -negative subject were not found. Furthermore, the most recent large-scale study failed to demonstrate any differences between preclinical AD subjects, who later developed AD, and dementia-free individuals, using plasma $A\beta_{1-40}$, $A\beta_{1-42}$, and $A\beta_{1-42}/A\beta_{1-40}$, and therefore concluded there was no clinical utility in these plasma $A\beta$ biomarkers¹⁷. Thus, there has not yet been any study which successfully validates the clinical utility of blood-based $A\beta$ markers.

Pilot study

At the starting point of this international collaboration, a blinded pilot study was conducted to judge whether the project was worthy to proceed or not, by testing the plasma biomarker performance in a small-scale sample provided by AIBL. The AIBL group selected 20 $A\beta^+$ subjects (9 AD and 11 CN, age 72.2 ± 3.7 yrs., 14 females, 13 *APOE4* carriers) and 20 $A\beta^-$ subjects (20 CN, age 72.1 ± 3.6 yrs., 12 females, 8 *APOE4* carriers) at their choice, and sent 40 plasma samples to Koichi Tanaka Mass Spectrometry Research Laboratory (Shimadzu) without any clinical or imaging information. After the IP-MS measurements, the composite biomarker values were generated before getting any sample information from AIBL. For each $APP_{669-711}/A\beta_{1-42}$ and $A\beta_{1-40}/A\beta_{1-42}$, the mean and SD values of the pilot 40 samples were computed and used for the normalization to z-scores. Then, individual z-scores of $APP_{669-711}/A\beta_{1-42}$ and $A\beta_{1-40}/A\beta_{1-42}$ were averaged and used as a value for the composite biomarker. The receiver operating characteristic (ROC) analyses demonstrated that the area under the curves (AUCs) were high in $APP_{669-711}/A\beta_{1-42}$ (0.923) and $A\beta_{1-40}/A\beta_{1-42}$ (0.930), and the highest in the composite biomarker (0.975) (Extended Data Fig. 1b). Compared with both $APP_{669-711}/A\beta_{1-42}$ and $A\beta_{1-40}/A\beta_{1-42}$, the composite biomarker showed significant improvements in the net reclassification improvement (NRI) ($P < 0.001$ and $P = 0.007$, respectively) and in the integrated discrimination improvement (IDI) ($P < 0.001$ and $P = 0.002$, respectively).

Afterwards, the weights of $APP_{669-711}/A\beta_{1-42}$ and $A\beta_{1-40}/A\beta_{1-42}$ were estimated using the binominal logistic regression analysis inputting the z-scores of these two biomarkers as variables. Results demonstrated that both normalized values of $APP_{669-711}/A\beta_{1-42}$ and $A\beta_{1-40}/A\beta_{1-42}$ significantly contributed to the model with coefficients of 2.60 ($P = 0.037$) and 2.57 ($P = 0.017$), respectively. This indicated the pre-determined weight of 1:1 was appropriate to use for the main analyses.

Supplementary Discussion

Differences in discriminating power between PIB-PET and F-18 amyloid radiotracers

Our results showed that the performance of the plasma biomarkers varied between the PET tracers used for determining brain A β status. Compared to PIB, the performance estimates for FLUTE and FBP yielded about 10% lower accuracy. Given that there were no significant differences in the two independent PIB datasets, it is unlikely that this difference is attributable to variability in the biomarker performance, but rather the consequence of higher variance and lower performance of the ^{18}F -tracers compared to PIB¹⁸⁻²¹. When Mormino and colleagues (2014 + Suppl Materials)²¹ compared the discriminating power of FBP and PIB, they found that the discriminating power of FBP was much lower than that of PIB. For example, the “ambiguity area” (the area between high and low A β burden as determined by each tracer) is 10 times larger for FBP than for PIB. In other words, F-18 radiotracers are less sensitive than PIB to discriminate between high and low A β burden at the intermediate or peri-threshold levels. Thus, these differences in discriminating power likely explain the lower biomarker performances when compared to ^{18}F -tracers observed in our study. It is also likely that the higher variance and lower performance of the ^{18}F -tracers affected the results of the correlation analyses, where the correlation coefficients were relatively lower in the ^{18}F -tracers compared with PIB (Fig. 3a). This finding is consistent with our previous report²² that the correlations between CSF-A β_{1-42} levels and FLUTE, and FBP SUVR values (Spearman's ρ -0.51 and -0.38, respectively) were lower than that of PIB (Spearman's ρ -0.76).

While this difference was already known, it was important not to restrict the analysis to PIB, but to also include ^{18}F A β tracers because these are the ones approved for clinical use and the ones being used in therapeutic trials.

Possible explanations for the high performances of the plasma A β biomarkers

The high performance of the plasma A β assays may be attributable to the following factors: 1) Analytic factors: It has been difficult to measure plasma A β_{1-42} and A β_{1-40} levels accurately using the conventional ELISA-based techniques which employ only antibody affinity, because compared with CSF, the concentration of total protein in plasma is about 100-fold higher²³, and that of A β about 50-fold lower²⁴. Our IP-MS method can accurately distinguish the mass of each A β peptide fragment in addition to the enrichment by IP²⁵. There have been two other studies that also applied IP-MS technique for plasma A β assays. Although the study by Pannee et al.²⁶ was not robust enough to discriminate AD from control individuals, a very recent study by Ovod et al.²⁷ has demonstrated that the plasma ratio of A β_{42} /A β_{40} using liquid chromatography tandem mass spectrometry technique could predict brain A β burden with AUC of 0.8865 ($n = 41$). The reliability of our IP-MS method for the quantitative biomarker measurements, and detailed

methodological differences with other IP-MS studies are discussed in next section. 2) Post-analytical factors: In our IP-MS system, plasma A β ₁₋₄₂ level by itself showed strong biomarker performance, however, the ratios of plasma A β ₁₋₄₂ with reference peptides, APP₆₆₉₋₇₁₁/A β ₁₋₄₂ and A β ₁₋₄₀/A β ₁₋₄₂ significantly improved the performance. In addition, the composite score further enhanced the performances in the AIBL dataset and also increased correlations with A β -PET SUVR values, performing as the most stable biomarker. Many pathophysiological conditions including inflammation and renal function¹⁵, hypertension, diabetes and ischemic heart disease²⁸, as well as circadian fluctuations²⁴, influence A β plasma levels, and are thought to be responsible for the inter-individual variances of plasma A β levels, which are considered to be much larger than those observed in CSF. In addition, it is known that some individuals might express anti-A β auto-antibodies^{3,29} or anti-mouse antibody (HAMA) which could bind 6E10³⁰, and thereby mask or block 6E10 binding in IP. However, since these reference peptides have similar amino acid sequence and molecular size as A β ₁₋₄₂, these individual systemic factors are largely canceled out by calculating ratios, therefore minimizing inter-individual variances.

Validity of our IP-MS method for quantitative biomarker measurements

The reliability of our IP-MS methods for quantitative biomarker measurements were carefully validated by several steps. First, we plotted standard curves of synthetic peptides, A β ₁₋₄₂, APP₆₆₉₋₇₁₁ and A β ₁₋₄₀, in PBS containing 3 mg/mL bovine serum albumin, by measuring the normalized signal intensity with stable-isotope-labeled (SIL)-A β ₁₋₃₈ using the IP-MS method, that is, treated by IP and then measured in MALDI-TOF MS as plasma sample. The results showed very solid linearity with the coefficient of determination (R^2) 0.996 to 0.998 (Extended Data Fig. 5a). Second, we measured normalized intensity of endogenous A β ₁₋₄₂, A β ₁₋₄₀, and APP₆₆₉₋₇₁₁, which are present in human plasma (Tennessee Blood Services), and tested the linear relationship with peptide concentration by diluting the plasma with PBS. The results also demonstrated very solid linearity for all peptides with the coefficient of determination (R^2) 0.986 to 0.996 (Extended Data Fig. 5b). This linearity appears as good as that reported using ELISA³¹. Third, we analyzed the dose-dependency of the normalized intensity for each synthetic peptide of A β ₁₋₄₂, A β ₁₋₄₀, and APP₆₆₉₋₇₁₁ which were spiked into the human plasma. The results also showed very good linearity for each peptide with 0.999 to 1.000 of R^2 (Extended Data Fig. 5c), which values are as high or even higher than those reported by Pannee et al³² measured in CSF. Note that the regression lines have intercepts in the y-axis because of the existence of each endogenous peptide in the plasma. On the other hand, the regression lines reported by Pannee et al³² do not show the intercepts. This was because they used stable isotope labeled peptides, which are not natural forms, so they can be measured separately from the endogenous peptides in the CSF on the basis of their mass. These results indicate that the quantitative nature of our IP-MS method is very reliable in a human

plasma assay within specific ranges shown in Extended Data Fig. 5c. In addition, if we convert the normalized signal intensity to absolute peptide concentration using the standard curves shown in Extended Data Fig. 5a, then the averaged concentration of A β ₁₋₄₂ of the A β ⁻ and A β ⁺ groups in NCGG+AIBL overall dataset would be 38.5 $\pm$ 5.7 (14.8%) [mean $\pm$ SD (% coefficient variance)] ng/L and 28.3 $\pm$ 2.5 (8.8%) ng/L, respectively. These values are similar with those reported by Ovod et al.²⁷ using IP-MS with liquid chromatography tandem mass spectrometry (LC-MS/MS). The averaged concentration of A β ₁₋₄₀ of the A β ⁻ and A β ⁺ groups in the NCGG+AIBL overall dataset are estimated as 233.5 $\pm$ 43.1 (18.5%) ng/L and 220.7 $\pm$ 36.8 (16.7%) ng/L, respectively. These values lie between the reported values of other IP-MS based studies by Pannee et al.²⁶ and Ovod et al.²⁷.

There are considerable differences between our IP-MS methodology and those reported by Pannee et al.^{26,32}, and Ovod et al.²⁷. First, we used MALDI-TOF for mass spectrometry whereas other groups used LC-MS/MS technology. In general, LC-MS/MS is considered to be more sensitive in protein quantification, however, MALDI-TOF MS can also be used for reliable protein quantification^{25,33-36}. We also used SIL-A β ₁₋₃₈ as a common internal standard for all A β -related peptides, whereas other groups used individual internal standards corresponding to each peptide (SIL-A β ₁₋₄₂ for A β ₁₋₄₂, and SIL-A β ₁₋₄₀ for A β ₁₋₄₀). Since each peptide has different self-assembly tendency and some degree of peptide loss may occur during the measurements due to self-aggregation, especially for A β ₁₋₄₂, one could argue that our method partially sacrifices the reliability of quantitative peptide measurements. However, this possibility is very unlikely, because aggregation of A β ₁₋₄₂ was not observed within a period of 48h at a concentration of less than 300 mg/L (pH2.5, room temperature)³⁷, which is 6,000,000 and 1,500,000 times higher than the reported concentration of plasma A β ₁₋₄₂ (~50 ng/L)¹⁷ and the target range of the IP-MS measurement for A β ₁₋₄₂ (~40 pM = ~180ng/L), respectively. In addition, the aggregation of A β ₁₋₄₂ is also known to be pH dependent. At pH 7.4, the self-assembly tendency becomes much weaker than the above-mentioned acidic environment (pH2.5); the concentration to allow the aggregation of A β ₁₋₄₂ is reported to be about 750 mg/L³⁷. Since the IP process of our method is performed at pH7.4 (on ice, 1h x 2 times), it is very improbable that aggregation of A β ₁₋₄₂ and/or A β ₁₋₄₀ affects the reliable quantification of our IP-MS method. In addition, the very solid linearity shown in Extended Data Fig. 5c clearly indicated that our IP-MS method is very reliable in peptide quantification within the range less than 40pM for A β ₁₋₄₂ and APP₆₆₉₋₇₁₁, and less than 160pM for A β ₁₋₄₀. Furthermore, we conducted additional analyses to compare our method with the method using individual internal standards. We employed commercially-available SIL-A β ₁₋₄₂ and SIL-A β ₁₋₄₀ (rPeptide, Bogart, GA) and measured biomarkers for 19 subjects' plasma samples derived from NCGG. These subjects, including 1 AD, 7 MCI, and 11 CN cases, were a part of the NCGG dataset in the main analyses, but the time point for the plasma sampling was different. According

to PIB-PET imaging, 9 and 10 of them were classified as $A\beta^+$ and $A\beta^-$, respectively. The biomarker measurements were performed in totally blinded manner as in the main analyses. Extended Data Fig. 5d shows correlations of the $A\beta_{1-40}$, $A\beta_{1-42}$, and $A\beta_{1-40}/A\beta_{1-42}$ values between the two methods. All of them showed statistically significant strong correlations with the coefficients $r > 0.8$. Moreover, ROC analyses demonstrated that both methods showed high performances in predicting the individual $A\beta$ status ($A\beta^+$ or $A\beta^-$) (Extended Data Fig. 5e and f). These results indicate that both methods can generate reliable biomarkers. We cannot state which method is superior, but our method has one advantage when computing peptide ratios, which are our target biomarker. For example, the biomarker value for $A\beta_{1-40}/A\beta_{1-42}$ is computed from the following formula with the signal intensity.

$$(A\beta_{1-40}/SIL-A\beta_{1-38}) / (A\beta_{1-42}/SIL-A\beta_{1-38}) = A\beta_{1-40}/A\beta_{1-42}$$

As shown in the formula, the factor $SIL-A\beta_{1-38}$ is completely erased. This indicates that using the common internal standard can cancel out any implicit errors for added $SIL-A\beta_{1-38}$ amounts caused by production, preparation, and/or handling, when computing the peptide ratio. On the other hand, the method using individual internal standards cannot cancel out these types of errors. Therefore, our IP-MS method allows for a more reliable and robust generation of peptide ratios, which can be computed for each sample without requiring any standard regression formula to convert a signal intensity to an absolute plasma concentration value. This was also why we used normalized signal intensity values, but not peptide concentration values, for the biomarker analyses.

Possible causes of the site differences

The between-site differences are largely influenced by both pre-analytic and analytic factors. In particular, pre-analytic factors are known to be the largest source of variability in laboratory testing³⁸. In our study, there were substantial differences in the procedures and processes from blood sampling to plasma storage applied to the samples between NCGG and AIBL (see Methods, Blood processing and plasma storage). However, the influence of the analytic factor on our assay appears to be similar or even smaller compared with the CSF assay, because the coefficients of variance (CVs) of our intra- and inter-day assays yielded 1.6% to 10.7% (see Methods, Plasma $A\beta$ measurements), whereas these were 5% to 19% in within-laboratory CSF assays³⁹. Therefore, we consider the main causes of the between-site differences of our study are mainly due to the pre-analytic factors.

Assessments of clinical utility of the plasma biomarker for clinical trials

Using the validation AIBL PIB data, we assessed clinical utility of the composite biomarker

assuming two specific settings, screening for prodromal AD or preclinical AD for clinical trials.

In the setting of prodromal AD screens, the prevalence of $A\beta^+$ MCI (prodromal AD) among MCI individuals is about 66%⁴⁰ in the age range similar to this study (around 75 years). With this prevalence, the Diagnostic performance plots (DP-plots) of the composite biomarker within MCI individuals in the AIBL PIB data was drawn including PPV and NPV (Extended Data Fig. 8a). In this case, the representative optimal cut-off points, which were determined by the Youden's index (Y-point) in this dataset, appeared to represent the optimal cut-off point indicating, sensitivity = 0.900, specificity = 0.923, PPV = 0.958, and NPV = 0.826. With these values, the following estimate might be reasonable: if a clinical trial requires 300 prodromal AD individuals, $A\beta$ -PET scans are needed for about 455 MCI individuals without the plasma biomarker. But if about 506 MCI individuals are screened by the plasma biomarker, $A\beta$ -PET scans would be needed only for about 313 plasma biomarker positive individuals. The plasma biomarker may therefore be able to reduce unnecessary PET scans from 155 to 13.

In a preclinical AD screen scenario, the prevalence of $A\beta^+$ individuals among general elderly people aged around 75 years is about 30%⁴⁰. The DP-plots of the composite biomarker within CN in the AIBL PIB data at prevalence = 30% are shown in Extended Data Fig. 8b. In this case, performances of the composite biomarker at the Y-point were; sensitivity = 0.880, specificity = 0.868, PPV = 0.741, and NPV = 0.944. With these values, the following estimate can be made: if a clinical trial requires 300 individuals at risk for AD, $A\beta$ -PET scans are needed for 1000 general elderly people. But if about 1137 elderly individuals are screened by the plasma biomarker, $A\beta$ -PET scans would be needed for about 405 plasma biomarker positive individuals. In this setting, the plasma biomarker may be able to reduce unnecessary PET scans from 700 to 105. Other estimations can be derived by shifting a cut-off to higher specificity/PPV point where the composite biomarker performs with sensitivity = 0.520, specificity = 0.974, PPV = 0.894, and NPV = 0.826 (red dashed vertical line in Extended Data Fig. 8b, left). At this cut-off value, a blood-based screen for 1924 cognitively normal people would detect about 336 plasma biomarker-positive individuals who need $A\beta$ -PET scans to detect 300 $A\beta^+$ individuals at risk for AD. Although it depends on the cost for the plasma biomarker measurements, this estimate might provide the best cost-performance outcome.

Plasma biomarker performance in AD and other types of dementing disorders

In order to test the clinical utility of the plasma biomarkers in the differential diagnosis of AD, we conducted an additional study with a new dataset. Since the current study's cohorts (NCGG and AIBL) are focusing on the AD continuum, there are only a small number of non-AD cases, therefore, we needed to add another cohort as a source for this analysis. The additional cohort is a multicenter study aiming at validating the clinical utility of the plasma biomarker that extended

the target to all types of dementing disorders. The study just started in 2016 involving three Japanese institutes, NCGG, The Tokyo Metropolitan Institute of Gerontology (TMIG), and Kindai University Faculty of Medicine (Kindai). The study protocol is similar to NCGG's cohort described in the main text and all participants underwent PIB-PET imaging. The study was approved by the appropriate institutional ethics committee, and written informed consent was obtained from all participants (or the respective legal guardian) prior to participation.

There were 51 new subjects ($n = 9$, 23 and 19 from NCGG, TMIG, and Kindai, respectively) who were clinically classified, without biomarker information, as 31 AD (age 75.6 ± 6.0 , 23 females) and 20 non-AD (age 74.4 ± 5.6 , 8 females) cases (Extended Data Figure 8c and d). Plasma samples and PIB-PET data of these cases were analyzed in a blinded and independent manner as described in the main text. According to PIB-PET imaging (CAPAIBL, cut-off SUVR = 1.4), AD cases were classified as $A\beta^+$ ($n = 22$) and $A\beta^-$ ($n = 9$), respectively. With the $A\beta$ -PET information, these 22 $A\beta^+$ AD cases are considered as high-likelihood AD, whereas the 9 $A\beta^-$ cases are deemed as dementia unlikely due to AD. In the non-AD group, there were 8 $A\beta^+$ and 12 $A\beta^-$ cases. These 8 $A\beta^+$ cases are therefore deemed as mixed dementias or atypical AD, whereas the $A\beta$ negativity supports 12 cases as other types of dementing disorders. Note that the primary endpoint of this additional analysis is the same as the main study, the performance of plasma $A\beta$ biomarkers in determining the individual status of $A\beta$ deposition ($A\beta^+$ or $A\beta^-$) using PIB-PET as the standard of truth, and not to distinguish between AD and non-AD individuals. Thus, there were 30 $A\beta^+$ and 21 $A\beta^-$ cases and these numbers satisfied the sample size as described in the Methods, Sample size considerations section.

The ROC analysis demonstrated that the composite biomarker showed the highest performance (Extended Data Figure 8e). Using the same common cut-off value (0.376), as determined in the Discussion section (and Extended Data Fig. 7) of the original manuscript, the composite biomarker showed 96.7% sensitivity, 81.0% specificity, and 90.2% accuracy in overall data to predict the individual $A\beta$ status ($A\beta^+$ or $A\beta^-$) (Extended Data Figure 8f and g). The performances were also high for both within AD or within non-AD group analyses. These results suggest that the plasma biomarker may be useful in distinguishing individuals with or without $A\beta$ burden, and while independent of clinical classification, helpful information for the accurate differential diagnosis as well as determining therapeutic strategies.

We acknowledge the results still need to be further validated with larger sample size. This study analyzed the mixture of several types of dementing disorders with a small number of cases together, and it could be a similar to routine clinical practice. However, future studies focusing on larger groups of specific types of dementia are also needed.

Optimal generation of the composite biomarker

In this study, we pre-determined the weight of the $\text{APP}_{669-711}/\text{A}\beta_{1-42}$ and $\text{A}\beta_{1-40}/\text{A}\beta_{1-42}$ normalized values (z-scores) as 1:1, however, there could have been better proportions calculated to generate the composite biomarker. To test this possibility, we estimated the weight optimal to each NCGG PIB and AIBL PIB dataset using the binominal logistic regression analysis inputting z-scores of these two biomarker data sets as variables. Results demonstrated that in the NCGG PIB data, the coefficients for $\text{APP}_{669-711}/\text{A}\beta_{1-42}$ and $\text{A}\beta_{1-40}/\text{A}\beta_{1-42}$ were 1.14 ($P = 0.052$) and 3.59 ($P < 0.0001$), whereas in the AIBL PIB data, they were 3.04 ($P < 0.0001$) and 1.95 ($P < 0.0001$), respectively. We then tested the impact of these weightings on the biomarker performance. We generated two different composite biomarkers using the weight estimated by the NCGG data (1.14:3.59) and by the AIBL data (3.04:1.95), and compared the performance with the original composite biomarker (1:1 weight) using the ROC analyses. The results demonstrated that in the NCGG discovery data, the impact of different weightings was very small, i.e. the AUCs were almost identical (0.962 to 0.969) irrespective of the weights (Extended Data Fig. 9a left, and 9c left). In the validation AIBL PIB data, the weight estimated by the discovery NCGG data (1.14:3.59) slightly diminished the performances from the original composite biomarker (AUC 0.941 to 0.919, and accuracy 0.883 to 0.847). However, the weight (3.04:1.95) estimated by its own dataset (AIBL PIB data) did not improve the performance (AUC 0.943 and accuracy 0.883) (Extended Data Fig. 9a right, and 9c right). These results suggest that in this study, the original weight 1:1 was optimal or close to optimal.

In generating the composite biomarker, the discovery NCGG dataset was used for a standard database that all individual z-scores, including the AIBL dataset, of the $\text{APP}_{669-711}/\text{A}\beta_{1-42}$ and $\text{A}\beta_{1-40}/\text{A}\beta_{1-42}$ were computed using the mean and SD of the NCGG data. This was because we considered that referring to a shared database created by the discovery data should be better for strict validation. However, since there were considerable site differences between NCGG and AIBL in the plasma A β levels, referring a database created by its own site may improve the performance of the composite biomarker. To test this, we generated alternative composite biomarker by using mean and SD of the AIBL PIB data, and compared with the original composite biomarker using the ROC analyses. Results demonstrated that the performance was almost identical between the two normalization methods (Extended Data Fig. 9b). These results point to the potential general versatility of the composite biomarker.

However, optimal generation of the composite biomarker should be further validated in other large-scale datasets, preferably under the controlled standardized operating procedures (SOP). In those future studies, adding information of other A β peptides such as $\text{A}\beta_{1-38}$ and $\text{A}\beta_{1-39}$, which are also measurable by IP-MS²⁵ but were not analyzed in this study, might also improve biomarker performance. Otherwise, the performance of the $\text{A}\beta_{1-40}/\text{A}\beta_{1-42}$ ratio alone might prove to be

sufficient.

Additional experiment 1: Secretion of A β -related peptides from neuronal cells

Methods

Cell Culture

BE(2)-C cell (CRL-2268) was purchased from American Type Culture Collection. Cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) with high glucose (WAKO) supplemented with 10% heat-inactivated fetal bovine serum (FBS; Hyclone), 50 units/mL penicillin (Invitrogen), and 50 mg/mL streptomycin (Invitrogen) at 37 °C under humidified air containing 5% CO₂. We routinely check mycoplasma contamination by DAPI staining and PCR analysis. For collection of conditioned medium, BE(2)-C cells were seeded at 3×10^7 cells/10 cm dish. After 120 hours of incubation, the conditioned medium was centrifuged at 1,200 rpm for 3 min. The culture supernatants were obtained and stored at -80 °C until use.

Culture supernatant A β -related peptide measurement

Measurement of A β -related peptides in culture supernatants were performed by IP-MS as described in the Methods, "Plasma A β measurements" section except for using 20 pM SIL-A β 1-38 peptide as an internal standard. MS/MS analysis for identification of A β -related peptides was obtained using a MALDI-QIT reflectron TOF mass spectrometer (AXIMA Resonance, Shimadzu/KRATOS) in the positive ion mode and MS/MS fragments were generated by collision-induced dissociation with argon gas. 2,5-dihydroxybenzoic acid (DHB) was used instead of CHCA as the MALDI matrix. Peak lists were created from the raw MS/MS spectra by Mascot Distiller (Matrix Science), followed by analysis using Mascot software Version 2.4 (Matrix Science). The significance threshold for peptide identification was set at $P < 0.05$.

Results

A β -related peptides produced from human neuroblastoma cell line BE(2)-C were analyzed by IP-MS (Extended Data Fig. 6a). Mass spectrum of the culture supernatant from BE(2)-C cells was very similar to that of human plasma from A β ⁺ subjects. APP₆₆₉₋₇₁₁ as well as A β ₁₋₄₀ and A β ₁₋₄₂ were detected in the culture supernatant from BE(2)-C cell, whereas no A β -related peptide except for an internal standard was detected in medium without cell culture. These A β -related peptides were identified using MS/MS analysis. This result indicated that APP₆₆₉₋₇₁₁ is secreted from BE(2)-C cell to the culture supernatant.

Discussion

APP₆₆₉₋₇₁₁ has been found in human plasma^{25,41}, but it was unclear whether APP₆₆₉₋₇₁₁ is produced from neuronal cells. We demonstrated the secretion of APP₆₆₉₋₇₁₁ in BE(2)-C cells. Furthermore,

Beyer et al. also reported APP₆₆₉₋₇₁₁ is produced by SH-SY5Y neuroblastoma cells transfected with wild-type APP695⁴². These results confirmed that the APP₆₆₉₋₇₁₁ can be generated from neuronal cells.

Additional experiment 2: Determination of self-assembly tendency of APP₆₆₉₋₇₁₁

Determination of self-assembly tendency of APP₆₆₉₋₇₁₁

Methods

Peptides

A β ₁₋₄₂ and APP₆₆₉₋₇₁₁ were purchased from Peptide Institute (Ibaragi, Japan). The peptides were dissolved in 0.02% ammonia on ice, and any large aggregates that may have acted as a seed for aggregation were removed by ultracentrifugation in 500 μ L polyallomer tubes at 540,000 $\times$ g and 4 °C for 3 h. The peptide concentrations of the supernatants were determined in triplicate by the Micro BCA protein assay (Pierce, Rockford, IL). The supernatants were collected and stored at –80 °C prior to being used. Just before the experiment, the stock solution was mixed with an equal volume of double concentrated PBS [16.0 g/L NaCl, 0.40 g/L KCl, 2.3 g/L Na₂HPO₄, and 0.4 g/L KH₂PO₄ (pH 7.4)].

Thioflavin T (Th-T) Assay

The peptides (15 μ M) were incubated in PBS at 37 °C. Peptide assembly into fibrils was monitored by the Th-T assay^{43,44}. The sample (final A β concentration of 0.5 μ M) was added to a 5 μ M Th-T solution in 50 mM glycine buffer (pH 8.5). Fluorescence at 490 nm was measured at an excitation wavelength of 446 nm at 25 °C. The blank fluorescence (buffer) was subtracted.

Size Exclusion Chromatography

Each sample was centrifuged at 10,000 $\times$ g for 10 min and the supernatant (25 μ L) was injected into a Sephacryl S-300 HR column with an exclusion limit of 150,000Da (GE Healthcare Bioscience, NJ), which had been pretreated with an excess of BSA to block the non-specific binding of peptides. The column was eluted with PBS, pH7.4, at a flow rate of 0.5 mL/min and the peptides were detected by measuring UV absorbance at 220 nm. Molecular mass was estimated with FITC-dextran (MW 4,400, 21,200 and 42,000) as standards.

Circular Dichroism (CD)

CD spectra of the peptides (15 μ M) were measured on a Jasco J-820 apparatus at 37 °C, using a 1-mm path length quartz cell to minimize the absorbance due to buffer components. Eight scans were averaged for each sample. The blank spectra (buffer) were subtracted.

Results

Self-assembly tendency of the peptides was monitored by the Th-T assay (Extended Data Fig. 6b), size exclusion chromatography (Extended Data Fig. 6c), and CD spectroscopy (Extended Data Fig. 6d). APP₆₆₉₋₇₁₁ was much less prone to self-aggregate than A β ₁₋₄₂. In the case of A β ₁₋₄₂, the Th-T fluorescence started to increase after a lag time of 3 h and reached plateau at 6 h. No monomer was detected in the supernatant after 6 h, suggesting the complete formation of amyloid

fibrils. CD spectra were characteristic of a random coil structure until 3 h. After 6 h, the presence of a minimum at 220 nm indicated the formation of β -sheets. In contrast, in the case of APP₆₆₉₋₇₁₁, the Th-T fluorescence gradually increased after 6 h, and the monomer fraction was concomitantly decreased. The formation of β -sheets was only partial, and a significant fraction of monomer was still detectable even at 24 h.

Discussion

It remains to be elucidated why the APP₆₆₉₋₇₁₁/A β ₁₋₄₂ and A β ₁₋₄₀/A β ₁₋₄₂ ratios in the plasma highly correlate with the presence of amyloid in the brain. Given that APP₆₆₉₋₇₁₁ has almost the same size as A β ₁₋₄₂ but it lacks C-terminal 2 hydrophobic amino acids of A β ₁₋₄₂ just like A β ₁₋₄₀⁴¹, it is possible to assume that its self-assembly tendency is extremely lower than that of A β ₁₋₄₂ as previously reported with A β ₁₋₄₀⁴⁵. Indeed, this possibility has been confirmed in this study (Extended Data Fig. 6b, c, d). Thus, APP₆₆₉₋₇₁₁ as well as A β ₁₋₄₀ likely functions as a good reference to monitor the assembly and deposition of A β ₁₋₄₂ in the brain.

References

1. Jack, C.R., Jr., *et al.* Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* **7**, 257-262 (2011).
2. Villemagne, V.L., *et al.* Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *Lancet Neurol* **12**, 357-367 (2013).
3. Sevigny, J., *et al.* The antibody aducanumab reduces A β plaques in Alzheimer's disease. *Nature* **537**, 50-56 (2016).
4. Sperling, R., Mormino, E. & Johnson, K. The evolution of preclinical Alzheimer's disease: implications for prevention trials. *Neuron* **84**, 608-622 (2014).
5. Cummings, J., *et al.* Re-engineering Alzheimer clinical trials: Global Alzheimer's platform network. *J Prev Alz Dis* **3**, 114-120 (2016).
6. Ritchie, C.W., *et al.* Development of interventions for the secondary prevention of Alzheimer's dementia: the European Prevention of Alzheimer's Dementia (EPAD) project. *Lancet Psychiat* **3**, 179-186 (2016).
7. Ellis, K.A., *et al.* The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. *Int Psychogeriatr* **21**, 672-687 (2009).
8. Saji, N., Sakurai, T., Suzuki, K., Mizusawa, H. & Toba, K. ORANGE's challenge: developing wide-ranging dementia research in Japan. *Lancet Neurol* **15**, 661-662 (2016).
9. Cohen, J.P., Dong, J., Lu, C.Y. & Chakravarthy, R. Restricting access to florbetapir: Medicare coverage criteria for diagnostics and drugs are inconsistent. *BMJ* **351**, h3333 (2015).
10. Monserrate, A.E., *et al.* Factors associated with the onset and persistence of post-lumbar puncture headache. *JAMA Neurol* **72**, 325-332 (2015).
11. Henriksen, K., *et al.* The future of blood-based biomarkers for Alzheimer's disease. *Alzheimers Dement* **10**, 115-131 (2014).
12. O'Bryant, S.E., *et al.* Blood-based biomarkers in Alzheimer disease: Current state of the science and a novel collaborative paradigm for advancing from discovery to clinic. *Alzheimers Dement* (2016).
13. Toledo, J.B., Shaw, L.M. & Trojanowski, J.Q. Plasma amyloid beta measurements - a desired but elusive Alzheimer's disease biomarker. *Alzheimer's Research & Therapy* **5**, 8 (2013).
14. Fei, M., Jianghua, W., Rujuan, M., Wei, Z. & Qian, W. The relationship of plasma A β levels to dementia in aging individuals with mild cognitive impairment. *J Neurol Sci* **305**, 92-96 (2011).
15. Rembach, A., *et al.* Changes in plasma amyloid β in a longitudinal study of aging and

- Alzheimer's disease. *Alzheimers Dement* **10**, 53-61 (2014).
16. Swaminathan, S., *et al.* Association of plasma and cortical amyloid beta is modulated by APOE ϵ 4 status. *Alzheimers Dement* **10**, e9-e18 (2014).
 17. L  vheim, H., *et al.* Plasma concentrations of free amyloid- β cannot predict the development of Alzheimer's disease. *Alzheimers Dement*, doi:10.1016/j.jalz.2016.12.004 (2017).
 18. Vandenberghe, R., *et al.* ^{18}F -flutemetamol amyloid imaging in Alzheimer disease and mild cognitive impairment: a phase 2 trial. *Ann Neurol* **68**, 319-329 (2010).
 19. Wong, D.F., *et al.* *In vivo* imaging of amyloid deposition in Alzheimer disease using the radioligand [^{18}F]AV-45 (florbetapir F 18). *J Nuclear Medicine* **51**, 913-920 (2010).
 20. Landau, S.M., *et al.* Amyloid PET imaging in Alzheimer's disease: a comparison of three radiotracers. *Eur J Nucl Med Mol Imaging* **41**, 1398-1407 (2014).
 21. Mormino, E.C., *et al.* Amyloid and APOE epsilon4 interact to influence short-term decline in preclinical Alzheimer disease. *Neurology* **82**, 1760-1767 (2014).
 22. Li, Q.X., *et al.* Alzheimer's disease normative cerebrospinal fluid biomarkers validated in PET amyloid- β characterized subjects from the Australian Imaging, Biomarkers and Lifestyle (AIBL) study. *J Alzheimers Dis* **48**, 175-187 (2015).
 23. Hu, S., Loo, J.A. & Wong, D.T. Human body fluid proteome analysis. *Proteomics* **6**, 6326-6353 (2006).
 24. Huang, Y., *et al.* β -amyloid dynamics in human plasma. *Arch Neurol* **69**, 1591-1597 (2012).
 25. Kaneko, N., Yamamoto, R., Sato, T.A. & Tanaka, K. Identification and quantification of amyloid β -related peptides in human plasma using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Proc Jpn Acad Ser B Phys Biol Sci* **90**, 104-117 (2014).
 26. Pannee, J., *et al.* The amyloid- β degradation pattern in plasma--a possible tool for clinical trials in Alzheimer's disease. *Neuroscience letters* **573**, 7-12 (2014).
 27. Ovod, V. *et al.* Amyloid beta concentrations and stable isotope labeling kinetics of human plasma specific to central nervous system amyloidosis. *Alzheimers Dement* **13**, 841-849 (2017).
 28. Janelidze, S., *et al.* Plasma β -amyloid in Alzheimer's disease and vascular disease. *Scientific Reports* **6**, 26801 (2016).
 29. M  nning, U. *et al.* Alzheimer amyloid BA4-protein reactive antibodies in human sera and CSF. *Neurobiology of aging* **11**, 338 (1990).
 30. Kricka, L.J. Human anti-animal antibody interferences in immunological assays. *Clinical chemistry* **45**, 942-956 (1999).
 31. Perez-Grijalba, V., *et al.* Validation of Immunoassay-Based Tools for the Comprehensive Quantification of Abeta40 and Abeta42 Peptides in Plasma. *J Alzheimers Dis* **54**, 751-762

- (2016).
32. Pannee, J., *et al.* A selected reaction monitoring (SRM)-based method for absolute quantification of A β 38, A β 40, and A β 42 in cerebrospinal fluid of Alzheimer's disease patients and healthy controls. *J Alzheimers Dis* **33**, 1021-1032 (2013).
 33. Gelfanova, V. *et al.* Quantitative analysis of amyloid- β peptides in cerebrospinal fluid using immunoprecipitation and MALDI-ToF mass spectrometry. *Brief Funct Genomic Proteomic* **6**, 149-158, (2007).
 34. Laarakkers, C.M., *et al.* Improved mass spectrometry assay for plasma hepcidin: detection and characterization of a novel hepcidin isoform. *PloS one* **8**, e75518 (2013).
 35. Pan, S., *et al.* Application of targeted quantitative proteomics analysis in human cerebrospinal fluid using a liquid chromatography matrix-assisted laser desorption/ionization time-of-flight tandem mass spectrometer (LC MALDI TOF/TOF) platform. *Journal of proteome research* **7**, 720-730 (2008).
 36. Bucknall, M., Fung, K.Y. & Duncan, M.W. Practical quantitative biomedical applications of MALDI-TOF mass spectrometry. *Journal of the American Society for Mass Spectrometry* **13**, 1015-1027 (2002).
 37. Burdick, D., *et al.* Assembly and aggregation properties of synthetic Alzheimer's A4/beta amyloid peptide analogs. *The Journal of biological chemistry* **267**, 546-554 (1992).
 38. O'Bryant, S.E., *et al.* Guidelines for the standardization of preanalytic variables for blood-based biomarker studies in Alzheimer's disease research. *Alzheimers Dement* **11**, 549-560 (2015).
 39. Mattsson, N., *et al.* CSF biomarker variability in the Alzheimer's Association quality control program. *Alzheimers Dement* **9**, 251-261 (2013).
 40. Rowe, C.C., *et al.* Amyloid imaging results from the Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging. *Neurobiology of aging* **31**, 1275-1283 (2010).
 41. Kaneko, N., *et al.* Novel plasma biomarker surrogating cerebral amyloid deposition. *Proc Jpn Acad Ser B Phys Biol Sci* **90**, 353-364 (2014).
 42. Beyer, I., *et al.* Solid-phase synthesis and characterization of N-terminally elongated A β _{3-x} – peptides. *Chemistry* **22**, 8685-8693 (2016).
 43. LeVine, H., 3rd. Thioflavine T interaction with synthetic Alzheimer's disease β -amyloid peptides: detection of amyloid aggregation in solution. *Protein Sci* **2**, 404-410 (1993).
 44. Naiki, H. & Gejyo, F. Kinetic analysis of amyloid fibril formation. *Methods Enzymol* **309**, 305-318 (1999).
 45. Jarrett, J.T., Berger, E.P. & Lansbury, P.T., Jr. The C-terminus of the β protein is critical in amyloidogenesis. *Ann N Y Acad Sci* **695**, 144-148 (1993).