



Grey zone amyloid burden affects memory function: the SCIENCE project

J. L. Ebenau¹ · S. C. J. Verfaillie² · K. A. van den Bosch¹ · T. Timmers¹ · L. M. P. Wesselman¹ · M. van Leeuwenstijn¹ · H. Tuncel² · S. V. S. Golla² · M. M. Yaqub² · A. D. Windhorst² · N. D. Prins^{1,3} · F. Barkhof^{2,4} · P. Scheltens¹ · W. M. van der Flier^{1,5} · B. N. M. van Berckel^{1,2}

Received: 3 June 2020 / Accepted: 20 August 2020 / Published online: 4 September 2020

© The Author(s) 2020

Abstract

Purpose To determine thresholds for amyloid beta pathology and evaluate associations with longitudinal memory performance with the aim to identify a grey zone of early amyloid beta accumulation and investigate its clinical relevance.

Methods We included 162 cognitively normal participants with subjective cognitive decline from the SCIENCE cohort (64 ± 8 years, 38% F, MMSE 29 ± 1). Each underwent a dynamic [¹⁸F] florbetapir PET scan, a T1-weighted MRI scan and longitudinal memory assessments (RAVLT delayed recall, *n* = 655 examinations). PET scans were visually assessed as amyloid positive/negative. Additionally, we calculated the mean binding potential (BP_{ND}) and standardized uptake value ratio (SUV_{r50–70}) for an a priori defined composite region of interest. We determined six amyloid positivity thresholds using various data-driven methods (resulting thresholds: BP_{ND} 0.19/0.23/0.29; SUV_r 1.28/1.34/1.43). We used Cohen's kappa to analyse concordance between thresholds and visual assessment. Next, we used quantiles to divide the sample into two to five subgroups of equal numbers (median, tertiles, quartiles, quintiles), and operationalized a grey zone as the range between the thresholds (0.19–0.29 BP_{ND}/1.28–1.43 SUV_r). We used linear mixed models to determine associations between thresholds and memory slope.

Results As determined by visual assessment, 24% of 162 individuals were amyloid positive. Concordance with visual assessment was comparable but slightly higher for BP_{ND} thresholds (range kappa 0.65–0.70 versus 0.60–0.63). All thresholds predicted memory decline (range beta – 0.29 to – 0.21, all *p* < 0.05). Analyses in subgroups showed memory slopes gradually became steeper with higher amyloid load (all *p* for trend < 0.05). Participants with a low amyloid burden benefited from a practice effect (i.e. increase in memory), whilst high amyloid burden was associated with memory decline. Memory slopes of individuals in the grey zone were intermediate.

Conclusion We provide evidence that not only high but also grey zone amyloid burden subtly impacts memory function. Therefore, in case a binary classification is required, we suggest using a relatively low threshold which includes grey zone amyloid pathology.

Keywords [¹⁸F] florbetapir · Amyloid · Subjective cognitive decline · Grey zone · Cognition

This article is part of the Topical Collection on Neurology

✉ J. L. Ebenau
j.ebenau@amsterdamumc.nl

¹ Alzheimer Centre, Department of Neurology, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, The Netherlands

² Department of Radiology & Nuclear Medicine, Amsterdam Neuroscience, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, The Netherlands

³ Brain Research Centre, Amsterdam, The Netherlands

⁴ UCL Institutes of Neurology and Healthcare Engineering, London, UK

⁵ Department of Epidemiology & Biostatistics, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, The Netherlands

Introduction

The presence of pathological amyloid beta depositions is one of the hallmarks of Alzheimer's disease (AD) and amyloid pathology is thought to play an important role in its pathophysiology [1, 2]. Indeed, high amyloid burden in cognitively normal individuals is associated with a greater risk of cognitive decline, particularly of memory function [3–9]. Furthermore, individuals with subjective cognitive decline (SCD) are more often amyloid positive than the general population and are at increased risk of cognitive decline and dementia [10, 11]. Therefore, individuals with SCD form an ideal population to study the effects of 'early' amyloid deposition on cognition.

Amyloid beta pathology can be assessed in vivo by [^{18}F] florbetapir positron emission tomography (PET) using visual assessment in a dichotomous manner, i.e. positive versus negative [12]. However, the accuracy depends on the expertise of the trained reader [13], and visual assessment of scans with early amyloid accumulation can be challenging. Classification of amyloid positivity can also be determined using a threshold as an alternative to visual assessment. The standardized uptake value ratio (SUVr) is a widely used method for estimating amyloid burden semi-quantitatively using a static scan procedure. Dynamic scanning allows for calculation of binding potential (BP_{ND}) which provides a more exact quantification of specific binding to amyloid beta [14]. BP_{ND} has been shown to be less prone to overestimation compared with SUVr and is more reliable when studying early amyloid accumulation [15–17]. So far, different SUVr thresholds for [^{18}F] florbetapir have been proposed [18–23], but these thresholds are highly variable (range SUVr thresholds 1.08–1.34). BP_{ND} thresholds have not been published yet.

Dichotomizing amyloid burden into a negative and positive status can be useful in clinical and research settings, but it disregards the potential significance of early (pathological) amyloid accumulation [24]. Recent studies show that even in individuals that are initially labelled as amyloid negative, the amyloid accumulation slope is associated with memory decline [25, 26]. It is uncertain whether this suggests that current thresholds are simply too high and lower thresholds would be able to correctly classify individuals, or that there is a more gradual association between memory decline and amyloid burden. The latter would point towards a ‘dose-dependent risk’ with a grey zone amyloid burden reflecting an at-risk state for AD.

In the current study, we aimed to define thresholds for amyloid positivity using data-driven methods based on both SUVr and BP_{ND} . Subsequently, we compared each of these classifications with visual assessment of amyloid positivity and determined associations with memory function over time. In addition, we identified a ‘grey zone’ of amyloid burden in cognitively normal individuals, and investigated its clinical significance, by exploring the nature of the relationship between amyloid levels in the subthreshold range and memory slope.

Method

Population

We included 162 cognitively normal participants with SCD from the Subjective Cognitive Impairment Cohort (SCIENCE) within the Amsterdam Dementia Cohort at the Alzheimer Centre Amsterdam [27, 28]. All subjects with [^{18}F] florbetapir PET, magnetic resonance imaging (MRI) scan and cognitive

data available were included. One hundred and fifty-two participants were referred to the memory clinic by their general physician, a neurologist or a geriatrician, and underwent an extensive standardized diagnostic workup that included a neurologic and neuropsychological examination, laboratory testing and brain MRI [28, 29]. In a consensus meeting, participants were labelled SCD when cognitive performance appeared within normal limits compared with peers, and criteria were not met for mild cognitive impairment (MCI), dementia or other neurological or psychiatric diseases that could possibly cause cognitive complaints. At annual follow-up visits, neuropsychological testing was repeated and diagnoses were re-evaluated. In addition, 10 participants were included as research participants via the Dutch Brain Research Registry (hersenonderzoek.nl). They also experienced cognitive complaints in absence of a diagnosis of MCI or dementia, and received the same baseline workup.

Neuropsychological assessment

We previously showed that the relationship between amyloid burden and cognitive decline was strongest for the memory domain, especially for the Rey auditory verbal learning task (RAVLT) delayed recall [7]. Therefore, for this study, we used the RAVLT delayed recall as a measure for memory function. We used visits conducted before as well as after the PET scan to accurately estimate the memory slope, resulting in longitudinal cognitive data covering 3.8 ± 3.1 years. Concurrent time points were defined as the visit closest to the PET scan date (median -0.19 (IQR -0.38 – 0.14 years)). We used two different versions of the RAVLT, between which we alternated at the annual follow-up visits. In total, 655 neuropsychological examinations of 162 participants were available ($149 \geq 2$ visits; range 1–10; median 3 visits).

Questionnaires

Within the SCIENCE cohort, a number of questionnaires are administered to evaluate subjective cognitive complaints, mental health, instrumental activities of daily living and lifestyle [27]. For this study, we used the cognitive change index (CCI, 20 questions, range 0–80) to quantify the degree of subjective cognitive complaints. We additionally used the geriatric depression scale (GDS, 15 questions, range 0–15) to evaluate depressive symptoms. For both questionnaires, a higher score reflects more severe symptoms.

PET acquisition and image analysis

PET scans were acquired on an Ingenuity PET-CT ($n = 115$) or a Gemini TF PET-CT ($n = 47$; Philips, Best, the Netherlands) scanner. Dynamic PET emission scans of

90 min ($n = 137$) were obtained starting directly after tracer injection of approximately 370 M Becquerel (MBq) [^{18}F] florbetapir. During the course of the study, our group showed that the scan duration could be reduced without compromising the reliability of results [14]. Therefore, the more recent scans ($n = 21$) had a duration of 70 min. Furthermore, in four participants, the scan was terminated early (three after 60 min, one after 79 min) due to participant-related issues. These scans were still used since they had an uninterrupted 60 min of scanning [14]. Head movement when lying in the camera was monitored with laser beams, and if necessary, the position of the head was corrected. Data were reconstructed with a standard LOR RAMLA reconstruction algorithm into 22 frames, and images were corrected for scatter, random coincidences, attenuation, decay and dead time. Images were reconstructed with a matrix size of $128 \times 128 \times 90$ and a voxel size of $2 \times 2 \times 2 \text{ mm}^3$. Isotropic 3-dimensional T1-weighted MR images (GE Discovery MR750 3 T ($n = 58$), PETMR 3 T ($n = 71$), Signa 1.5 T ($n = 6$), Signa 3 T ($n = 2$), Titan 3 T ($n = 24$) and external scan ($n = 1$)) were co-registered to PET images using the Vinci software (Max Planck Institute, Cologne, Germany). Regions of interest (ROIs) were defined on the co-registered MRI using the Hammers probability atlas [30] in PVElab. Receptor parametric mapping (RPM) was used to generate BP_{ND} images with cerebellar grey matter as a reference region [16, 31, 14, 17]. We extracted BP_{ND} and SUVr_{50-70} values in the following a priori defined regions: orbitofrontal, temporal, parietal, anterior cingulate, posterior cingulate and precuneus [21]. An SUVr time interval of 50–70 min post-injection was chosen because this is commonly used, and our group showed before that SUVr becomes constant from 40 min onward [14, 18]. We subsequently averaged the values of the a priori defined regions into one volume weighted mean cortical BP_{ND} or SUVr value. The difference in time between MRI and PET was generally within 1 year (median time difference 0.22 years (IQR = 0.49–0.55)).

Threshold derivation

SUV images were visually assessed as ‘positive’ or ‘negative’ by a trained and experienced nuclear medicine physician (BvB) who was blinded for clinical information, based on standards provided by the manufacturer [32]. Next, we used different data-driven methods to obtain thresholds for amyloid positivity for both BP_{ND} and SUVr . First we used the R studio function `normalmixEM` to fit Gaussian mixture models (GMM) with 1–9 components. Bayesian information criterion (BIC) indicated a model with 2 components as being the most optimal fit to our data. A threshold was derived representing the mean of the calculated μ of both components. The calculated thresholds were similar when we used the proportions derived from visual assessment (24% and 76%) as a starting

value for mixture weights. This resulted in cut-off points of 0.23 (BP_{ND}) and 1.34 (SUVr).

Next, we used K-means clustering. We assumed the data consisted of two clusters. We derived two cut-off values, the first representing the 90th percentile of the cluster with low amyloid burden, and the second representing the 10th percentile of the cluster with high amyloid burden. The cut-off values were purely data-driven, and information about visual assessment of scans was not used for these thresholds. This resulted in a low threshold (0.19 BP_{ND} and 1.28 SUVr), and a high threshold (0.29 BP_{ND} and 1.43 SUVr). Subsequently, we took the area between the lower and higher thresholds derived by K-means clustering to operationalize a grey zone. Figure 1 shows a summary of all derived thresholds and visualizes the predefined grey zone.

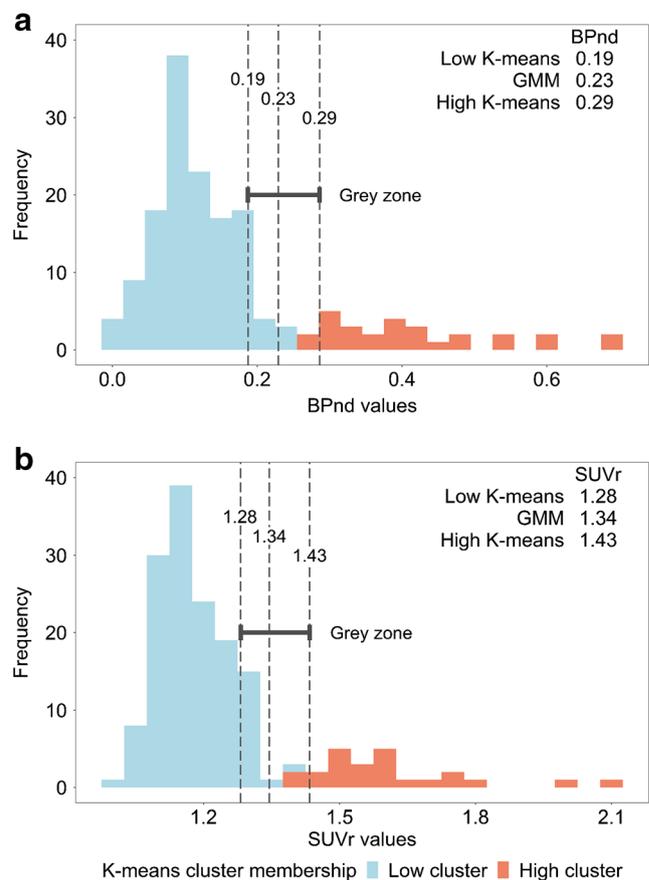


Fig. 1 Visualization of thresholds and grey zone. Two frequency histograms containing all mean BP_{ND} values (**a**) and SUVr values (**b**), with K-means cluster membership visualized by different colours (blue and red). The dashed lines represent the 6 different thresholds (for BP_{ND} : 0.19, 0.23 and 0.29, for SUVr : 1.28, 1.34 and 1.43). The grey zone was operationalized as the range between the lowest and the highest thresholds derived through K-means clustering. For BP_{ND} , 121 participants had an amyloid burden lower than the low K-means threshold, 15 participants had grey zone amyloid burden and 26 participants had an amyloid burden higher than the high K-means threshold. For SUVr , the numbers were 125, 15 and 22 respectively

Statistics

We used *t* test, Mann-Whitney *U* and chi-square where appropriate to compare demographic measures between amyloid positive and amyloid negative groups, based on visual assessment. We used Cohen's kappa to determine the degree of concordance between visual assessment on the one hand and the six thresholds on the other hand. We used linear mixed models (LMM) to assess the associations between amyloid status (visual and data-driven) and memory slopes. Separate models were run for each of the seven ways of defining amyloid positivity. Amyloid status, time and the interaction between amyloid status and time were included as independent variables, age, sex, education and scanner type were included as covariates, and scores on the RAVLT delayed memory task were used as dependent variables. Intercept and time were included as random factors, as this resulted in a better fit. Using these models, we estimated the annual change over time for both a negative and a positive amyloid status. We compared models based on betas, *p* values and Akaike information criterion (AIC).

Subsequently, using an increasing number of quantiles, we divided the sample into two, three, four and five equal-sized distributions (i.e. subgroups) to explore whether there is a gradual association between amyloid burden and memory slope. We subsequently used LMM to estimate memory slopes for each subgroup. Separate analyses were run for each quantile-based division. Subgroups (entered as dummies), time and the interaction between subgroups and time were included as independent variables, age, sex, education and scanner type were included as covariates, and RAVLT delayed recall score was used as dependent variable. In addition, we ran models including subgroups as continuous variables, and present the resulting *p* value for trend.

All analyses were done using SPSS version 26 and R studio version 1.1.463. For the estimated trends, we used the R studio function of *emtrends*. *p* values < 0.05 were considered significant.

Results

Baseline demographics

At baseline, individuals were on average 64 ± 8 years old, 63 (39%) were female and MMSE was 29 ± 1 (Table 1). Among 162 individuals, 38 (24%) were amyloid positive as defined by visual assessment. Amyloid positive (A+) individuals were on average older and more often APOE4 carrier than amyloid negative (A-) individuals.

We presented thresholds, frequency and kappa values in Table 2. When we applied the different thresholds, the amyloid positivity rates ranged from 26 (16%) to 41 (25%) for

BP_{ND} thresholds, and from 22 (14%) to 37 (23%) for SUVR thresholds. For BP_{ND} as well as SUVR, the low K-means threshold resulted in the highest percentage of A+ individuals, and the high K-means threshold in the lowest percentage of A+ individuals. The grey zone, operationalized as the range between the lowest and highest K-means thresholds, consisted of 9% of individuals for both BP_{ND} and SUVR. Cohen's kappa showed that there was substantial concordance between visual assessment and each of the six thresholds. Upon visual inspection, concordance was highest for BP_{ND} thresholds, but confidence intervals overlapped.

Amyloid positivity thresholds in relation to memory slopes

We investigated the association between different definitions of amyloid positivity and memory slopes. We found each operationalization of amyloid positivity was associated with rate of decline on the RAVLT delayed recall (Table 3). Models in which A+ was defined by visual assessment or BP_{ND} thresholds performed somewhat better than models based on SUVR (i.e. lower AIC values).

Relationship between grey zone amyloid burden and memory slope

Next, we categorised participants based on an increasing number of quantiles to evaluate whether the association between amyloid burden and memory slope is based on a gradual change in amyloid burden. For all models, a gradually lower annual memory performance is seen with increasing amyloid levels (all *p* for trend < 0.05). Subgroups with the lowest amyloid burden (1st half, 1st third, 1st quarter, 1st fifth and K-low) showed a practice effect, with an increase in memory performance over time, whilst subgroups with the highest amyloid burden (3rd third, 4th quarter, 5th fifth and K-high) showed memory decline over time (Fig. 2). Memory slopes of individuals in the grey zone were intermediate, with betas closer to zero or negative, shown most clearly in the grey zone (0.19–0.29 BP_{ND}) and the 4th fifth (0.14–0.22 BP_{ND}, 0.21–1.31 SUVR, Fig. 2).

Discussion

In this sample of cognitively normal individuals with SCD, we observed that grey zone amyloid burden contains relevant clinical information. Furthermore, we obtained thresholds for amyloid positivity based on both SUVR and BP_{ND}, which corresponded well with visual assessment of amyloid deposition.

We investigated the association between grey zone amyloid burden and memory function. We found that cognitively

Table 1 Baseline demographics by amyloid status based on visual assessment

	Total <i>n</i> = 162	Amyloid negative <i>n</i> = 124	Amyloid positive <i>n</i> = 38
Age, mean (SD)	64 (8)	63 (8)	68 (8)*
Sex, <i>n</i> female (%)	63 (39%)	47 (38%)	16 (42%)
Education, mean (SD) ^a	6 (1)	6 (1)	6 (1)
APOE4 status, <i>n</i> carrier (%)	54 (36%)	31 (26%)	23 (68%)*
CCI, mean (SD) ^{a,b}	21.8 (14.5)	21.4 (15.0)	23.0 (13.1)
GDS, mean (SD) ^{a,b}	2.4 (2.0)	2.5 (2.1)	2.3 (1.9)
MMSE, mean (SD) ^{a,b}	28.9 (1.2)	28.9 (1.2)	28.7 (1.2)
RAVLT delayed, mean (SD) ^b	9.1 (3.0)	9.3 (3.0)	8.3 (3.2)
Amyloid load (BP _{ND}), mean (SD)	0.16 (0.13)	0.11 (0.06)	0.33 (0.18)*
Amyloid load (SUV _r), mean (SD)	1.24 (0.18)	1.17 (0.08)	1.45 (0.26)*

Amyloid status was determined by visual assessment of the [¹⁸F] florbetapir PET scan. Education is rated using the Dutch Verhage system [33]. Amyloid load represents the volume-weighted mean cortical value in a composite region of a priori defined regions (orbitofrontal, temporal, parietal, anterior cingulate, posterior cingulate and precuneus), with cerebellar grey matter as reference region. CCI and GDS test scores were available for 159 and 96 participants respectively

CCI cognitive change index, SCF subjective cognitive functioning, GDS geriatric depression scale, MMSE mini-mental state examination, RAVLT Rey auditory verbal learning task, BP_{ND} binding potential, SUV_r standardized uptake value ratio

^a Mann-Whitney *U* test. All other analyses were performed using *t* test and Chi-square

^b Score on concurrent test

* *p* < 0.01 for difference between amyloid negative and positive individuals

healthy individuals with low amyloid levels showed improved memory performance over time, which could be due to a practice effect. By contrast, individuals with substantial amyloid burden showed memory decline over time. Individuals with grey zone amyloid burden had slopes in between, showing neither decline, nor improvement in memory. This implies that these individuals did not benefit from a practice effect, like amyloid negative individuals do. The absence of a practice effect is not an innocent finding, as it has previously been demonstrated as a predictor of future deterioration [34–37]. Although for all subgroups, the estimated annual change was relatively small, the fact that differences could already be measured in this very early stage provides evidence for

the concept of a grey zone. Furthermore, some individuals might already experience a subclinical decline in test scores, whilst the test scores themselves are still within normal limits. This illustrates the relevance of longitudinal research to capture within-subject changes over time. Comparison with other studies is complicated because there is not one universal grey zone definition. Studies that focused on peri- or subthreshold amyloid levels have had different approaches, for example studying amyloid negative subthreshold individuals [25, 26, 38], or CSF/PET discordant cases [39]. In a recent article, the grey zone is proposed as a region of uncertainty around the threshold for which more data is needed to actually estimate the risk of cognitive decline or clinical progression [40]. These

Table 2 Cut-off values for different methods

	Derivation method	Threshold	<i>n</i> positive (%)	Kappa (95% CI)
Visual assessment			38 (24%)	
BP _{ND}	Low K-means	0.19	41 (25%)	0.65 (0.51–0.79)
	GMM	0.23	30 (19%)	0.70 (0.57–0.84)
	High K-means	0.29	26 (16%)	0.65 (0.51–0.80)
SUV _r	Low K-means	1.28	37 (23%)	0.60 (0.45–0.75)
	GMM	1.34	25 (15%)	0.63 (0.48–0.78)
	High K-means	1.43	22 (14%)	0.60 (0.44–0.75)

Cohen’s kappa was used to determine the degree of concordance between visual assessment and the six different thresholds

BP_{ND} binding potential, GMM Gaussian mixture modelling, SUV_r standardized uptake value ratio

Table 3 Relationship between different amyloid positivity thresholds and longitudinal performance on a memory task

	Threshold	Amyloid status	Estimated annual change	AIC
Visual assessment		Negative	0.19 (0.05)	2933.7
		Positive	− 0.28 (0.09)**	
BP _{ND}	0.19	Negative	0.19 (0.05)	2938.5
		Positive	− 0.22 (0.08)**	
	0.23	Negative	0.17 (0.05)	2935.8
		Positive	− 0.28 (0.11)**	
	0.29	Negative	0.15 (0.05)	2938.2
		Positive	− 0.28 (0.12)**	
SUV _r	1.28	Negative	0.16 (0.05)	2941.2
		Positive	− 0.21 (0.10)**	
	1.34	Negative	0.14 (0.05)	2940.0
		Positive	− 0.28 (0.12)**	
	1.43	Negative	0.14 (0.05)	2943.0
		Positive	− 0.29 (0.13)**	

Values given are beta (SE), as estimated by linear mixed models (predictor: amyloid status, outcome: score on RAVLT delayed recall). Numbers reflect annual change in raw score points. Models are adjusted for age, sex, education and scanner type

AIC Akaike information criterion, SE standard error, RAVLT Rey auditory verbal learning task, BP_{ND} binding potential, SUV_r standardized uptake value ratio

** p value < 0.01. p value represents the significance of the difference between a positive amyloid status compared with a negative amyloid status

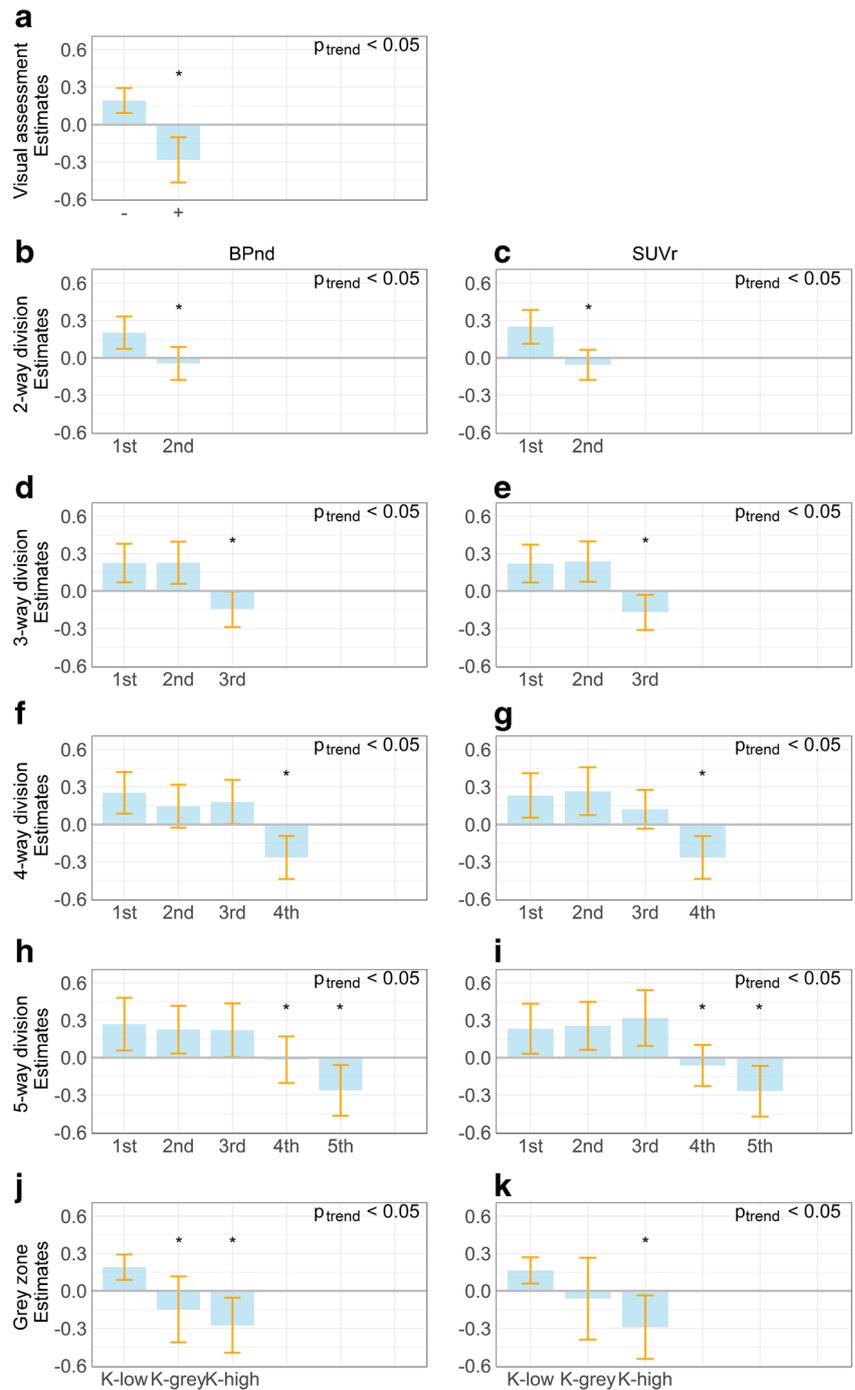
previous studies found that individuals in the subthreshold range can be on the path to further neurodegeneration (i.e. atrophy, tau pathology, hypometabolism) [38, 41], and are at risk of further amyloid accumulation, cognitive decline and clinical progression [25, 26, 39]. In the present study, we defined the grey zone making use of two thresholds obtained in a data-driven way. In a second approach, we subdivided the data using divisions based on quantiles. Irrespective of the approach, our findings showed that the negative relationship between amyloid and memory performance is not merely driven by the small number of individuals with high amyloid burden, but rather that the variability in amyloid burden, even within normal limits, has potential clinical value.

We used different data-driven methods, such as Gaussian mixture modelling and K-means clustering, to derive cut-off values for amyloid positivity. We found thresholds of 0.19, 0.23 and 0.29 for BP_{ND}, and thresholds of 1.28, 1.34 and 1.43 for SUV_r. Literature has generated inconsistent findings with respect to amyloid thresholds, ranging from 1.08 to 1.34 for SUV_r, with 1.10 being reported most frequently [12, 18–23, 42–45]. The large variability indicates that thresholds may to some extent rely on methodology, image processing pipeline used and study sample. For example, the partial volume correction method [46] and the choice of ROIs [47] affect the degree of amyloid burden. For this reason, we used a commonly used ‘meta ROI’ [21, 24, 25], which is able to clearly distinguish AD patients from cognitively normal controls. However, small differences can be seen across studies [26, 48, 49]. In addition, thresholds are dependent on sample

characteristics [50]. We aimed to minimize this effect with our choice of robust data-driven methods. Although our thresholds seemed substantially higher than the aforementioned thresholds, all corresponded equally well to visual assessment. We show that dichotomized BP_{ND} values may even correspond to visual assessment somewhat better, which is consistent with the findings of a previous [¹⁸F] flutemetamol PET study [13]. All amyloid positivity thresholds predicted future memory decline, which is consistent with another study [51], although models with BP_{ND} thresholds and visual assessment seemingly resulted in a slightly better fit. Because of the underlying gradual association between amyloid burden and memory function, apparently the height of the threshold does not necessarily have a substantial effect on the association between amyloid positivity and memory function.

Strengths of this study include that we used two measures of amyloid quantification, BP_{ND} and SUV_r, and that we applied various data-driven approaches. BP_{ND} has been shown to be less sensitive to differences in flow and we found a good concordance with visual assessment. Using BP_{ND} and SUV_r as continuous measures enabled us to thoroughly explore the grey zone, which is not possible with a strict binary division like visual assessment. Furthermore, we had a large, well-defined cohort, with a relatively long follow-up. Limitations include the lack of a gold standard such as pathology confirmation. Notwithstanding, we used visual assessment for comparison analyses which has been shown to correlate very well with pathology [12, 52]. Furthermore, we used memory decline as outcome measure, as opposed to clinical progression

Fig. 2 Estimated longitudinal change on RAVLT delayed recall. Bar graphs showing estimated longitudinal change for performance on RAVLT delayed recall over time. The sample was divided into subgroups using visual assessment (a), an increasing number of quantiles (b–i), and the predefined grey zone (j–k). **a** Visual assessment. **b** Two-way division (BP_{ND}). **c** Two-way division (SUV_r). **d** Three-way division (BP_{ND}). **e** Three-way division (SUV_r). **f** Four-way division (BP_{ND}). **g** Four-way division (SUV_r). **h** Five-way division (BP_{ND}). **i** Five-way division (SUV_r). **j** Grey zone (according to K-means thresholds) (BP_{ND}). **k** Grey zone (SUV_r). Bars represent predicted annual change in raw test score, and error bars represent 95% confidence interval. P_{trend} represents p value for trend. *: p value represents significance (< 0.05) of the difference between the subgroup under investigation and the reference category within the subgroup division (1st subgroup)



to MCI or dementia. This might have led to less robust results because memory performance may be reversible, particularly in cases with limited amyloid burden. On the other hand, it might take a relatively long time before a substantial part of this sample shows clinical progression or cognitive impairment. Nevertheless, since we had cognitive data covering on average 3.8 years, our models should give an accurate estimation of memory slope. Using these methods, we were able to capture subtle decline, and found evidence for a diminished

practice effect in grey zone amyloid burden. Last, although BP_{ND} and SUV_r are widely used continuous measures, it is difficult to compare tracers and sites due to the large variability in methodology previously mentioned. The aim of the Centiloid Project is to provide a method with which this is possible. In this method, all outcome measures are standardized by scaling them to a 0 to 100 scale [53, 54]. For future research, it would be of interest to explore how our results translate to the Centiloid scale [55].

Our demonstration of the potential significance of grey zone amyloid burden may have several clinical consequences. Especially for individuals with amyloid burden within the grey zone, a single threshold might not be very good at distinguishing individuals with a high and a low risk of cognitive decline [4, 40]. When a binary division is warranted, our results imply that only lower thresholds that include the grey zone capture all individuals at risk of memory decline, which corresponds to a previous study that showed existing thresholds for Pittsburgh compound B (PIB) seem too high [24]. When a high threshold is used (e.g. 0.29 BP_{ND}) that classifies 16% as amyloid positive, 9% of individuals that are actually also at risk are labelled amyloid negative. This means a total of 25% of individuals is at risk of future deterioration. When considering that the 4th subgroup of the 5-way division already demonstrates a diminished practice effect, even up to 40% might be at risk (4th and 5th subgroup together). This could have consequences for clinical trials that only include amyloid positive individuals. Excluding grey zone individuals would lower recruitment rates and means loss of valuable information. In addition, these subjects could benefit the most from disease modifying drugs as they are very early in the disease course.

In summary, we showed that various thresholds correspond well to visual assessment in our sample, particularly BP_{ND} thresholds. We furthermore show that not only a high amyloid burden but also grey zone amyloid burden has an effect on longitudinal memory function. We therefore suggest, when the same methodology is used, to use a low BP_{ND} threshold of 0.19 when a binary classification is needed, to also include the grey zone.

Acknowledgements Research of Alzheimer Centre Amsterdam is part of the neurodegeneration research program of Amsterdam Neuroscience.

Funding Open access funding provided by Amsterdam UMC (Vrije Universiteit Amsterdam). Alzheimer Centre Amsterdam is supported by Stichting Alzheimer Nederland and Stichting VUmc fonds. Wiesje van der Flier holds the Pasman chair. The clinical database structure was developed with funding from Stichting Dioraphte. Frederik Barkhof is supported by the NIHR biomedical research centre at UCLH. The SCIENCe project is supported by a research grant from Gieskes-Strijbis fonds. PET scans were funded by a research grant from AVID.

Data availability Any data used within the article may be shared upon reasonable request.

Compliance with ethical standards

Conflict of interest J.L. Ebenau reports no conflict of interest. Dr. Sander C.J. Verfaillie reports no conflict of interest. Karlijn A. van den Bosch reports no conflict of interest. Tessa Timmers reports no conflict of interest. Dr. Linda M.P. Wesselman reports no conflict of interest. Mardou van Leeuwenstijn reports no conflict of interest. Hayel Tuncel reports no conflict of interest. Dr. Sandeep V.S. Golla reports no conflict of interest. Dr. Maqsood M. Yaqub reports no conflict of interest. Dr.

Albert D. Windhorst reports no conflict of interest. Dr. Niels D. Prins reports consulting, advisory and speaker fees from Boehringer Ingelheim, Envivo, Janssen, Novartis, Probiadrug, Sanofi, Takeda, Kyowa Kirin Pharmaceutical Development, DSMB of AbbVie's M15-566, grants from Alzheimer Nederland (all paid directly to his institution) outside the submitted work. Dr. Prins is CEO and co-owner of the Brain Research Centre, Amsterdam, The Netherlands. Dr. Frederik Barkhof is a consultant for Biogen-Idec, Janssen Alzheimer Immunotherapy, Bayer-Schering, Merck-Serono, Roche, Novartis, Genzyme and Sanofi-Aventis; has received sponsoring from European Commission-Horizon 2020, National Institute for Health Research-University College London Hospitals Biomedical Research Centre, Scottish Multiple Sclerosis Register, TEVA, Novartis and Toshiba; and serves on the editorial boards of *Radiology*, *Brain*, *Neuroradiology*, *Multiple Sclerosis Journal*, and *Neurology*. Dr. Philip Scheltens has acquired grant support (for the institution) from Biogen. In the past 2 years, he has received consultancy/speaker fees (paid to the institution) from Probiadrug Biogen, EIP Pharma, Merck AG. Dr. Wiesje M. van der Flier's research programs have been funded by ZonMW, the Netherlands Organization of Scientific Research, Alzheimer Nederland, Cardiovascular Onderzoek Nederland, Stichting Dioraphte, Gieskes-Strijbis fonds, Pasman stichting, Boehringer Ingelheim, Life-MI, AVID, Biogen MA and Combinostics. All funding is paid to her institution. Dr. Bart van Berckel has received funding from ZonMW, the Netherlands Organization of Scientific Research, the Centre of Translational Molecular Imaging and Avid Radiopharmaceuticals. All funding is paid to his institution.

Ethics approval The research was approved by the Medical Ethics Review Committee of Amsterdam UMC. This study was performed in line with the principles of the Declaration of Helsinki.

Consent to participate Written informed consent was obtained from all patients included in the study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA Research Framework: toward a biological definition of Alzheimer's disease. *Alzheimers Dement*. 2018;14(4):535–62. <https://doi.org/10.1016/j.jalz.2018.02.018>.
2. Selkoe DJ, Hardy J. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol Med*. 2016;8(6):595–608. <https://doi.org/10.15252/emmm.201606210>.
3. Villemagne VL, Burnham S, Bourgeat P, Brown B, Ellis KA, Salvado O, et al. Amyloid beta deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *Lancet Neurol*. 2013;12(4):357–67. [https://doi.org/10.1016/s1474-4422\(13\)70044-9](https://doi.org/10.1016/s1474-4422(13)70044-9).

4. Jack CR Jr, Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS, et al. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol.* 2013;12(2):207–16. [https://doi.org/10.1016/s1474-4422\(12\)70291-0](https://doi.org/10.1016/s1474-4422(12)70291-0).
5. van Harten AC, Smits LL, Teunissen CE, Visser PJ, Koene T, Blankenstein MA, et al. Preclinical AD predicts decline in memory and executive functions in subjective complaints. *Neurology.* 2013;81(16):1409–16. <https://doi.org/10.1212/WNL.0b013e3182a8418b>.
6. van Harten AC, Visser PJ, Pijnenburg YA, Teunissen CE, Blankenstein MA, Scheltens P, et al. Cerebrospinal fluid Abeta42 is the best predictor of clinical progression in patients with subjective complaints. *Alzheimers Dement.* 2013;9(5):481–7. <https://doi.org/10.1016/j.jalz.2012.08.004>.
7. Timmers T, Ossenkuppele R, Verfaillie SCJ, van der Weijden CWJ, Slot RER, Wesselman LMP, et al. Amyloid PET and cognitive decline in cognitively normal individuals: the SCIENCE project. *Neurobiol Aging.* 2019;79:50–8. <https://doi.org/10.1016/j.neurobiolaging.2019.02.020>.
8. Vos SJB, Xiong C, Visser PJ, Jaseleec MS, Hassenstab J, Grant EA, et al. Preclinical Alzheimer's disease and its outcome: a longitudinal cohort study. *Lancet Neurol.* 2013;12(10):957–65. [https://doi.org/10.1016/S1474-4422\(13\)70194-7](https://doi.org/10.1016/S1474-4422(13)70194-7).
9. Mormino EC, Betensky RA, Hedden T, Schultz AP, Amariglio RE, Rentz DM, et al. Synergistic effect of beta-amyloid and neurodegeneration on cognitive decline in clinically normal individuals. *JAMA Neurol.* 2014;71(11):1379–85. <https://doi.org/10.1001/jamaneurol.2014.2031>.
10. Buckley RF, Sikkes S, Villemagne VL, Mormino EC, Rabin JS, Burnham S, et al. Using subjective cognitive decline to identify high global amyloid in community-based samples: a cross-cohort study. *Alzheimers Dement.* 2019;11:670–8. <https://doi.org/10.1016/j.dadm.2019.08.004>.
11. Slot RER, Sikkes SAM, Berkhof J, Brodaty H, Buckley R, Cavado E, et al. Subjective cognitive decline and rates of incident Alzheimer's disease and non-Alzheimer's disease dementia. *Alzheimers Dement.* 2019;15(3):465–76. <https://doi.org/10.1016/j.jalz.2018.10.003>.
12. Clark CM, Pontecorvo MJ, Beach TG, Bedell BJ, Coleman RE, Doraiswamy PM, et al. Cerebral PET with florbetapir compared with neuropathology at autopsy for detection of neuritic amyloid-beta plaques: a prospective cohort study. *Lancet Neurol.* 2012;11(8):669–78. [https://doi.org/10.1016/s1474-4422\(12\)70142-4](https://doi.org/10.1016/s1474-4422(12)70142-4).
13. Collij LE, Konijnenberg E, Reimand J, Kate MT, Braber AD, Alves IL, et al. Assessing amyloid pathology in cognitively normal subjects using (18)F-flutemetamol PET: comparing visual reads and quantitative methods. *J Nucl Med.* 2019;60(4):541–7. <https://doi.org/10.2967/jnumed.118.211532>.
14. Golla SS, Verfaillie SC, Boellaard R, Adriaanse SM, Zwan MD, Schuit RC, et al. Quantification of [(18)F]florbetapir: a test-retest tracer kinetic modelling study. *J Cereb Blood Flow Metab.* 2019;39(11):2172–80. <https://doi.org/10.1177/0271678x18783628>.
15. van Berckel BN, Ossenkuppele R, Tolboom N, Yaqub M, Foster-Dingley JC, Windhorst AD, et al. Longitudinal amyloid imaging using 11C-PiB: methodologic considerations. *J Nucl Med.* 2013;54(9):1570–6. <https://doi.org/10.2967/jnumed.112.113654>.
16. Lammertsma AA. Forward to the past: the case for quantitative PET imaging. *J Nucl Med.* 2017;58(7):1019–24. <https://doi.org/10.2967/jnumed.116.188029>.
17. Verfaillie SCJ, Golla SSV, Timmers T, Tuncel H, van der Weijden CWJ, Schober P, et al. Repeatability of parametric methods for [18F] florbetapir imaging in Alzheimer's disease and healthy controls: a test-retest study. *J Cereb Blood Flow Metab.* 2020;0271678X20915403. <https://doi.org/10.1177/0271678X20915403>.
18. Joshi AD, Pontecorvo MJ, Clark CM, Carpenter AP, Jennings DL, Sadowsky CH, et al. Performance characteristics of amyloid PET with florbetapir F 18 in patients with Alzheimer's disease and cognitively normal subjects. *J Nucl Med.* 2012;53(3):378–84. <https://doi.org/10.2967/jnumed.111.090340>.
19. Chiotis K, Saint-Aubert L, Boccardi M, Gietl A, Picco A, Varrone A, et al. Clinical validity of increased cortical uptake of amyloid ligands on PET as a biomarker for Alzheimer's disease in the context of a structured 5-phase development framework. *Neurobiol Aging.* 2017;52:214–27. <https://doi.org/10.1016/j.neurobiolaging.2016.07.012>.
20. Landau SM, Breault C, Joshi AD, Pontecorvo M, Mathis CA, Jagust WJ, et al. Amyloid-beta imaging with Pittsburgh compound B and florbetapir: comparing radiotracers and quantification methods. *J Nucl Med.* 2013;54(1):70–7. <https://doi.org/10.2967/jnumed.112.109009>.
21. Joshi AD, Pontecorvo MJ, Lu M, Skovronsky DM, Mintun MA, Devous MD Sr. A semiautomated method for quantification of F 18 florbetapir PET images. *J Nucl Med.* 2015;56(11):1736–41. <https://doi.org/10.2967/jnumed.114.153494>.
22. Camus V, Payoux P, Barre L, Desgranges B, Voisin T, Tauber C, et al. Using PET with 18F-AV-45 (florbetapir) to quantify brain amyloid load in a clinical environment. *Eur J Nucl Med Mol Imaging.* 2012;39(4):621–31. <https://doi.org/10.1007/s00259-011-2021-8>.
23. Fleisher AS, Chen K, Liu X, Roontiva A, Thiyyagura P, Ayutyanont N, et al. Using positron emission tomography and florbetapir F18 to image cortical amyloid in patients with mild cognitive impairment or dementia due to Alzheimer disease. *Arch Neurol.* 2011;68(11):1404–11. <https://doi.org/10.1001/archneurol.2011.150>.
24. Villeneuve S, Rabinovici GD, Cohn-Sheehy BI, Madison C, Ayakta N, Ghosh PM, et al. Existing Pittsburgh Compound-B positron emission tomography thresholds are too high: statistical and pathological evaluation. *Brain.* 2015;138(Pt 7):2020–33. <https://doi.org/10.1093/brain/awv112>.
25. Landau SM, Horng A, Jagust WJ. Memory decline accompanies subthreshold amyloid accumulation. *Neurology.* 2018;90(17):e1452–e60. <https://doi.org/10.1212/wnl.0000000000005354>.
26. Farrell ME, Chen X, Rundle MM, Chan MY, Wig GS, Park DC. Regional amyloid accumulation and cognitive decline in initially amyloid-negative adults. *Neurology.* 2018;91(19):e1809–e21. <https://doi.org/10.1212/wnl.0000000000000649>.
27. Slot RER, Verfaillie SCJ, Overbeek JM, Timmers T, Wesselman LMP, Teunissen CE, et al. Subjective Cognitive Impairment Cohort (SCIENCE): study design and first results. *Alzheimers Res Ther.* 2018;10(1):76. <https://doi.org/10.1186/s13195-018-0390-y>.
28. van der Flier WM, Scheltens P. Amsterdam dementia cohort: performing research to optimize care. *J Alzheimers Dis.* 2018;62(3):1091–111. <https://doi.org/10.3233/jad-170850>.
29. van der Flier WM, Pijnenburg YA, Prins N, Lemstra AW, Bouwman FH, Teunissen CE, et al. Optimizing patient care and research: the Amsterdam Dementia Cohort. *J Alzheimers Dis.* 2014;41(1):313–27. <https://doi.org/10.3233/jad-132306>.
30. Hammers A, Allom R, Koeppe MJ, Free SL, Myers R, Lemieux L, et al. Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Hum Brain Mapp.* 2003;19(4):224–47. <https://doi.org/10.1002/hbm.10123>.
31. Gunn RN, Lammertsma AA, Hume SP, Cunningham VJ. Parametric imaging of ligand-receptor binding in PET using a simplified reference region model. *NeuroImage.* 1997;6(4):279–87. <https://doi.org/10.1006/nimg.1997.0303>.

32. Highlights of prescribing information: amyvid (florbetapir F18 injection). Food and Drug Administration. https://www.accessdata.fda.gov/drugsatfda_docs/label/2012/202008s000lbl.pdf.
33. Verhage F, Van Der Werff JJ. An analysis of variance based on the Groninger Intelligence Test Scores. *Ned Tijdschr Psychol.* 1964;19:497–509.
34. Hassenstab J, Ruvolo D, Jasielc M, Xiong C, Grant E, Morris JC. Absence of practice effects in preclinical Alzheimer's disease. *Neuropsychology.* 2015;29(6):940–8. <https://doi.org/10.1037/neu0000208>.
35. Verfaillie SCJ, Pichet Binette A, Vachon-Preseu E, Tabrizi S, Savard M, Bellec P, et al. Subjective cognitive decline is associated with altered default mode network connectivity in individuals with a family history of Alzheimer's disease. *Biol Psychiatry Cogn Neurosci Neuroimaging.* 2018;3(5):463–72. <https://doi.org/10.1016/j.bpsc.2017.11.012>.
36. Duff K, Foster NL, Hoffman JM. Practice effects and amyloid deposition: preliminary data on a method for enriching samples in clinical trials. *Alzheimer Dis Assoc Disord.* 2014;28(3):247–52. <https://doi.org/10.1097/WAD.0000000000000021>.
37. Jutten RJ, Grandoit E, Foldi NS, Sikkes SAM, Jones RN, Choi SE, et al. Lower practice effects as a marker of cognitive performance and dementia risk: a literature review. *Alzheimers Dement.* 2020;12(1):e12055. <https://doi.org/10.1002/dad2.12055>.
38. Leal SL, Lockhart SN, Maass A, Bell RK, Jagust WJ. Subthreshold amyloid predicts tau deposition in aging. *J Neurosci.* 2018;38(19):4482–9. <https://doi.org/10.1523/jneurosci.0485-18.2018>.
39. Palmqvist S, Scholl M, Strandberg O, Mattsson N, Stomrud E, Zetterberg H, et al. Earliest accumulation of beta-amyloid occurs within the default-mode network and concurrently affects brain connectivity. *Nat Commun.* 2017;8(1):1214. <https://doi.org/10.1038/s41467-017-01150-x>.
40. McRae-McKee K, Udeh-Momoh CT, Price G, Bajaj S, de Jager CA, Scott D, et al. Perspective: clinical relevance of the dichotomous classification of Alzheimer's disease biomarkers: should there be a "gray zone"? *Alzheimers Dement.* 2019;15(10):1348–56. <https://doi.org/10.1016/j.jalz.2019.07.010>.
41. Insel PS, Ossenkoppele R, Gessert D, Jagust W, Landau S, Hansson O, et al. Time to amyloid positivity and preclinical changes in brain metabolism, atrophy, and cognition: evidence for emerging amyloid pathology in Alzheimer's disease. *Front Neurosci.* 2017;11:281. <https://doi.org/10.3389/fnins.2017.00281>.
42. Fakhry-Darian D, Patel NH, Khan S, Barwick T, Svensson W, Khan S, et al. Optimisation and usefulness of quantitative analysis of (18)F-florbetapir PET. *Br J Radiol.* 2019;92(1101):20181020. <https://doi.org/10.1259/bjr.20181020>.
43. Habert MO, Bertin H, Labit M, Diallo M, Marie S, Martineau K, et al. Evaluation of amyloid status in a cohort of elderly individuals with memory complaints: validation of the method of quantification and determination of positivity thresholds. *Ann Nucl Med.* 2018;32(2):75–86. <https://doi.org/10.1007/s12149-017-1221-0>.
44. Carbonell F, Zijdenbos AP, Charil A, Grand'Maison M, Bedell BJ. Optimal target region for subject classification on the basis of amyloid PET images. *J Nucl Med.* 2015;56(9):1351–8. <https://doi.org/10.2967/jnumed.115.158774>.
45. Sakr FA, Grothe MJ, Cavado E, Jelistratova I, Habert M-O, Dyrba M, et al. Applicability of in vivo staging of regional amyloid burden in a cognitively normal cohort with subjective memory complaints: the INSIGHT-preAD study. *Alzheimers Res Ther.* 2019;11(1):15. <https://doi.org/10.1186/s13195-019-0466-3>.
46. Thomas BA, Erlandsson K, Modat M, Thurfjell L, Vandenberghe R, Ourselin S, et al. The importance of appropriate partial volume correction for PET quantification in Alzheimer's disease. *Eur J Nucl Med Mol Imaging.* 2011;38(6):1104–19. <https://doi.org/10.1007/s00259-011-1745-9>.
47. Schmidt ME, Chiao P, Klein G, Matthews D, Thurfjell L, Cole PE, et al. The influence of biological and technical factors on quantitative analysis of amyloid PET: points to consider and recommendations for controlling variability in longitudinal data. *Alzheimers Dement.* 2015;11(9):1050–68. <https://doi.org/10.1016/j.jalz.2014.09.004>.
48. Mormino EC, Kluth JT, Madison CM, Rabinovici GD, Baker SL, Miller BL, et al. Episodic memory loss is related to hippocampal-mediated beta-amyloid deposition in elderly subjects. *Brain.* 2009;132(Pt 5):1310–23. <https://doi.org/10.1093/brain/awn320>.
49. Jagust WJ, Landau SM, Shaw LM, Trojanowski JQ, Koeppe RA, Reiman EM, et al. Relationships between biomarkers in aging and dementia. *Neurology.* 2009;73(15):1193–9. <https://doi.org/10.1212/WNL.0b013e3181bc010c>.
50. Duara R, Loewenstein DA, Lizarraga G, Adjouadi M, Barker WW, Greig-Custo MT, et al. Effect of age, ethnicity, sex, cognitive status and APOE genotype on amyloid load and the threshold for amyloid positivity. *Neuroimage Clin.* 2019;22:101800. <https://doi.org/10.1016/j.nicl.2019.101800>.
51. Schreiber S, Landau SM, Fero A, Schreiber F, Jagust WJ. Initiative ftAsDN. Comparison of visual and quantitative florbetapir F 18 positron emission tomography analysis in predicting mild cognitive impairment outcomes. *JAMA Neurol.* 2015;72(10):1183–90. <https://doi.org/10.1001/jamaneurol.2015.1633>.
52. Ossenkoppele R, Jansen WJ, Rabinovici GD, Knol DL, van der Flier WM, van Berckel BN, et al. Prevalence of amyloid PET positivity in dementia syndromes: a meta-analysis. *Jama.* 2015;313(19):1939–49. <https://doi.org/10.1001/jama.2015.4669>.
53. Klunk WE, Koeppe RA, Price JC, Benzinger TL, Devous MD Sr, Jagust WJ, et al. The Centiloid Project: standardizing quantitative amyloid plaque estimation by PET. *Alzheimers Dement.* 2015;11(1):1–15.e1–4. <https://doi.org/10.1016/j.jalz.2014.07.003>.
54. Navitsky M, Joshi AD, Kennedy I, Klunk WE, Rowe CC, Wong DF, et al. Standardization of amyloid quantitation with florbetapir standardized uptake value ratios to the Centiloid scale. *Alzheimers Dement.* 2018;14(12):1565–71. <https://doi.org/10.1016/j.jalz.2018.06.1353>.
55. Hanseeuw BJ, Malotau V, Dricot L, Quenon L, Sznajer Y, Cerman J, et al. Defining a Centiloid scale threshold predicting long-term progression to dementia in patients attending the memory clinic: an [18F] flutemetamol amyloid PET study. *Eur J Nucl Med Mol Imaging.* 2020. <https://doi.org/10.1007/s00259-020-04942-4>.

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