Title

Existing Pittsburgh Compound-B positron emission tomography thresholds are too high: statistical and pathological evaluation.

Journal

Brain : a journal of neurology, 138(Pt 7)

ISSN

0006-8950

Authors

Villeneuve, Sylvia
Rabinovici, Gil D
Cohn-Sheehy, Brendan I
et al.

Publication Date

2015-07-01

DOI

10.1093/brain/awv112

Peer reviewed

OXFORD UNIVERSITY PRESS | Brain

Existing Pittsburgh Compound B - Positron Emission Tomography Thresholds are Too High: Statistical and Pathological Validation

Journal:	*Brain*
Manuscript ID:	BRAIN-2014-02124.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	13-Feb-2015

Complete List of Authors:	Villeneuve, Sylvia; University of California Berkeley, Helen Wills Neuroscience Institute Rabinovici, Gil; UCSF, Neurology; Helen Wills Neuroscience Institute, UC Berkeley, Cohn-Sheehy, Brendan; Memory & Aging Center, University of California, San Francisco, Department of Neurology; Helen Wills Neuroscience Institute, UC Berkeley, Madison, Cindee; Helen Wills Neuroscience Institute, University of California Berkeley, Berkeley, CA, Ayakta, Nagehan; Helen Wills Neuroscience Institute, University of California Berkeley, Berkeley, CA, ; UCSF Memory & Aging Center, Department of Neurology Ghosh, Pia; Helen Wills Neuroscience Institute, UC Berkeley, ; UCSF Memory & Aging Center, Department of Neurology La Joie, Renaud; Inserm U1077, ; University of California, Berkeley, Helen Wills Neuroscience Institute Arthur-Bentil, Samia; Helen Wills Neuroscience Institute, University of California Berkeley, Berkeley, CA, Vogel, Jacob; University of California Berkeley, Helen Wills Neuroscience Institute Marks, Shawn; Helen Wills Neuroscience Institute, Lehmann, Manja; UCL, Dementia Research Centre; UCSF, Neurology Rosen, Howard; University of California at San Francisco, Neurology Reed, Bruce; Univeristy of California, Davis, Neurology; Olichney, John; UC Davis, Boxer, Adam; UCSF, Miller, Bruce; UCSF, Neurology; UCSF, Neurology Borys, Ewa; Stritch School of Medicine, Loyola University, Pathology Jin, Lee-Way; Alzheimer's Disease Center, University of California at Davis, Pathology Huang, Eric; UCSF, Pathology Lea, Grinberg; University of California, San Francisco, CA 941143, USA, Neurology DeCarli, Charles; Univeristy of California, Davis, Neurology Seeley, William; UCSF Memory & Aging Center, Department of Neurology Jagust, William; University of California Berkeley,

Subject category:	Dementia
To search keyword list, use whole or part words followed by an *:	Amyloid imaging < DEMENTIA, Alzheimer's disease < DEMENTIA, Dementia: biomarkers < DEMENTIA, Neuropathology < DEMENTIA, Neurodegeneration: biomarkers < NEURODEGENERATION: CELLULAR AND MOLECULAR, beta-Amyloid < NEURODEGENERATION: CELLULAR AND MOLECULAR

Existing Pittsburgh Compound B - Positron Emission Tomography Thresholds are Too High: Statistical and Pathological Validation

Sylvia Villeneuve, [1*] Gil D. Rabinovici,[1-3*] Brendan I. Cohn-Sheehy, [1-3] Cindee Madison, [1] Nagehan Ayakta, [1,3] Pia M. Ghosh,[1-3] Renaud La Joie, [1] Samia Kate Arthur-Bentil, [1,2] Jacob W. Vogel, [1] Shawn M. Marks, [1] Manja Lehmann,[1,3] Howard. J. Rosen,[3] Bruce Reed,[4,5] John Olichney,[4] Adam L. Boxer,[3] Bruce L. Miller,[3] Ewa Borys,[6] Lee-Way Jin,[7] Eric J. Huang,[8] Lea T. Grinberg,[3,8] Charles DeCarli,[4] William W. Seeley,[3,8] and William Jagust[1,2]

[1] Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley CA, USA

[2] Lawrence Berkeley National Laboratory, Berkeley, California, USA

[3] Memory and Aging Center, Department of Neurology, University of California, San Francisco

[4] Department of Neurology, School of Medicine, University of California, Davis

[5] Veterans Affairs Northern California Health Care System, Martinez, CA

[6] Dept. of Pathology, Stritch School of Medicine, Loyola University Chicago

[7] Dept. of Pathology and Laboratory Medicine, University of California, Davis

[8] Department of Pathology, University of California, San Francisco

* These authors contributed equally to this work

Correspondence to:
Sylvia Villeneuve

Helen Wills Neuroscience Institute
132 Barker Hall, MC# 3190
University of California, Berkeley, CA 94720 USA
E-mail: villeneuve.sylvia@gmail.com
Telephone: 510-643-6616
Fax: 510-642-3192

Character count title: 111; word count abstract: 384; word count text: 5967; number of tables: 3; number of figure: 5, number of video:1.

SUMMARY

β-Amyloid (Aβ), a hallmark of Alzheimer's disease, begins accumulating up to two decades before the onset of dementia, and can be detected in vivo applying Aβ PET tracers such as carbon-11 labelled Pittsburgh Compound-B (PIB). A variety of thresholds have been applied in the literature to define PIB-PET positivity, but the ability of these thresholds to detect early Aβ deposition is unknown, and validation studies comparing PIB thresholds to post-mortem amyloid burden are lacking. In this study we first derived thresholds for amyloid PET positivity using PIB-PET in 154 cognitively normal older adults with four complementary approaches: (1) reference values from a young control group aged between 20 and 30 years, (2) a Gaussian mixture model that assigned each subject a probability of being Aβ-positive or Aβ-negative based on PIB index uptake, (3) a k-means cluster approach that clustered subjects into Aβ-positive or Aβ-negative based on PIB uptake in different brain regions (features), and (4) an iterative voxel-based analysis that further explored the spatial pattern of early Aβ PET signal. Next, we tested the sensitivity and specificity of the derived thresholds in 50 individuals who underwent PIB-PET during life and brain autopsy (mean time PET to autopsy 3.1 ± 1.8 years). Amyloid at autopsy was classified using Consortium to Establish a Registry for Alzheimer's Disease (CERAD) criteria, unadjusted for age. The analytic approaches yielded low thresholds ($SUVR_{low}$ =1.21, DVR_{low} =1.08) that represent the earliest detectable PIB signal, as well as high thresholds ($SUVR_{high}$ =1.40, DVR_{high} =1.20) that are more conservative in defining PIB-PET positivity. In voxel-wise contrasts, elevated PIB retention was first noted in the medial frontal cortex, then the precuneus, lateral frontal and parietal lobes, and finally the lateral temporal lobe. When compared to post-mortem amyloid burden, low proposed thresholds were more sensitive than high thresholds (sensitivities:

DVR_{low} 81.0%, $SUVR_{low}$ 83.3%; DVR_{high} 61.9%, $SUVR_{high}$ 62.5%) for CERAD moderate-to-frequent neuritic plaques, with similar specificity (DVR_{low} 95.8%; $SUVR_{low}$, DVR_{high} and $SUVR_{high}$ 100.0%). A receiver operator characteristic analysis identified optimal DVR (1.06) and SUVR (1.20) thresholds that were nearly identical to the a priori DVR_{low} and $SUVR_{low}$. In summary, we found that frequently applied thresholds for PIB-positivity (typically at or above DVR_{high} and $SUVR_{high}$) are overly stringent in defining amyloid positivity. Lower thresholds in this study resulted in higher sensitivity while not compromising specificity.

INTRODUCTION

PET amyloid imaging has had a profound effect on aging and dementia research. The first publication of a β-amyloid-selective imaging agent, carbon-11 labelled Pittsburgh Compound-B [^{11}C]PIB (Klunk *et al.*, 2004) opened the door to *in vivo* detection of a core aspect of Alzheimer's disease pathology. Soon afterwards, [^{18}F]-labeled amyloid imaging agents were developed and commercialized, widely increasing the availability of this technology. The ability to detect and quantify fibrillar brain amyloid-β (Aβ) *in vivo* has helped to establish models of disease pathophysiology and biomarker progression (Jack *et al.*, 2010, Bateman *et al.*, 2012, Jack *et al.*, 2013) and guide the design of clinical trials (Salloway *et al.*, 2014). These studies offer the potential for a more complete understanding of the pathophysiology of Alzheimer's disease, along with the hope of early therapeutic intervention in people who harbor amyloid pathology but do not yet express cognitive decline.

In studies of amyloid imaging, many investigators use an overall measure of radioligand retention in the brain in order to dichotomize subjects into "amyloid positive" and "amyloid negative" categories. However, Aβ deposition occurs on a continuum; at present there is no clear *a priori* way to separate individuals who have pathologically relevant Aβ deposition from those who do not. Nevertheless, there are important reasons to consider categorical classification of individual subjects. Classification of individuals as amyloid "positive" or "negative" is relevant for clinical diagnosis, for inclusion of subjects in anti-amyloid therapeutic trials, and for distinguishing Aβ –dependent and Aβ –independent changes in cognition and in brain structure and function. Measurements of 1.4 - 1.5 standardized uptake value ratio (SUVR) units have often been traditionally used in the literature to identify Aβ-

positive subjects using PET scanning with PIB. These thresholds are based on different categorization approaches such as the natural data breakpoints, the upper confidence limit observed in cognitively normal older adults, the lower confidence limits found in patients with clinical Alzheimer's disease dementia, iterative outlier removal, hierarchical clustering or Gaussian mixture modeling (GMM) (Pike *et al.*, 2007, Aizenstein *et al.*, 2008, Jack *et al.*, 2008, Hedden *et al.*, 2009, Rowe *et al.*, 2010, Villemagne *et al.*, 2011, Jack *et al.*, 2012, Nordberg *et al.*, 2013, Mormino *et al.*, 2014). It is unknown if these thresholds truly allow the earliest possible detection point of pathologically relevant Aβ-PET signal. Identifying subjects with amyloid deposition as early as possible is important to truly understand the Alzheimer's disease pathophysiological cascade, and to exclude early amyloid accumulators from studies focusing on normal cognitive aging and "suspected non-Alzheimer's disease pathology" (SNAP)(Jack *et al.*, 2012).

The main aim of this study was to identify and validate thresholds that detect pathologically relevant PIB signal as early as possible. A secondary aim was to examine the spatial pattern of early Aβ PET signal. As a first step, we applied four distinct statistical approaches to define a low threshold for Aβ PET positivity based on [^{11}C]PIB data from 154 older adults. Thresholds were defined based on (1) reference values from a young control group, (2) a GMM, (3) a k-means cluster approach, and (4) an iterative voxel-based analysis. The GMM and the cluster analyses also allowed derivation of a higher threshold that might be favored when reducing the rate of false positives is more important than detecting early PIB-PET signal. The cluster and the voxel-wise analyses further allowed examination of the spatial pattern of early Aβ PET signal. Because Aβ is hypothesized to start

accumulating long before cognitive impairment is clinically evident, we derived the thresholds based on data acquired in cognitively normal older adults. Thresholds were derived for the two most common methods of PIB-PET quantification: SUVR units, and distribution volume ratios (DVR).

Evaluating sensitivity and specificity of amyloid PET thresholds requires a "standard of truth" which, for the detection of brain Aβ, necessitates pathological examination of the brain (Clark *et al.*, 2011). As a second step we therefore applied the derived low and high SUVR and DVR thresholds to PIB-PET scans from 50 individuals enrolled in longitudinal studies of aging and dementia who underwent amyloid imaging and were also followed to autopsy. Classifications based on the proposed thresholds were compared to the burden of amyloid at autopsy as measured by the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) scale (Mirra *et al.*, 1991). CERAD is a semi-quantitative scale of Aβ neuritic plaque (NPs), fibrillar amyloid aggregates considered to be the primary Aβ species that bind PIB *in vivo* (Ikonomovic *et al.*, 2008, Ikonomovic *et al.*, 2012, Ni *et al.*, 2013).

We hypothesized that the proposed low thresholds would have higher sensitivity for significant Aβ burden (defined as CERAD moderate-to-frequent NPs), whereas the proposed high thresholds would have higher specificity and result in fewer false positives in brains with CERAD absent to sparse NPs. We also hypothesized that cortical hubs of brain connectivity would be key regions of early PiB signal, supporting the idea that regions of high connectivity are prone to earlier Aβ deposition (Buckner *et al.*, 2009).

MATERIALS AND METHODS

Threshold derivation

Participants

Table 1 shows the demographic data of 154 cognitively normal elderly and 18 young adults included in the threshold derivation step. Older adults were from the Berkley Aging Cohort (BAC) and from ongoing studies at the University of California Davis (UCD) Alzheimer's Disease Center (see supplementary material for more details). Written informed consent was obtained from participants under protocols approved by the Institutional Review Boards of all participating institutions.

MR imaging and preprocessing

Structural T1-weighted MR images were obtained on different scanners (supplementary material). All MPRAGE scans were processed with FreeSurfer version 5.1 (http://surfer.nmr.mgh. harvard.edu/) to derive regions of interest (ROIs) in each subject's native space using the Desikan-Killiany atlas (Desikan *et al.*, 2006). These ROIs where then used to extract regional cortical PIB values.

PET imaging and preprocessing

All subjects underwent PIB-PET imaging at Lawrence Berkeley National Laboratory (supplementary material). PIB-PET data were preprocessed using the Statistical Parametric Mapping software package (SPM8; http://www.fil.ion. ucl.ac.uk/spm) using a previously published protocol (Villeneuve *et al.*, 2014). All subjects included in the threshold derivation received dynamic scans. DVRs were generated with Logan graphical analysis, PIB frames corresponding to 35–90 minutes post-injection and a native-space grey matter cerebellar mask as the reference region (Logan *et al.*, 1996,

Price *et al.*, 2005). SUVRs were calculated by dividing the mean uptake 50-70 minutes post-injection by the grey matter cerebellar mask.

For each subject, both a DVR and a SUVR "PIB index" were derived from the native-space image by averaging the weighted mean value from Freesurfer-derived ROIs in frontal, temporal, parietal and posterior cingulate cortex using the Desikan-Killiany atlas (Desikan *et al.*, 2006) (supplementary Fig. 1A). The PIB index thus includes cortical regions that show a high proclivity for PiB retention in AD and normal aging (Price *et al.*, 2005, Rabinovici *et al.*, 2010). ROI –specific values were also extracted from 76 ROIs from the same atlas (supplementary Table 1). Because of the linear correspondence found between the DVR and the SUVR values, all statistical analyses for the threshold derivation part of the manuscript were performed using DVRs and a regression line (SUVR = -0.54 + 1.62*DVR, R = 0.97) was applied to derive the SUVR cutoffs corresponding to the DVR thresholds.

Statistical Analysis

We investigated optimal PIB Index cutoffs to detect early PIB-PET signal using 4 different methods: (1) a reference group of young adults, (2) GMM analysis, (3) cluster analysis and (4) voxel-wise analysis.

Young adults analysis

There is strong evidence that adults under age 30 years are almost invariably free from Aβ accumulation (Kok *et al.*, 2009, Fleisher *et al.*, 2012). Thus, the first threshold we used in this study was defined as 2 standard deviations (SD) above the group of 18 young subjects aged between 20 and 30 years. This approach was used in

a previous publication from our group (Mormino *et al.*, 2012). In that previous publication, which included 11 of the 18 current young subjects, the cutoff was set at DVR = 1.08.

Gaussian mixture models (GMM) analysis

A GMM is a probabilistic model assuming that the overall data distribution can be estimated from a mixture of Gaussian distributions. Using that technique, we fit from 1 to 11 Gaussian distributions to our data and used a Bayesian information criterion (BIC) to assess the optimal number of Gaussian distributions represented in our data. We found that the best fit for our data was 2 Gaussian distributions, which is consistent with a previous report using different data sets (Mormino *et al.*, 2014). Then, each subject was assigned a probability of belonging to each Gaussian distribution. The two cutoffs derived using this technique represent the 90% probability of belonging in the low (representing the PIB-negative subjects) or the high (representing the PIB-positive subjects) distributions. While admittedly arbitrary, we chose a probability of 90% based on thresholds applied to define abnormal scan results in the literature (Jack *et al.*, 2012). Notably, results were highly similar while using a 95% probability (supplementary Table 3).

k-means Cluster analysis

This technique is similar to (and not totally independent from) the GMM analyses in the sense that it defines how we can cluster, or group, the data together. In this analysis, instead of examining the DVR PIB index we used the PIB DVR values extracted from the 76 ROIs defined by the Desikan-Killiany atlas. We restricted the analysis to two clusters, one that represents subjects with high Aβ deposition and the

other representing subjects with low Aβ deposition. The two cutoffs derived using this technique represent the 90th percentile of the low cluster (representing the PIB-negative subjects) or the 10th percentile of the high cluster (representing the PIB-positive subjects).

Voxel-wise analysis

In this analysis, we began by ranking all older normal subjects by their DVR PIB index. We then created a reference group of 22 subjects with a mean index of 1.00 to which we compared a series of subsequent groupings of the remaining participants (referred to here as the group of interest). In other words, after selecting the 22 subjects that comprised the reference group, we took the 22 subjects with the next highest PIB index values and performed voxel-wise contrasts between this group of interest and the reference group. We then dropped the subject from the group of interest with the lowest PIB index and added the subject with the next highest value and again performed a voxel-wise comparison between that group of interest of 22 subjects to the same reference group. We continued this process, iteratively creating groups of interest of 22 subjects by dropping the individual with the lowest PIB index and adding the one with the next highest index, such that the "new" group of interest differed from the previous group by only two subjects. In this way, groups of interest gradually moved up the scale of PIB index DVRs, always using the same initial reference group of 22 subjects (mean DVR=1.00) for comparison. This procedure was repeated until the subject with the highest DVR was included. A DVR of 1.00 was chosen for the reference group since this DVR value reflects a level of PIB retention in the cortex equivalent to the cerebellum grey matter, and thus indicates no specific tracer retention.

Voxel-wise analyses were performed using SPM8 software. First, the DVR scans were warped to the MDT2 template (Sun *et al.*, 2007). Then, images were smoothed (Gaussian kernel of 10 x 10 x 10 mm), masked to exclude non-grey matter voxels from the analyses, and two-sample t-tests were performed. All voxel-wise analyses were corrected for multiple comparisons using a Family-Wise Error (FWE) at $p < .05$ and cluster size $k \geq 150$.

The DVR threshold was defined as the mean PIB index DVR of the group of interest when a statistically significant signal of elevated PIB retention was first detected. Overall, 66 iterations were used to examine patterns of Aβ accumulation. In this model, we assume that increasing DVR represents temporal progression, an assumption that we recognize may not be entirely justified.

Threshold validation

Participants

Table 1 shows the demographic data of the 50 individuals who had both PIB-PET and autopsy. Participants were from the University of California, San Francisco Memory and Aging Center or the UCD Alzheimer's Disease Center (supplementary material). All but two of these participants were cognitively impaired and therefore not included in the threshold derivation step. Clinical diagnosis was established at a multi-disciplinary conference applying standard research criteria for mild cognitive impairment (MCI) and dementia syndromes (McKhann *et al.*, 1984, Roman *et al.*, 1993, Petersen, 2004, McKeith *et al.*, 2005, Albert *et al.*, 2011, Gorno-Tempini *et al.*,

2011, McKhann *et al.*, 2011, Rascovsky *et al.*, 2011, Armstrong *et al.*, 2013). This report reflects data on 50 individuals who had both PIB-PET and autopsy as of May, 2014.

MR imaging and preprocessing

Structural T1-weighted MR images were obtained on different scanners (supplementary material). T1-weighted images were used only for definition of the cerebellum reference region, using FreeSurfer v5.1 software, and for spatial normalization.

PET imaging and processing

All subjects underwent PET imaging on the same scanners with the same acquisition parameters as for the threshold derivation study. 45 of these subjects underwent dynamic PET imaging. The cerebellar grey normalized DVR image was then warped to an anatomical T1-based template in MNI space using the subjects' T1 MRIs as reference images. Grey matter segmentation was defined for each subject in template space, applying a probabilistic grey matter mask from the Montreal Neurological Institute (MNI) T1 template using Statistical Parametric Mapping (SPM8). In 5 subjects, dynamic data were not obtained, but PIB data were collected from 50-70 min following tracer injection for calculation of SUVRs (in one subject, only data from 55-70 min were available). SUVR and DVR images were subsequently processed identically. All analyses for the threshold validation part of the manuscript were preformed using both the DVR and the SUVR values.

Similar to the threshold derivation study, we estimated mean cortical PIB retention using a "PIB index". However, due to a high failure rate of FreeSurfer-based segmentation in MRIs derived from dementia subjects, the PIB index in this group was created by combining frontal, temporal, parietal and posterior cingulate regions defined within the Automated Anatomical Labeling (AAL) atlas (Tzourio-Mazoyer *et al.*, 2002) and extracting values in template space. The AAL-defined regions were highly analogous to those defined using the Desikan atlas (supplementary Fig. 1B), and the two methods yield highly correlated mean DVR values (supplementary materials).

Neuropathological Examination

Brain autopsies were performed at University of California, San Francisco Memory and Aging Center (N = 38), University of California Davis (9), University of Pennsylvania (1), University of California Los Angeles (1) and Mayo Clinic Jacksonville (1). Pathological assessments were performed using institution-specific protocols, as previously described (Chui *et al.*, 2006, Forman *et al.*, 2006, Grinberg *et al.*, 2013, Magaki *et al.*, 2014, Murray *et al.*, 2014). All autopsies included tissue sampling in regions relevant to the differential diagnosis of dementia based on published consensus criteria (Roman *et al.*, 1993, Newell *et al.*, 1999, McKeith *et al.*, 2005, Mackenzie *et al.*, 2010, Hyman *et al.*, 2012). Tissue staining included hematoxylin/eosin and at least one of the following stains: thioflavin S, modified Bielschowsky, Gallyas silver stain, or immunohistochemistry for Aβ (the latter was available in 41/50 cases and is considered equivalent to older stains for the purposes of CERAD staging in updated pathological criteria (Hyman *et al.*, 2012). NP densities were based on the assessment of sections stained with these methods. CERAD scores

were based on the highest density of NPs found at autopsy. Sections were rated unadjusted for age, as follows: 1-5 NPs in a 100x field were classified as CERAD-sparse, 6-14 as CERAD-moderate and ≥15 as CERAD-frequent (Montine *et al.*, 2012). Immunohistochemistry for hyperphosphorylated tau, α-synuclein, ubiquitin, and transactive response DNA binding protein 43 (TDP-43) was performed based on institutional protocols. Alzheimer's disease neuropathology was further characterized by Braak stage (Braak and Braak, 1998) and the National Institute on Aging – Reagan criteria (1997, Hyman and Trojanowski, 1997). In some cases neuropathologists had access to the clinical histories and thus may not have been blinded to PIB-PET results.

Autopsy reports were reviewed by an experienced neurologist (GDR) who extracted the primary and contributing neuropathological diagnoses, as well as CERAD, Braak and NIA-Reagan scores for each case. The presence and degree of cortical diffuse plaques (DPs) and cerebral amyloid angiopathy (CAA) in brain parenchyma (i.e. not isolated to leptomeninges) was also recorded. Staging applying the updated NIA-AA pathological criteria for Alzheimer's disease (Hyman *et al.*, 2012) were only available for a subset of participants since many of the autopsies preceded publication of these new criteria.

Estimation of Sensitivity and Specificity

We assessed the relationship between classification of subjects as PIB-PET positive or negative applying the derived low and high thresholds, and compared PET categorization to the classification of subjects by autopsy as positive or negative for significant NPs. To conform to previous PET-pathology studies (Clark *et al.*, 2012, Curtis *et al.*, 2015), cases with CERAD scores of "absent" or "sparse" were

categorized as pathologically negative for significant NPs, while "moderate" and "frequent" CERAD scores were considered positive. Sensitivity and specificity for each of these thresholds were estimated by the appropriate observed proportion, and 95% confidence intervals were generated based on the assumption that they follow a binomial distribution. In order to explore whether the *a priori* thresholds we selected were truly optimal, in a separate exploratory analysis we used receiver-operating characteristic (ROC) curves to determine the PIB thresholds that maximized overall classification accuracy of CERAD moderate-frequent cases versus CERAD absent-sparse cases.

Additional Statistical Analyses

For both the threshold derivation and validation steps, group differences in continuous variables were examined using Student's t-tests. Group differences in dichotomous variables were compared using chi-square or Fisher's exact tests. Statistical analyses were implemented in PASW 21.0 (SPSS Inc.).

RESULTS

Threshold derivation

Young adults

The mean DVR of the young subjects was 1.01 (SD = .03), leading to a cutoff of 1.07 that is 2 standard deviations above the mean. Fitting a regression line between DVRs and SUVRs (SUVR = -0.54 + 1.62*DVR, R = 0.97) showed that a DVR cutoff of 1.07 is equivalent to an SUVR cutoff of 1.19.

Table 2 shows the cutoffs for the four analytic techniques applied in this study.

Gaussian mixture models

The two Gaussian distributions are presented in Fig. 1. Based on the model, each subject was assigned a probability of belonging to either distribution. Two cutoffs were derived: the lower threshold was DVR = 1.09 and corresponds to a 90% probability of belonging in the low Aβ distribution; and the higher threshold was DVR = 1.21 and corresponds to a 90% probability of belonging to the high Aβ distribution. Applying the DVR vs. SUVR regression line identified a low SUVR cutoff of 1.23 and a high SUVR cutoff of 1.42.

Cluster analysis

The two clusters are presented in Fig. 2A, representing subjects with high and low Aβ. Each row of each cluster represents one subject and each column represents one significant feature (ROI), for a total of 13 features. Similar to the GMM approach, each subject was assigned a probability of belonging to cluster A (low Aβ) or cluster B (high Aβ). Two cutoffs were derived: the lower threshold was DVR=1.09 and corresponds to the 90th percentile of cluster A (low Aβ); and the higher threshold was DVR = 1.19 and corresponds to the 10th percentile of Cluster B (high Aβ). Fitting a regression line between the DVR and the SUVR data suggest a low SUVR cutoff of 1.23 and a high SUVR cutoff of 1.39.

The 13 features represent the brain regions that best (or most significantly) discriminate two clusters among the 76 ROIs tested in the model are shown in Fig. 2B. The anterior cingulate (left and right) and the precuneus (left and right) stood out as key regions that differentiate the two clusters. The ROIs representing the 13

significant features, that are mainly symmetrical, as well as their weights in the model are shown in supplementary Table 2.

Voxel-wise analysis

The first significant difference between the reference group (n = 22, mean DVR = 1.00, SD = 0.01) and the groups of interest (n = 22) was found when the group of interest had a mean PIB index DVR of 1.07 (SD = 0.01) (Fig. 3, Video). In this contrast, a cluster of significant PIB retention was limited to the medial frontal cortex; at very slightly higher levels of mean PIB index DVRs (1.08), significant clusters were found in the precuneus, followed by the lateral frontal and parietal lobes which appeared when the mean PIB index DVR was near 1.14 (SUVR 1.31), and finally the lateral temporal lobe (Fig. 3, Video). The DVR of 1.07 defining the first significant cluster of voxels corresponds to an SUVR of 1.19.

Threshold validation

In the separate autopsy validation cohort, the mean interval between PET scanning and death was 3.1 years (range: 0.2 – 6.4 years). At autopsy, the most common neuropathological diagnoses were frontotemporal lobar degeneration, Alzheimer's disease (often with mixed pathology) and cerebrovascular disease (Table 1). The distribution of CERAD scores was relatively bimodal, with most subjects classified as CERAD absent or CERAD frequent. Only 4 subjects fell in the CERAD moderate category.

The overall prevalence of Aβ deposits (in the form of NPs, DPs or CAA) at autopsy was 84%, including 64% with NPs, 80% with DPs (data missing in 4 cases) and 50%

with CAA (data missing in 6 cases). Some degree of amyloid pathology was present at autopsy in 100% of ApoE4 carriers versus 77% of non-carriers (p=0.09).

Fig. 4 shows examples of PIB-PET scans in subjects representing the spectrum of post-mortem amyloid burden. PIB-PET in the subject with CERAD sparse plaques had DVR and SUVR values well below the low thresholds (Fig. 4, top row). PIB-PET in the subject with CERAD moderate plaques had DVR and SUVR values positive by the low thresholds and negative by the high thresholds (middle row). This subject has relatively focal retention of PIB in the left temporal and parietal lobes. PIB-PET in the subject with CERAD frequent NPs was well above the high DVR and SUVR thresholds (bottom row).

Sensitivity and specificity for low and high DVR and SUVR thresholds are shown in Table 3. While DVR_{low} and $SUVR_{low}$ showed reasonable sensitivity for CERAD moderate-to-frequent NPs (81% - 83%), DVR_{high} and $SUVR_{high}$ thresholds showed surprisingly low sensitivity (~62%). Conversely, all thresholds showed very high specificity. Even when applying liberal thresholds, there was only one false positive result by DVR_{low}, and none applying $SUVR_{low}$. Overall, DVR and SUVR values were highly correlated ($r=0.98$, $p<0.001$), and both measures showed comparable sensitivity and specificity at both low and high thresholds.

Fig. 5 shows the relationships between DVR/SUVR values and CERAD scores. For DVR_{high}, there were 8 false negatives (compared to 4 for DVR_{low}), while for $SUVR_{high}$ there were 9 false negatives (compared to 4 for $SUVR_{low}$). Similar plots comparing values to Thal amyloid stages and NIA-AA pathological criteria for AD in the subset

of patients in which these were available are shown in supplementary Fig. 4. While a direct comparison between the SUVR and DVR methods of PIB quantification is outside the scope of our study, it is worth noting that these methods yielded highly correlated results, and performed comparably in predicting post-mortem amyloid burden.

Supplementary Fig. 3 provides examples of misclassified cases. There was only one false positive case in our series. Visual inspection of the image suggests that the borderline positive DVR value may be due to relatively high PIB retention in white matter, which could contaminate grey matter PIB signal via partial volume effects. Scans were classified as false negatives (by one or both thresholds) for a variety of reasons including focal amyloid accumulation, cortical atrophy and CAA in the grey matter of the cerebellum (see supplementary material for more details). PET to pathology intervals were no different in patients that were "false negatives" by one or more standards versus other patients in the cohort.

The results of the ROC analysis are shown in supplementary Table 4. The empirically derived optimal DVR (DVR = 1.06, AUC 0.89, sensitivity = 85.7%, specificity = 95.8%) and SUVR (SUVR = 1.20, AUC 0.88, sensitivity = 83.3%, specificity = 100%) thresholds were nearly identical to the *a priori* DVR_{low} and $SUVR_{low}$ thresholds, thus yielding similar sensitivity and specificity values.

DISCUSSION

The selection of threshold values to determine amyloid positivity has important implications for studying mechanisms of Alzheimer's disease pathogenesis and for

diagnosis and selection of subjects for clinical trials. Although Aβ deposition occurs on a continuum, it is often necessary to dichotomize subjects as amyloid positive or negative. Using different data analysis approaches, we found strong evidence that an optimal threshold for early Aβ detection with PIB-PET should be set at a DVR index of 1.08 (or an SUVR of 1.21). Our threshold derivation approaches also provide evidence that a DVR cutoff of 1.20 (SUVR of 1.40) may be valuable when the priority is to minimize false positive subjects. The validity of the low cutoff to detect fibrillar Aβ plaque pathology was confirmed by an autopsy study of 50 individuals. In fact, the low DVR threshold of 1.08 was optimal at detecting moderate-to-frequent NPs (specificity = 95.5%; sensitivity = 81%) while the higher DVR threshold was surprisingly insensitive to this burden of amyloid pathology (sensitivity = 61.9%; specificity =100%). The results in the aggregate suggest that the currently used SUVR thresholds of 1.40 (and higher) and DVR thresholds of 1.20 (and higher) are insensitive and likely misclassify many individuals with substantial Aβ NP accumulation as amyloid negative.

The threshold derived from a group of young older adults highly likely to be free of Aβ led to a DVR threshold of 1.07 or an SUVR threshold of 1.19, similar to the threshold defined in a previous study from our group that included subset of the same subjects (Mormino *et al.*, 2012). The voxel-wise analysis approach also yielded a DVR threshold of 1.07 which represented the lowest mean DVR at which statistical PIB signal was detected. This seemingly low threshold value for DVR index has face validity because regional PIB retention in the cases that would be classified as low positives occurs in medial frontal and medial parietal cortex, brain regions known to show early Aβ deposition. Furthermore, both a GMM and a cluster analyses

suggested a similar DVR cutoff of 1.09 (SUVR = 1.23) to detect early PIB-PET signal. Based on these results, we propose that an optimal threshold for the early detection of Aβ with PIB-PET should be a DVR of 1.08, which represents the mean point between 1.07 and 1.09. Based on an independent longitudinal study using a similar processing method (Villemagne *et al.*, 2013), there would be an estimated 7 year time window between our value defining a lower threshold (DVR = 1.08, SUVR = 1.21) and the time when a person reaches an SUVR of 1.40, allowing for potential earlier therapeutic amyloid lowering interventions.

While a DVR cutoff of 1.08 (SUVR = 1.19) might be optimal to detect early PIB signal, it might not be ideal in all circumstances and should be used with caution. Indeed, a DVR cutoff of 1.08 will likely increase the number of false positive cases due to partial volume effects of white matter binding, or other measurement errors. Selection of thresholds for any given study needs to establish whether false positives or false negatives are more problematic. If a study, or a clinical trial, needs to emphasize specificity over sensitivity our GMM and cluster results analyses support the widespread practice of using a DVR of 1.20 (SUVR = 1.40). While this higher threshold might misclassify early accumulators, it will still capture subjects many years before Aβ burden approaches a plateau (Villemagne *et al.*, 2011, Jack *et al.*, 2013). Labeling individuals who fall below this high threshold as Aβ-negative is however problematic. In our cognitively normal cohort, 19% of individuals falling below a DVR of 1.20 were classified as positive using the lower threshold. In studies in which it is crucial to detect Aβ accumulation as early as possible (e.g. in studies making inferences about neurodegenerative processes that are independent of Aβ), low sensitivity can substantially bias conclusions when high thresholds are used.

Relatively few studies have compared PIB-PET results during life with post-mortem amyloid burden. In a small series of 6 individuals, 5 cognitively normal and 1 with dementia, mean cortical DVR values were not strongly associated with NP scores, although regional scores, especially in the precuneus, were notable for NP accumulation above DVR values of 1.20 (our DVR_{high}) (Sojkova *et al.*, 2011, Driscoll *et al.*, 2012). This sample was too small to estimate sensitivity and specificity. A number of larger studies have assessed thresholds for F18-labelled amyloid PET tracers in end-of-life populations. In an imaging-pathology series involving 59 subjects studied with the amyloid PET tracer [^{18}F]Florbetapir (Clark *et al.*, 2012), a pre-specified threshold SUVR of 1.10 (whole cerebellum reference) had a sensitivity of 97% and specificity of 100% for CERAD moderate-to-frequent NPs. This threshold is comparable to a PIB SUVR of 1.40 (cerebellar grey reference, or our $SUVR_{high}$) (Landau *et al.*, 2014). It is difficult to directly compare our results to those of [^{18}F]Florbetapir given the many differences not only in tracers but in methods of analysis and subject selection. Recent efforts to standardize and cross-validate amyloid PET values across tracers and analytic methodologies, including the proposed "Centiloid" scale (Klunk *et al.*, 2015), will help aggregate results from clinicopatahological studies of amyloid imaging in an effort to optimize early detection.

In our series, the pathology-defined group with the highest rate of false negatives was the CERAD moderate group, all of whom were classified as negative by the high thresholds. Subjects with intermediate levels of plaque pathology are likely to be the most problematic for detection, and only 8% of our cases (and 15% of cases in the

florbetapir series (Clark *et al.*, 2012)) were in this category. In contrast, in a large sample of subjects recruited from the community (with normal cognition, MCI, or dementia), intermediate-level Alzheimer's disease pathology (by CERAD and NIA-Reagan criteria) accounted for between 20-30% of cases with Alzheimer's disease as the sole pathological diagnosis (Schneider *et al.*, 2009). Thus, existing studies, which have recruited primarily either end-of-life or tertiary dementia center populations, appear to under-sample a very common intermediate degree of amyloid pathology that is most problematic for image-guided classification. Larger series including more individuals with intermediate pathology will be required to accurately assess imaging-pathological correlations. It is important to note that while a low number of intermediate cases may have biased our study to over-estimate sensitivity, our cohort included a larger proportion of individuals with sparse NP (16%). Reassuringly, there were no false positives in this group (Fig. 5), which bodes well for the specificity of our proposed low thresholds in future studies.

Many individuals with intermediate DVR/SUVR values (i.e. in between the low and high thresholds) had advanced amyloid pathology at autopsy. Cases in which binding appeared patchy on PET were usually found to have more extensive amyloid deposition at autopsy. Individuals with primarily diffuse plaques or CAA were not classified as PIB-positive, though PIB binds DPs and CAA *in vitro* (Lockhart *et al.*, 2007), and rare "false positive" PIB scans have been reported in individuals with florid deposition of DPs or pure CAA (Kantarci *et al.*, 2012, Ducharme *et al.*, 2013). Overall we conclude that PIB primarily detects relatively advanced Aβ NP deposition, and that even early PIB signal may merely represent the "tip of the iceberg" of

underlying amyloid pathology. This observation suggests that a negative scan, particularly with widely used high thresholds, does not exclude amyloid pathology.

A secondary goal of this study was to examine the spatial pattern of earliest Aβ accumulation detectable with PIB. In order to capture "progression" of Aβ signal, we ranked subjects based on their PIB Index DVR values and performed iterative voxel-wise contrasts with a low DVR group using a sliding window. We found that early PIB binding is seen in anterior cingulate cortex and other medial frontal regions (Fig. 3, Video), consistent with previous reports in preclinical AD (Sperling *et al.*, 2009). The anterior cingulate and precuneus were also the regions showing the highest discriminant power in in the cluster analyses (Fig. 2). Anterior cingulate and precuneus are highly inter-connected cortical hubs, and early PIB signal in these regions reinforces the hypothesis that regions of high connectivity are prone to earlier Aβ deposition (Buckner *et al.*, 2009). Overall, our cross-sectional analysis suggests that Aβ may spread from medial to lateral frontal and parietal regions, later involving the lateral temporal lobe. By the time a person reaches the widely used SUVR thresholds of 1.40 and 1.50, Aβ is already widespread across most of association cortex, relatively sparing the occipital lobes and unimodal processing regions. Of course these longitudinal inferences are based on cross-sectional data and should thus be interpreted with caution. Furthermore, since PIB mainly binds to the fibrillar form of Aβ, we cannot exclude the possibility that other forms of Aβ may have different spatial patterns.

The major limitation of all studies trying to set a threshold for Aβ positivity is that these thresholds are highly dependent on the specific imaging methodology

employed. This includes the radiotracer used, the time period following injection when the images are acquired, the reference region, and whether partial volume correction is employed. For instance, previous studies showed that applying partial volume correction increased PIB values by 10 to 20% in normal controls (Rabinovici *et al.*, 2010, Villemagne *et al.*, 2011) while choosing the whole cerebellum as the reference region decreased the values by about 6% (unpublished data). Since our four analytical methods gave almost identical thresholds, one can argue that all of the techniques used in this study were equivalent. Therefore, using a group of young adults might be an efficient way to set a low threshold. Our autopsy validation cohort consisted mainly of patients with dementia, many of whom had significant cortical atrophy. This could potentially lower overall PIB retention values due to partial volume effects, thus under-estimating sensitivity and over-estimating specificity of low thresholds when they are applied to cognitively normal individuals who will likely have less brain atrophy. Many autopsies were performed prior to publication of the updated NIA-AA pathological criteria for AD, and Thal amyloid stages were available only in a subset of cases. Our mean PET to autopsy interval was relatively long compared to previous studies. While longer PET to post-mortem intervals have been proposed as a possible explanation for false negative cases (presumably due to continued Aβ aggregation between the scan and death) (Clark *et al.*, 2012), we found no evidence of this in our series, with no difference in the PET-to-autopsy interval between false negative cases and the rest of the cohort. PIB index was measured using different processing pipelines in the derivation and validation cohorts. However, these processing streams yield highly correlated values, and the robustness of the low thresholds across two processing platforms strengthens our results. Finally, our study was designed to define and test thresholds for mean cortical PIB retention.

Approaches that examine region-specific binding should be explored, as they may detect early signal prior to more global values becoming positive. We also did not include the striatum in the PIB index – this region may be sensitive to early Aβ accumulation in autosomal dominant forms of AD (Bateman *et al.*, 2012).

In summary we found that the low PIB DVR=1.08 (or SUVR=1.21) thresholds showed higher sensitivity than the higher thresholds typically applied in the literature (DVR=1.20; SUVR=1.40) without compromising specificity. Further PET to pathology correlative studies are needed to validate these findings, and to define optimal thresholds for other Aβ tracers and image analysis approaches.

ACKNOWLEDGMENTS

We would like to acknowledge Suzanne Baker, Mustafa Janabi, Kris Norton and James O'Neil for their support with PET scanning; Dennis Dickson, Mario Mendez, John Trojanowski and Harry Vinters for patient referrals and autopsies; and Helaine St-Amant and Angie Yi for help with data preparation and helpful discussions throughout the project.

FUNDING

This work was supported by National Institute on Aging grants K23-AG031861 and R01-AG045611 to G.D.R., P01-AG1972403 to B.LM. and W.W.S, P50-AG023501 to B.L.M., G.D.R and W.W.S., R01-AG032306 and K24-AG045333 to H.J.R, P01-AG12435, P30-AG10129, R01-AG021028 and R01-AG031563 to C.D., R01-AG031563 to B.R.; R01AG038791 to A.L.B., R01-AG034570 to WJJ; the Consortium for Frontotemporal Dementia Research to B.L.M. and W.W.S; the Tau Consortium to W.W.S., G.D.R and W.J.J.; John Douglas French Alzheimer's Foundation to G.D.R. and B.L.M.; State of California Department of Health Services Alzheimer's Disease Research Center of California grant 04-33516 to B.L.M; Hellman Family Foundation award to G.D.R.; and Canadian Institutes of Health Research post-doctoral fellowship to S.V.

REFERENCES

Consensus recommendations for the postmortem diagnosis of Alzheimer's disease. The National Institute on Aging, and Reagan Institute Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease. Neurobiol Aging. 1997;18(4 Suppl):S1-2.

Aizenstein HJ, Nebes RD, Saxton JA, Price JC, Mathis CA, Tsopelas ND, et al. Frequent amyloid deposition without significant cognitive impairment among the elderly. Archives of neurology. 2008;65(11):1509-17.

Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 2011;7(3):270-9.

Armstrong MJ, Litvan I, Lang AE, Bak TH, Bhatia KP, Borroni B, et al. Criteria for the diagnosis of corticobasal degeneration. Neurology. 2013;80(5):496-503.

Bateman RJ, Xiong C, Benzinger TL, Fagan AM, Goate A, Fox NC, et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. N Engl J Med. 2012;367(9):795-804.

Bateman RJ, Xiong C, Benzinger TL, Fagan AM, Goate A, Fox NC, et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. N Engl J Med. 2012;367(9):795-804.

Braak H, Braak E. Diagnostic criteria for neuropathologic assessment of Alzheimer's disease. Neurobiology of Aging. 1998;18(4 (supplement 1)):S85-S8.

Buckner RL, Sepulcre J, Talukdar T, Krienen FM, Liu H, Hedden T, et al. Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability,

and relation to Alzheimer's disease. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2009;29(6):1860-73.

Chui HC, Zarow C, Mack WJ, Ellis WG, Zheng L, Jagust WJ, et al. Cognitive impact of subcortical vascular and Alzheimer's disease pathology. Annals of neurology. 2006;60(6):677-87.

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.

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 neurology. 2012;11(8):669-78.

Clark CM, Schneider JA, Bedell BJ, Beach TG, Bilker WB, Mintun MA, et al. Use of florbetapir-PET for imaging beta-amyloid pathology. Jama. 2011;305(3):275-83.

Curtis C, Gamez JE, Singh U, Sadowsky CH, Villena T, Sabbagh MN, et al. Phase 3 Trial of Flutemetamol Labeled With Radioactive Fluorine 18 Imaging and Neuritic Plaque Density. JAMA Neurol. 2015.

Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage. 2006;31(3):968-80.

Driscoll I, Troncoso JC, Rudow G, Sojkova J, Pletnikova O, Zhou Y, et al. Correspondence between in vivo (11)C-PiB-PET amyloid imaging and postmortem, region-matched assessment of plaques. Acta Neuropathol. 2012;124(6):823-31.

Ducharme S, Guiot MC, Nikelski J, Chertkow H. Does a positive Pittsburgh Compound B scan in a patient with dementia equal Alzheimer disease? JAMA Neurol. 2013;70(7):912-4.

Fleisher AS, Chen K, Quiroz YT, Jakimovich LJ, Gomez MG, Langois CM, et al. Florbetapir PET analysis of amyloid-beta deposition in the presenilin 1 E280A autosomal dominant Alzheimer's disease kindred: a cross-sectional study. Lancet neurology. 2012;11(12):1057-65.

Forman MS, Farmer J, Johnson JK, Clark CM, Arnold SE, Coslett HB, et al. Frontotemporal dementia: clinicopathological correlations. Ann Neurol. 2006;59(6):952-62.

Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, et al. Classification of primary progressive aphasia and its variants. Neurology. 2011;76(11):1006-14.

Grinberg LT, Wang X, Wang C, Sohn PD, Theofilas P, Sidhu M, et al. Argyrophilic grain disease differs from other tauopathies by lacking tau acetylation. Acta Neuropathol. 2013;125(4):581-93.

Hedden T, Van Dijk KR, Becker JA, Mehta A, Sperling RA, Johnson KA, et al. Disruption of functional connectivity in clinically normal older adults harboring amyloid burden. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2009;29(40):12686-94.

Hyman BT, Phelps CH, Beach TG, Bigio EH, Cairns NJ, Carrillo MC, et al. National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2012;8(1):1-13.

Hyman BT, Trojanowski JQ. Consensus recommendations for the postmortem diagnosis of Alzheimer disease from the National Institute on Aging and the Reagan Institute Working Group on diagnostic criteria for the neuropathological assessment of Alzheimer disease. Journal of neuropathology and experimental neurology. 1997;56(10):1095-7.

Ikonomovic MD, Abrahamson EE, Price JC, Hamilton RL, Mathis CA, Paljug WR, et al. Early AD pathology in a [C-11]PiB-negative case: a PiB-amyloid imaging, biochemical, and immunohistochemical study. Acta Neuropathol. 2012;123(3):433-47.

Ikonomovic MD, Klunk WE, Abrahamson EE, Mathis CA, Price JC, Tsopelas ND, et al. Post-mortem correlates of in vivo PiB-PET amyloid imaging in a typical case of Alzheimer's disease. Brain. 2008;131(Pt 6):1630-45.

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.

Jack CR, Jr., Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. Lancet Neurol. 2010;9(1):119-28.

Jack CR, Jr., Knopman DS, Weigand SD, Wiste HJ, Vemuri P, Lowe V, et al. An operational approach to National Institute on Aging-Alzheimer's Association criteria for preclinical Alzheimer disease. Annals of neurology. 2012;71(6):765-75.

Jack CR, Jr., Lowe VJ, Senjem ML, Weigand SD, Kemp BJ, Shiung MM, et al. 11C PiB and structural MRI provide complementary information in imaging of Alzheimer's disease and amnestic mild cognitive impairment. Brain : a journal of neurology. 2008;131(Pt 3):665-80.

Jack CR, Jr., Wiste HJ, Lesnick TG, Weigand SD, Knopman DS, Vemuri P, et al. Brain beta-amyloid load approaches a plateau. Neurology. 2013;80(10):890-6.

Kantarci K, Yang C, Schneider JA, Senjem ML, Reyes DA, Lowe VJ, et al. Antemortem amyloid imaging and beta-amyloid pathology in a case with dementia with Lewy bodies. Neurobiol Aging. 2012;33(5):878-85.

Klunk WE, Engler H, Nordberg A, Wang Y, Blomqvist G, Holt DP, et al. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. Annals of neurology. 2004;55(3):306-19.

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. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2015;11(1):1-15 e4.

Kok E, Haikonen S, Luoto T, Huhtala H, Goebeler S, Haapasalo H, et al. Apolipoprotein E-dependent accumulation of Alzheimer disease-related lesions begins in middle age. Annals of neurology. 2009;65(6):650-7.

Landau SM, Thomas BA, Thurfjell L, Schmidt M, Margolin R, Mintun M, et al. Amyloid PET imaging in Alzheimer's disease: a comparison of three radiotracers. Eur J Nucl Med Mol Imaging. 2014.

Lockhart A, Lamb JR, Osredkar T, Sue LI, Joyce JN, Ye L, et al. PIB is a non-specific imaging marker of amyloid-beta (Abeta) peptide-related cerebral amyloidosis. Brain : a journal of neurology. 2007;130(Pt 10):2607-15.

Logan J, Fowler JS, Volkow ND, Wang GJ, Ding YS, Alexoff DL. Distribution volume ratios without blood sampling from graphical analysis of PET data. Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism. 1996;16(5):834-40.

Mackenzie IR, Neumann M, Bigio EH, Cairns NJ, Alafuzoff I, Kril J, et al. Nomenclature and nosology for neuropathologic subtypes of frontotemporal lobar degeneration: an update. Acta Neuropathol. 2010;119(1):1-4.

Magaki S, Yong WH, Khanlou N, Tung S, Vinters HV. Comorbidity in Dementia: Update of an Ongoing Autopsy Study. J Am Geriatr Soc. 2014.

McKeith IG, Dickson DW, Lowe J, Emre M, O'Brien JT, Feldman H, et al. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. Neurology. 2005;65(12):1863-72.

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

McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Jr., Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 2011;7(3):263-9.

Mirra SS, Heyman A, McKeel D, Sumi SM, Crain BJ, Brownlee LM, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part II. Standardization of the neuropathologic assessment of Alzheimer's disease. Neurology. 1991;41(4):479-86.

Montine TJ, Phelps CH, Beach TG, Bigio EH, Cairns NJ, Dickson DW, et al. National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease: a practical approach. Acta Neuropathol. 2012;123(1):1-11.

Mormino EC, Betensky RA, Hedden T, Schultz AP, Ward A, Huijbers W, et al. Amyloid and APOE {varepsilon}4 interact to influence short-term decline in preclinical Alzheimer disease. Neurology. 2014.

Mormino EC, Brandel MG, Madison CM, Rabinovici GD, Marks S, Baker SL, et al. Not quite PIB-positive, not quite PIB-negative: slight PIB elevations in elderly normal control subjects are biologically relevant. NeuroImage. 2012;59(2):1152-60.

Murray ME, Cannon A, Graff-Radford NR, Liesinger AM, Rutherford NJ, Ross OA, et al. Differential clinicopathologic and genetic features of late-onset amnestic dementias. Acta Neuropathol. 2014;128(3):411-21.

Newell KL, Hyman BT, Growdon JH, Hedley-Whyte ET. Application of the National Institute on Aging (NIA)-Reagan Institute criteria for the neuropathological diagnosis of Alzheimer disease. J Neuropathol Exp Neurol. 1999;58(11):1147-55.

Ni R, Gillberg PG, Bergfors A, Marutle A, Nordberg A. Amyloid tracers detect multiple binding sites in Alzheimer's disease brain tissue. Brain. 2013;136(Pt 7):2217-27.

Nordberg A, Carter SF, Rinne J, Drzezga A, Brooks DJ, Vandenberghe R, et al. A European multicentre PET study of fibrillar amyloid in Alzheimer's disease. European journal of nuclear medicine and molecular imaging. 2013;40(1):104-14.

Petersen RC. Mild cognitive impairment as a diagnostic entity. J Intern Med. 2004;256(3):183-94.

Pike KE, Savage G, Villemagne VL, Ng S, Moss SA, Maruff P, et al. Beta-amyloid imaging and memory in non-demented individuals: evidence for preclinical Alzheimer's disease. Brain : a journal of neurology. 2007;130(Pt 11):2837-44.

Price JC, Klunk WE, Lopresti BJ, Lu X, Hoge JA, Ziolko SK, et al. Kinetic modeling of amyloid binding in humans using PET imaging and Pittsburgh Compound-B.

Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism. 2005;25(11):1528-47.

Rabinovici GD, Furst AJ, Alkalay A, Racine CA, O'Neil JP, Janabi M, et al. Increased metabolic vulnerability in early-onset Alzheimer's disease is not related to amyloid burden. Brain : a journal of neurology. 2010;133(Pt 2):512-28.

Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain. 2011;134(Pt 9):2456-77.

Roman GC, Tatemichi TK, Erkinjuntii T. Vascular Dementia: Diagnostic criteria for research studies. Neurology. 1993;43:250-60.

Rowe CC, Ellis KA, Rimajova M, Bourgeat P, Pike KE, Jones G, et al. Amyloid imaging results from the Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging. Neurobiology of aging. 2010;31(8):1275-83.

Salloway S, Sperling R, Fox NC, Blennow K, Klunk W, Raskind M, et al. Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. N Engl J Med. 2014;370(4):322-33.

Schneider JA, Arvanitakis Z, Leurgans SE, Bennett DA. The neuropathology of probable Alzheimer disease and mild cognitive impairment. Ann Neurol. 2009;66(2):200-8.

Sojkova J, Driscoll I, Iacono D, Zhou Y, Codispoti KE, Kraut MA, et al. In vivo fibrillar beta-amyloid detected using [11C]PiB positron emission tomography and neuropathologic assessment in older adults. Arch Neurol. 2011;68(2):232-40.

Sperling RA, Laviolette PS, O'Keefe K, O'Brien J, Rentz DM, Pihlajamaki M, et al. Amyloid deposition is associated with impaired default network function in older persons without dementia. Neuron. 2009;63(2):178-88.

Sun FT, Schriber RA, Greenia JM, He J, Gitcho A, Jagust WJ. Automated template-based PET region of interest analyses in the aging brain. NeuroImage. 2007;34(2):608-17.

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage. 2002;15(1):273-89.

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 neurology. 2013;12(4):357-67.

Villemagne VL, Pike KE, Chetelat G, Ellis KA, Mulligan RS, Bourgeat P, et al. Longitudinal assessment of Abeta and cognition in aging and Alzheimer disease. Annals of neurology. 2011;69(1):181-92.

Villeneuve S, Reed BR, Wirth M, Haase CM, Madison CM, Ayakta N, et al. Cortical thickness mediates the effect of beta-amyloid on episodic memory. Neurology. 2014;82(9):761-7.

TABLES

Table 1: Participant demographics

	Threshold derivation study		Threshold validation study
	Young	Older adults	Older adults
n	18	154	50
Male (%)	8 (44%)	69 (45%)	33 (66%)
Age at PET	23.7 (2.9)	76.0 (6.2)	69.8 (9.6)
Education	13.2 (6.2)	16.1 (2.5)	15.7 (2.9)
MMSE	-	28.9 (1.2)	21.6 (6.8)
ApoE4 (%)	7 (47%)[a]	45 (31%)[b]	13 (27%)[c]
CDR	-	-	1.2 (0.8)
Time from PET to death (years)	-	-	3.1 (1.8)
PIB index DVR	1.01 (.03)	1.12 (.20)	1.14 (.29)[d]
PIB+ (%), DVR cutoff 1.08	0 (0%)	57 (37%)	18 (40%)[d]
PIB+ (%), DVR cutoff 1.20	0 (0%)	34 (22%)	13 (29%)[d]
Clinical diagnosis at PET	Normal (18)	Normal (154)	FTD[e] (28), AD (11), MCI (7), AD/SIVD (1), DLB (1), normal (2)
Primary neuropathological diagnosis	-	-	FTLD[c] (25), AD (6), CVD (6), AD/CVD (3), AD/DLB (3), AD/FTLD (4), TPD (1), AGD (1), no pathological findings (1)
CERAD (absent, sparse, moderate, frequent)	-	-	18, 8, 4, 20

Shown are mean (standard deviation) unless specified otherwise.

a - 3 subjects were not genotyped

b - 8 subjects were not genotyped

c - 2 subjects were not genotyped

d- DVR data missing for 5 subjects

e - Clinical syndromes included: corticobasal syndrome (CBS; 8), behavioral-variant FTD (6), FTD and amyotrophic lateral sclerosis (5), non-fluent variant primary progressive aphasia (nfvPPA; 5), nfvPPA/CBS (1), semantic variant PPA (3).

c- FTLD neuropathological subtypes: FTLD-TDP (12), corticobasal degeneration (7), Pick's disease (3), progressive supranuclear palsy (2), FTLD with non-specific 4 repeat tauopathy (1).

AD = Alzheimer's disease; AGD = argyrophillic grain disease; ApoE = Apolipoprotein E; CDR = Clinical Dementia Rating; CVD = cerebrovascular disease; DVR = distribution volume ratios; MCI = mild cognitive impairment; DLB = dementia with Lewy bodies; FTD = frontotemporal dementia; FTLD = frontotemporal lobar degeneration; MMSE = Mini mental state examination; PIB = Pittsburgh compound B; SIVD = subcortical ischemic vascular disease; TPD = tangle-predominant dementia.

Table 2: Cutoffs for capturing early amyloid PET positivity

	2 SD above Young (n=18)	Voxel-wise analysis	Cluster analysis	GMM analysis	**Optimal Cutoff**
DVR	1.07	1.07	1.09	1.09	**1.08**
SUVR	1.19	1.19	1.23	1.23	**1.21**

Shown are the optimal cutoffs derived using four independent data analysis techniques as well as the optimal thresholds proposed based on the results of these four analyses. The analyses were performed using the DVR data and a regression line (SUVR = -0.54 + 1.62*DVR) was fit between the DVR and the SUVR data to calculate the corresponding SUVR cutoffs.

DVR = distribution volume ratio; GMM = Gaussian mixture modeling; SUVR standardized uptake value ratio.

Table 3. Sensitivity and specificity for the detection of CERAD moderate-frequent neuritic plaques.

	Sensitivity	Specificity
Low Thresholds		
DVR = 1.08	**81.0%** (57.4% - 93.7%)	**95.8%** (76.9% - 99.8%)
SUVR = 1.21	**83.3%** (61.8% - 94.5%)	**100.0%** (84.0% - 100%)
High thresholds		
DVR = 1.20	**61.9%** (38.7% - 81.0%)	**100.0%** (82.8% -100%)
SUVR = 1.40	**62.5%** (40.8% - 80.4%)	**100.0%** (84.0% - 100%)

95% confidence intervals shown in parantheses.

DVR = Distribution Volume Rato; SUVR = standradized uptake value ratio

FIGURE LEGENDS:

Fig. 1: Gaussian mixture model (GMM) containing 2 mixtures (distributions) that best fit the Pittsburgh compound B (PIB) index values. The blue line represents the distribution associated with low Aβ values while the red line represents the distribution associated with high Aβ values. The two Gaussian distributions are superimposed on the subject density histogram for all PIB index values in older subjects.

Fig. 2: Cluster analysis containing 2 clusters (groups) representing individuals with low and high amyloid burden (A), as well as the 13 features (brain regions) that helped identify these two clusters (B). Cluster 1 (on the left) represents subjects with low Aβ values while cluster 2 (on the right) represents subjects with high Aβ values. Warmer colors are associated with higher distribution volume ratio (DVR) values (see the DVR color scale ranging from 0.5 to 2.5 on the right side of the figure). For each cluster, each row represents one subject and each column represents one of the 13 features that helped identify the two clusters. From left to right, the 13 features are: rostral anterior cingulate left hemisphere (lh), rostral anterior cingulate right hemisphere (rh), precuneus rh, precuneus lh, medial orbitofrontal rh, rostral middle frontal lh, rostral middle frontal rh, inferior parietal rh, medial orbitofrontal lh, superior orbitofrontal rh, posterior cingulate rh, superior orbitofrontal lh. These 13 features are also projected on a brain with the lighter colors corresponding to the features that have the highest weight in the model (see Table 3 for the weight of each feature in the model).

Fig. 3: Pattern of early detectable PIB binding in cognitively normal older adults. Each row of each image reflect a voxel-wise contrast of 22 subjects with the mean

values for the DVR/SUVR index listed on the left compared to a reference group (n = 22) with a DVR index of 1.00. Significant voxels first appeared when the group mean is DVR = 1.07 (see also Video showing all the voxel-wise analyses).

Threshold set at $p < .05$ after family-wise error correction, $k > 150$.

DVR = distribution volume ratio; SUVR standardized uptake value ratio

Fig. 4: PIB-PET versus post-mortem amyloid. Trans-axial PIB slices from a patient with CERAD sparse (top row), moderate (middle row) and frequent (bottom row) neuritic plaques. PIB-PET Trans-axial slices are presented in neurological orientation. Photomicrographs of Aβ immunohistochemistry are shown at 10x (top and bottom rows) or 20x (middle row) magnification.

DVR = Distribution Volume Rato; SUVR = standradized uptake value ratio; L = left; R = right.

Fig. 5: Scatterplots of PIB index DVR (left) and SUVR (right) versus CERAD rating. Low thresholds are signified by dotted horizontal lines, and high thresholds by solid horizontal lines.

ABS = absent; MOD = moderate; FREQ = frequent

Video In order to capture "progression" of Aβ, we ranked subjects based on their DVR index values and performed repeated voxel-wise analyses using a sliding window. Using this technique, we found that early PIB binding is found in the medial frontal lobe, including the anterior cingulate/orbitofrontal cortex, at a mean DVR of 1.07. Our findings also suggest that fibrillar Aβ may spread from medial to lateral frontal and parietal cortex, and later to the lateral temporal lobe. In this model, we

assume that increasing DVR represents temporal progression, an assumption that we recognize may not be entirely justified.

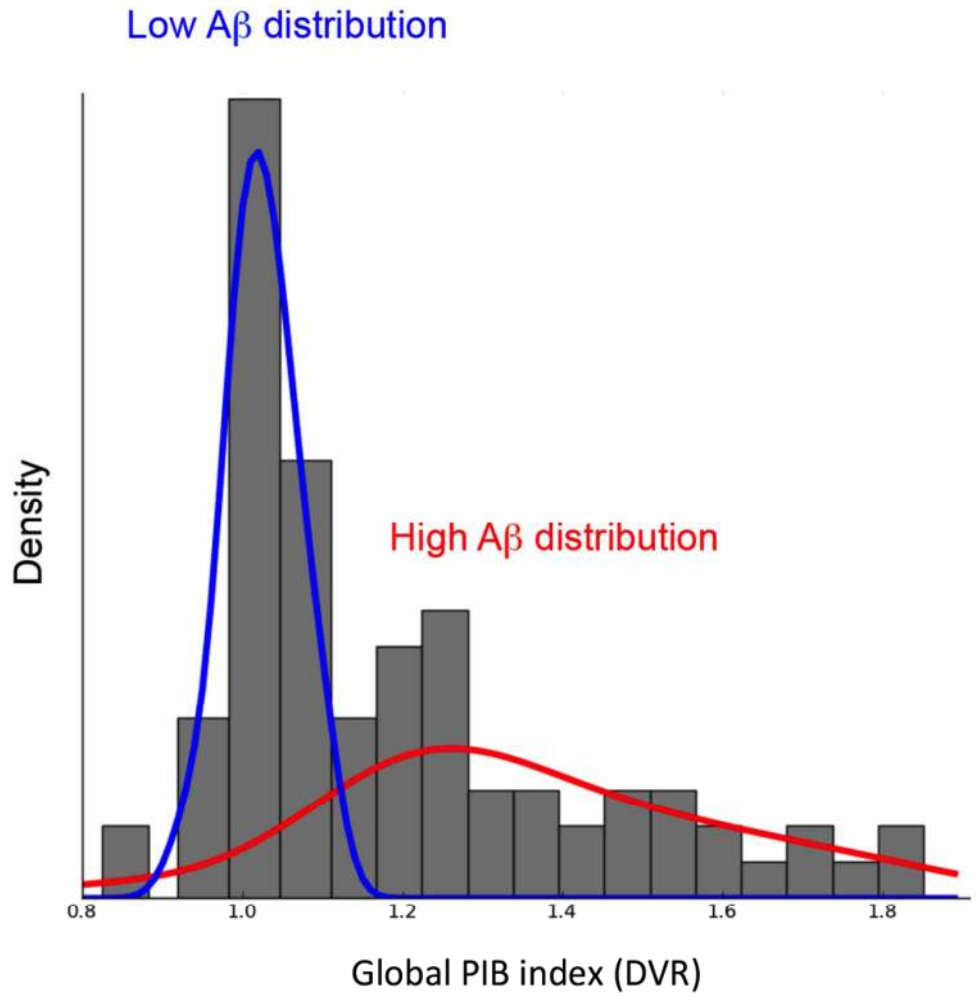

Fig. 1: Gaussian mixture model (GMM) containing 2 mixtures (distributions) that best fit the Pittsburgh compound B (PIB) index values. The blue line represents the distribution associated with low Aβ values while the red line represents the distribution associated with high Aβ values. The two Gaussian distributions are superimposed on the subject density histogram for all PIB index values in older subjects.
102x103mm (300 x 300 DPI)

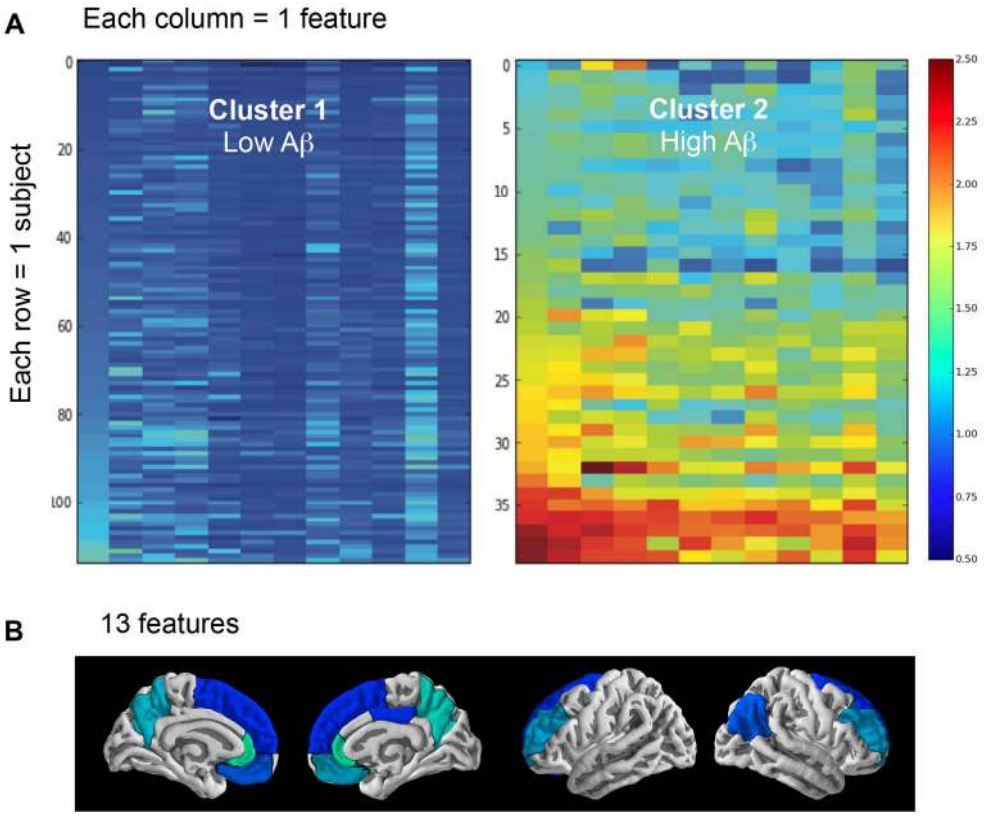

Fig. 2: Cluster analysis containing 2 clusters (groups) representing individuals with low and high amyloid burden (A), as well as the 13 features (brain regions) that helped identify these two clusters (B). Cluster 1 (on the left) represents subjects with low Aβ values while cluster 2 (on the right) represents subjects with high Aβ values. Warmer colors are associated with higher distribution volume ratio (DVR) values (see the DVR color scale ranging from 0.5 to 2.5 on the right side of the figure). For each cluster, each row represents one subject and each column represents one of the 13 features that helped identify the two clusters. From left to right, the 13 features are: rostral anterior cingulate left hemisphere (lh), rostral anterior cingulate right hemisphere (rh), precuneus rh, precuneus lh, medial orbitofrontal rh, rostral middle frontal lh, rostral middle frontal rh, inferior parietal rh, medial orbitofrontal lh, superior orbitofrontal rh, posterior cingulate rh, superior orbitofrontal lh. These 13 features are also projected on a brain with the lighter colors corresponding to the features that have the highest weight in the model (see Table 3 for the weight of each feature in the model).
176x144mm (300 x 300 DPI)

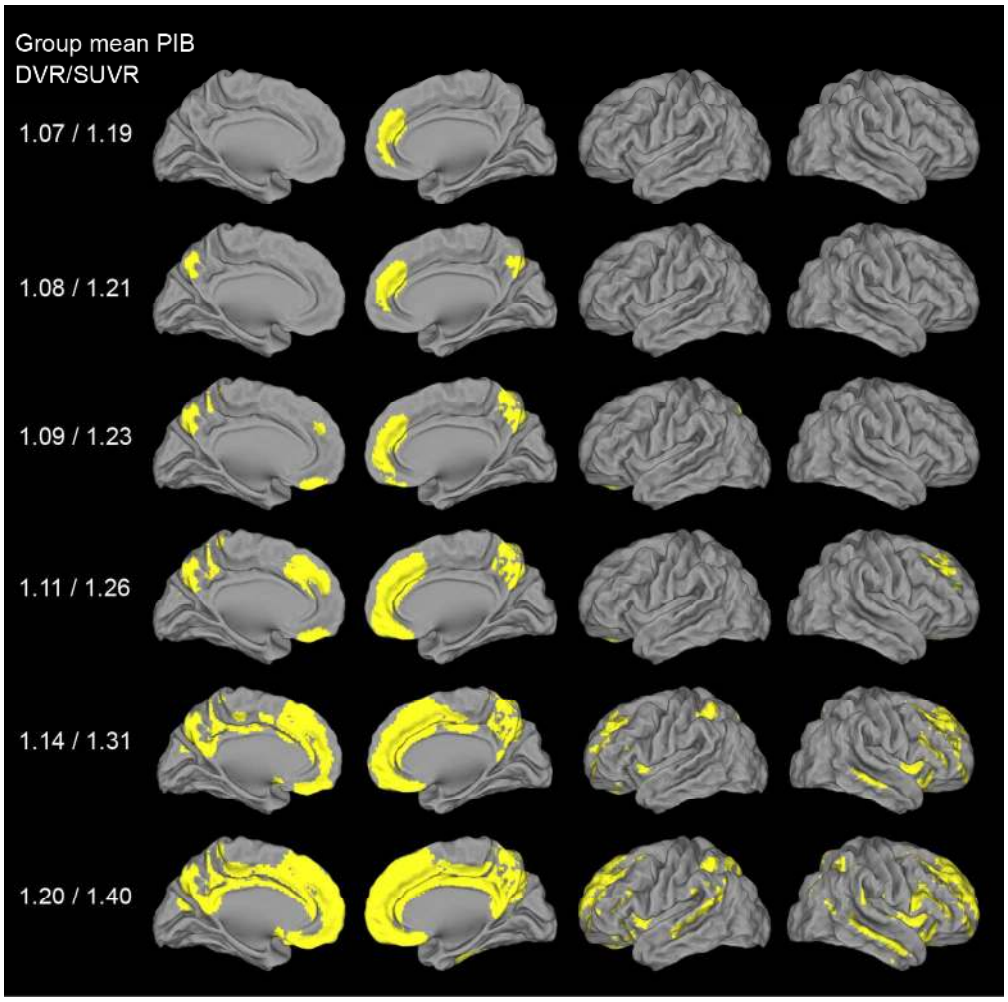

Fig. 3: Pattern of early detectable PIB binding in cognitively normal older adults. Each row of each image reflect a voxel-wise contrast of 22 subjects with the mean values for the global DVR/SUVR listed on the left compared to a reference group (n = 22) with a global DVR of 1.00. Significant voxels first appeared when the group mean is DVR = 1.07 (see also Video 1 showing all the voxel-wise analyses).
Threshold set at $p < .05$ after family-wise error correction, $k > 150$.
DVR = distribution volume ratio; SUVR standardized uptake value ratio
141x139mm (300 x 300 DPI)

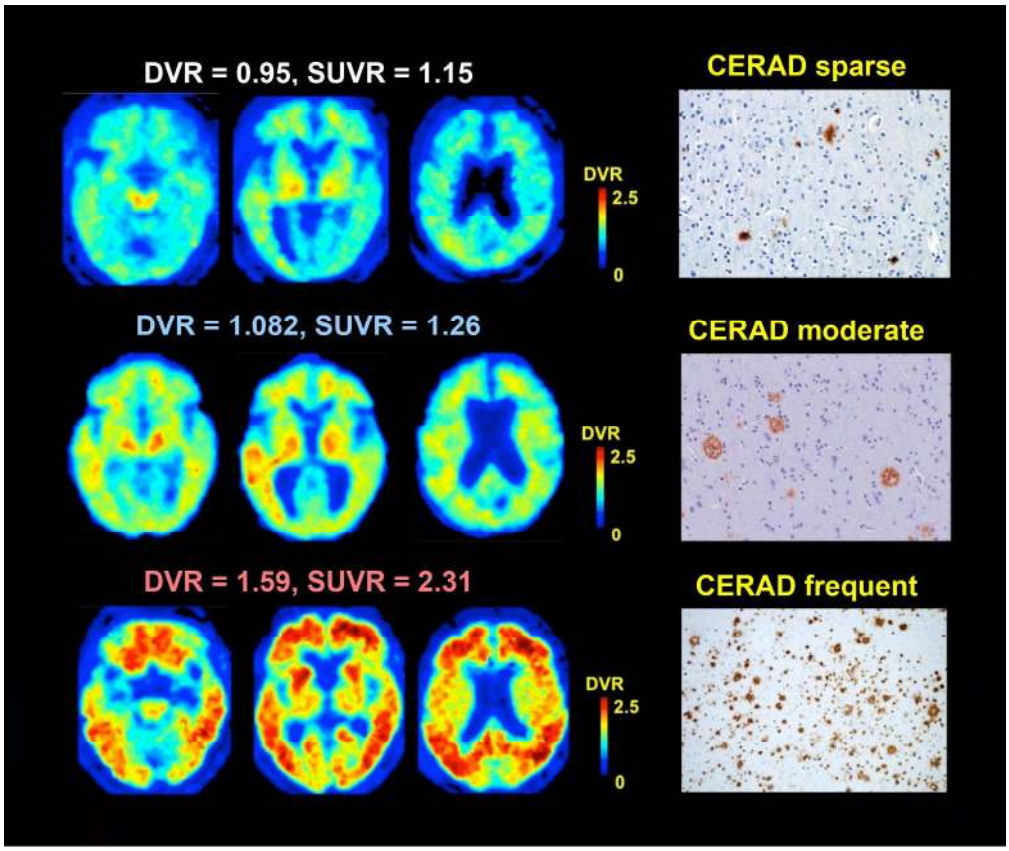

Fig 4: PIB-PET versus post-mortem amyloid. Trans-axial PIB slices from a patient with CERAD sparse (top row), moderate (middle row) and frequent (bottom row) neuritic plaques. PIB-PET Trans-axial slices are presented in neurological orientation. Photomicrographs of Aβ immunohistochemistry are shown at 10x (top and bottom rows) or 20x (middle row) magnification.
DVR = Distribution Volume Rato; SUVR = standradized uptake value ratio; L = left; R = right.
224x188mm (300 x 300 DPI)

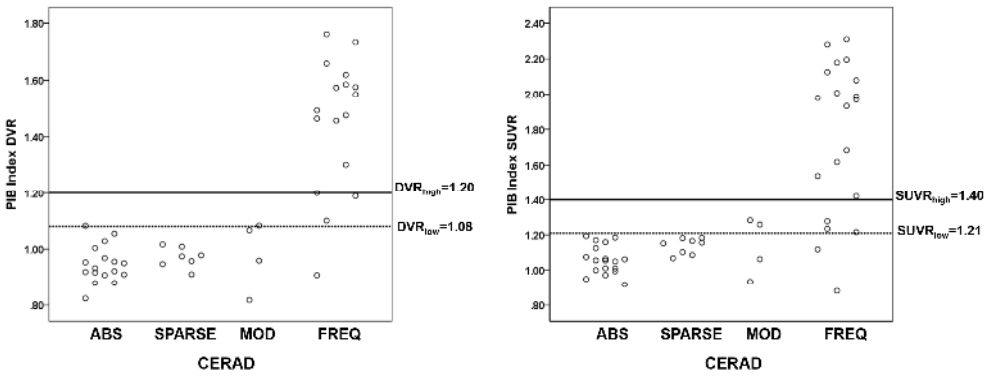

Fig 5: Scatterplots of PIB index DVR (left) and SUVR (right) versus CERAD rating. Low thresholds are signified by dotted horizontal lines, and high thresholds by solid horizontal lines.
ABS = absent; MOD = moderate; FREQ = frequent
241x90mm (300 x 300 DPI)

Supplementary Material

Supplementary Methods

Threshold derivation

Participants

For the derivation thresholds, 109 subjects were from the Berkley Aging Cohort (BAC) (Marchant *et al.*, 2012, Oh *et al.*, 2013, Wirth *et al.*, 2014), and 44 were recruited from ongoing studies at the University of California Davis (UCD) Alzheimer's Disease Center (Reed *et al.*, 2014, Villeneuve *et al.*, 2014). All young subjects were enrolled in the BAC. None of the subjects included in the current study were participants in ADNI.

The Mini Mental State Exam (MMSE) (Folstein *et al.*, 1975) was used as a standard clinical measure of global cognition. Cognitively normal elderly all had a MMSE ≥ 26 and, when applicable, a CDR of 0. The CDR was not available for the BAC subjects, but the latter had an exhaustive neuropsychological evaluation in order to exclude the presence of cognitive deficits (Oh *et al.*, 2013, Wirth *et al.*, 2014). Severe or unstable medical illness, neurologic or psychiatric disorders that could significantly affect brain function or structure, alcohol or substance abuse or significant head injury were also exclusion criteria. Participants from UCD and the Aging Brain project were recruited to emphasize vascular risk factors, and 13 subjects with cortical stroke were excluded from this study.

MR imaging and preprocessing

Structural T1-weighted MR images were obtained on two 1.5 T instruments (Siemens Avanto and GE Signa Genesis), a 3T instrument (Siemens Tim Trio) or a 4T

instrument (Siemens MedSpec Syngo); scans were used only for definition of regions of interest (ROIs) for PET data analysis. Acquisition parameters have been described previously (Marchant *et al.*, 2012, Wirth *et al.*, 2013, Villeneuve *et al.*, 2014).

PET imaging and processing

All subjects underwent PIB-PET imaging at Lawrence Berkeley National Laboratory on a Siemens ECAT EXACT HR PET scanner (Siemens Medical Systems, Erlangen Germany) or a Siemens Biograph scanner. PIB was synthesized at LBNL using a published protocol (Mathis *et al.*, 2003). After ~15 mCi of [^{11}C] PIB was injected into an antecubital vein, dynamic acquisition frames were obtained over 90 minutes in 3D mode (35 frames total: 4 × 15 seconds, 8 × 30 seconds, 9 × 60 seconds, 2 × 180 seconds, 10 × 300 seconds, and 2 × 600 seconds) (Villeneuve *et al.*, 2014).

For each subject, both a DVR "PIB index" and a SUVR "PIB index" were derived from the native-space image by averaging the weighted mean value (weighted by size of the ROI) from FreeSurfer-derived ROIs in frontal (cortical regions anterior to the precentral gyrus), temporal (middle and superior temporal regions), parietal (supramarginal gyrus, inferior/superior parietal lobules and precuneus) and posterior cingulate cortex (Villeneuve *et al.*, 2014) using the Desikan-Killiany atlas (Desikan *et al.*, 2006) (supplementary Fig. 1A). These regions consistently show high PIB retention in studies of AD and aging (Rabinovici *et al.*, 2010, Mormino *et al.*, 2011).

Threshold validation

Participants

For the validation thresholds portion of the study, participants were enrolled in longitudinal studies of aging and dementia at the University of California, San Francisco Memory and Aging Center (UCSF, N=41) or the University of California, Davis Alzheimer's Disease Center (UCD, N=9). Among the 50 subjects, PiB to autopsy data were presented on 11 patients from the UCSF cohort (Rabinovici *et al.*, 2011) and none of the UC Davis subjects in other previous publications. UCSF recruitment was enriched for patients with Alzheimer's disease or frontotemporal dementia (Rabinovici *et al.*, 2011), whereas at UCD recruitment focused on individuals at high vascular risk (Villeneuve *et al.*, 2014).

Fifty-six subjects underwent PET imaging and autopsy. Three subjects were excluded because they had been unable to complete an MRI examination necessary to generate quantitative values for PIB retention. Two subjects were excluded for technical reasons, one because of excessive head motion resulting in uninterpretable images, and one because of a failed injection. One additional subject underwent brain autopsy but did not have postmortem measures of Aβ (autopsy diagnosis was prion disease). Thus, the sample for analysis included 50 individuals with both quantitative PIB-PET data and Aβ assessment at autopsy.

The clinical evaluation at both centers included a structured neurobehavioral history and examination, caregiver interview, neuropsychological testing, apolipoprotein E (ApoE) genotyping and MRI (DeCarli *et al.*, 2008, Rabinovici *et al.*, 2011, Villeneuve *et al.*, 2014).

MR imaging and preprocessing

Structural T1-weighted MR images were obtained on one of two 1.5 T instruments (Siemens Avanto and Siemens Vision), a 3T instrument (Siemens Tim Trio) or a 4T instrument (Brucker). Acquisition parameters have been described previously (Rosen *et al.*, 2002, DeCarli *et al.*, 2008, Mueller *et al.*, 2009, Mormino *et al.*, 2012, Zhou *et al.*, 2012).

PET imaging and processing

To estimate mean cortical PIB retention, we created a "PIB index" region of interest (Rabinovici *et al.*, 2010) by combining the following regions defined within the Automated Anatomical Labeling (AAL) atlas (Tzourio-Mazoyer *et al.*, 2002): superior frontal gyrus, middle frontal gyrus, frontal superior/middle/inferior orbital gyri, superior medial frontal gyrus, inferior operculum, inferior triangularis, rolandic operculum, supplementary motor areas, rectus, olfactory bulb, insula, anterior cingulate, superior/middle/inferior temporal, superior/inferior parietal lobule, supramarginal gyrus, angular gyrus, posterior cingulate, precuneus, and middle cingulum (supplementary Fig. 1B). These regions consistently show high PIB retention in studies of AD and aging (Rabinovici *et al.*, 2010, Mormino *et al.*, 2011). We calculated mean SUVRs and DVRs per subject in template space across voxels with a grey-matter probability of at least 30% within the PIB index region of interest. In order to compare values to more widely used SUVR values, all subjects who underwent dynamic scanning also had data from 50-70 min post injection averaged and divided by the grey matter cerebellar mask to calculate SUVRs. In 3 subjects, regions of cortical stroke were masked and not included in the target ROI for calculation of PIB index. In sum, DVR data were available for 45 individuals and SUVR data for 50.

Supplementary Results

Threshold validation

Applying NIA-Reagan criteria, 28% of patients were negative for Alzheimer's disease, 14% had low-likelihood Alzheimer's disease, 6% had intermediate-likelihood Alzheimer's disease, 30% had high-likelihood Alzheimer's disease and 22% were unclassifiable due to lack of concordance between CERAD and Braak staging.

Supplementary Fig. 3 provides examples of misclassified cases. There was only one false positive case in our series (top row). The DVR value was just above the DVR_{low} threshold while the SUVR was just below the $SUVR_{low}$ threshold. Autopsy 5.3 years later revealed progressive supranuclear palsy. Aβ immunohistochemistry showed no evidence of NPs, DPs or CAA. Visual inspection of the image suggests that the borderline positive DVR value may be due to relatively high PIB retention in white matter, which could contaminate grey matter PIB signal via partial volume effects. Scans were classified as false negatives (by one or both thresholds) for a variety of reasons. Focal amyloid accumulation could lead to intermediate or low global DVR and SUVR measures (second and third rows, see also Fig. 4 middle row). Significant cortical atrophy also likely contributed to low values (third row). In one instance, borderline positive DVR (1.10) and SUVR (1.23) values in a patient with NIA-Reagan Alzheimer's disease high-likelihood were explained by the presence of significant CAA in the grey matter of the cerebellum, increasing tracer retention in the reference region and thus lowering cortical PIB values (data not shown). PET to pathology intervals were no different in patients that were "false negatives" by one or

more standards (N=11, mean PET-to-autopsy 2.7±1.2 years) versus other patients in the cohort (N=39, 3.1±1.9 years, p=0.36).

Supplementary Tables and Figures

Threshold derivation

Supplementary Table 1: ROIs included in the cluster analysis

1	Left-amygdala	**39**	Right-amygdala
2	Left-caudate	**40**	Right-caudate
3	Left-hippocampus	**41**	Right-hippocampus
4	Left-pallidum	**42**	Right-pallidum
5	Left-putamen	**43**	Right-putamen
6	Left-thalamus-Proper	**44**	Right-thalamus-Proper
7	Left-caudalanteriorcingulate	**45**	Right-caudalanteriorcingulate
8	Left-caudalmiddlefrontal	**46**	Right-caudalmiddlefrontal
9	Left-cuneus	**47**	Right-cuneus
10	Left-entorhinal	**48**	Right-entorhinal
11	Left-frontalpole	**49**	Right-frontalpole
12	Left-fusiform	**50**	Right-fusiform
13	Left-inferiorparietal	**51**	Right-inferiorparietal
14	Left-inferiortemporal	**52**	Right-inferiortemporal
15	Left-insula	**53**	Right-insula
16	Left-isthmuscingulate	**54**	Right-isthmuscingulate
17	Left-lateraloccipital	**55**	Right-lateraloccipital
18	Left-lateralorbitofrontal	**56**	Right-lateralorbitofrontal
19	Left-lingual	**57**	Right-lingual
20	Left-medialorbitofrontal	**58**	Right-medialorbitofrontal
21	Left-middletemporal	**59**	Right-middletemporal
22	Left-paracentral	**60**	Right-paracentral
23	Left-parahippocampal	**61**	Right-parahippocampal
24	Left-parsopercularis	**62**	Right-parsopercularis
25	Left-parsorbitalis	**63**	Right-parsorbitalis
26	Left-parstriangularis	**64**	Right-parstriangularis
27	Left-pericalcarine	**65**	Right-pericalcarine
28	Left-postcentral	**66**	Right-postcentral
29	Left-posteriorcingulate	**67**	Right-posteriorcingulate
30	Left-precentral	**68**	Right-precentral
31	Left-precuneus	**69**	Right-precuneus
32	Left-rostralanteriorcingulate	**70**	Right-rostralanteriorcingulate
33	Left-rostralmiddlefrontal	**71**	Right-rostralmiddlefrontal
34	Left-superiorfrontal	**72**	Right-superiorfrontal
35	Left-superiorparietal	**73**	Right-superiorparietal
36	Left-superiortemporal	**74**	Right-superiortemporal
37	Left-supramarginal	**75**	Right-supramarginal
38	Left-temporalpole	**76**	Right-temporalpole

Shown are the 76 ROIs included in the cluster analysis and extracted from the Desikan-Killiany atlas.

Supplementary Table 2: Weight of each feature in the cluster analyses.

Features	Weight
Rostral anterior cingulate lh	.038
Rostral anterior cingulate rh	.036
Precuneus rh	.033
Precuneus lh	.030
Medial orbitofrontal rh	.028
Rostral middle frontal lh	.028
Rostral middle frontal rh	.028
Inferior parietal rh	.026
Medial orbitofrontal lh	.025
Superior orbitofrontal rh	.022
Posterior cingulate rh	.022
Superior orbitofrontal lh	.021

Shown are the names and the weights of the 13 significant features in the cluster analyses.

lh: left hemisphere, rh = right hemisphere.

Supplementary Table 3: Cutoffs for capturing early amyloid PET positivity based on 95%, 90% and 50% thresholds.

	95%	90%	50 %
GMM_{low}	1.05	1.09	1.16
GMM_{high}	1.23	1.21	1.16
$Cluster_{low}$	1.11	1.09	1.02
$Cluster_{high}$	1.19	1.19	1.32

Shown are the cutoffs derived using the Gaussian mixture modeling (GMM) and the Cluster analyses while using 95%, 90% and 50% probabilities for the GMM and the 95% (5%), 90% (10%) and 50% (50%) percentile for the Cluster analyses. With a cutoff of 95%, the main results of the study would still be that an optimal cutoff to define a low threshold would be set at 1.08 (i.e., the mean of the four techniques would still be 1.08). Using 50%, than the mean of the four techniques would now be 1.09, which is still highly similar.

Threshold validation

Supplementary Table 4: Receiver operating characteristic (ROC) results for the detection of CERAD moderate-frequent neuritic plaques**.**

	DVR (N=45)	SUVR (N=50)
AUC	**.891** (0.769-1.000)	**.878** (0.760-0.997)
Optimized threshold	**1.06**	**1.20**
Sensitivity	**85.7%** (62.6%- 96.2%)	**83.3%** (61.8%-91.5%)
Specificity	**95.8%** (76.9%-99.8%)	**100%** (84.0%-100%)

95% confidence intervals shown in parantheses.

DVR = Distribution Volume Rato; SUVR = standradized uptake value ratio;

AUC = area under the curve

Threshold derivation

Supplementary Fig. 1

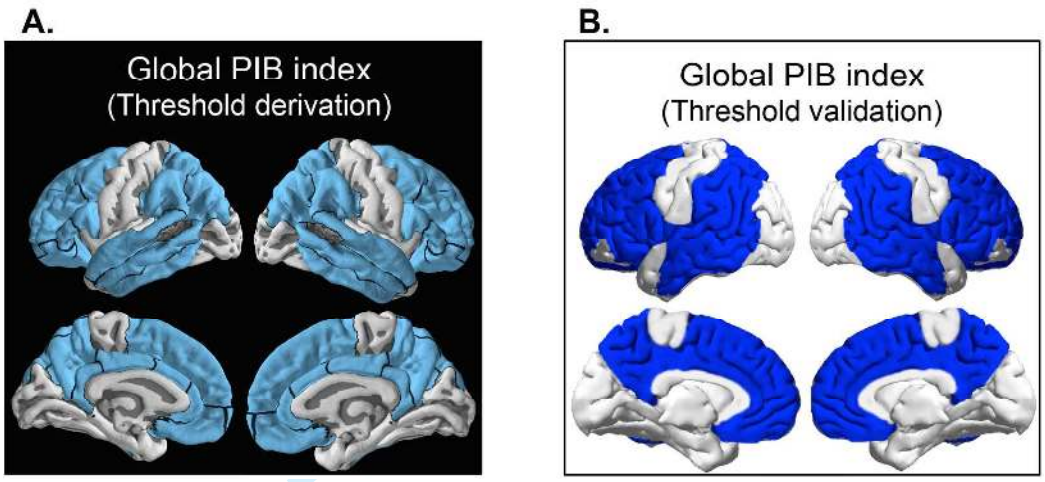

Figure legend: Shown is the global PIB index used in the threshold derivation part of the study (A) and in the threshold validation part of the study (B).

Supplementary Fig. 2

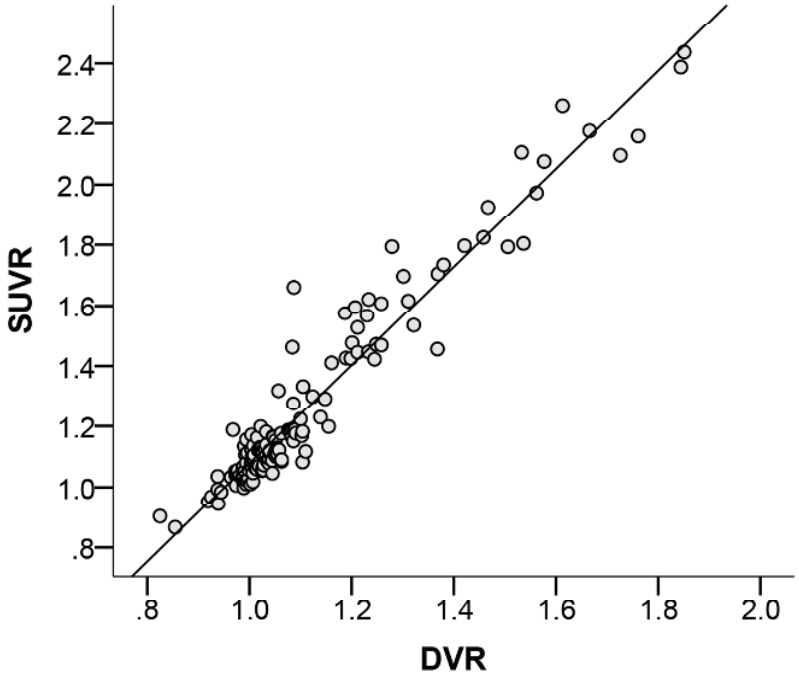

Figure legend: Shown is the correlation between the DVR and the SUVR values of the cognitively normal older adults included in the threshold derivation study (r^2=0.93).

Threshold validation

Supplementary Fig. 3

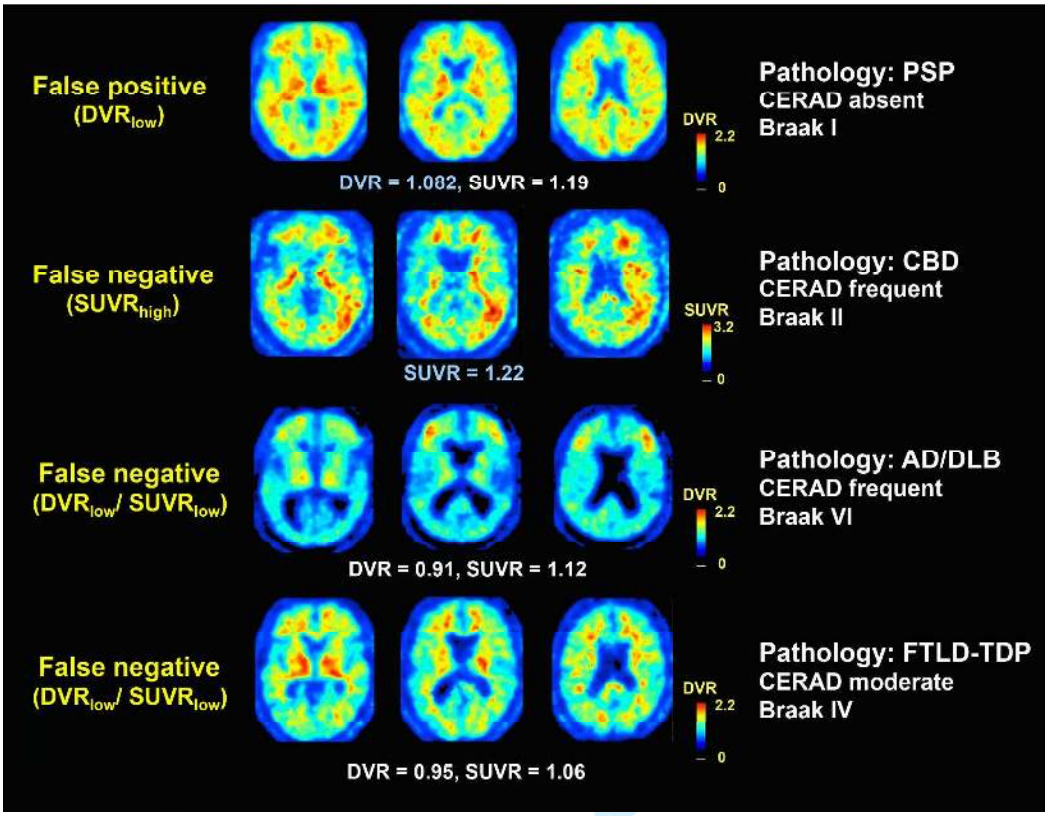

Figure legend: Examples of misclassified PIB-PET scans. The thresholds that misclassified the scans are listed in the left column. PIB-PET trans-axial slices are presented in neurological orientation in the middle column, with SUVR and DVR values listed below each scan. Primary neuroptahological diagnosis, CERAD and Braak scores are shown in the right column. See text for discussion.

PSP = progressive supranuclear palsy; CBD = corticobasal degeneration; AD = Alzheimer's disease; DLB = diffuse Lewy body disease; FTLD-TDP = frontotemporal lobar degeneration, TDP-43 positive; L = left; R = right.

Supplementary Fig. 4

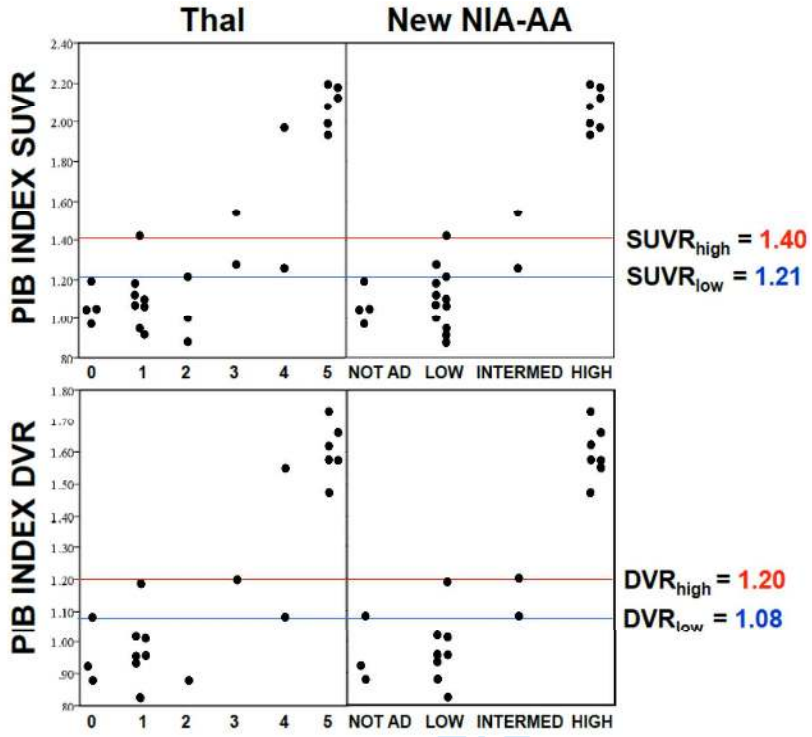

Figure legend: Scatterplots of PIB index SUVR (top) and DVR (bottom) versus Thal and the new HIA-AA rating criteria. Low thresholds are signified by blue lines and high thresholds by red lines.

Supplementary Fig. 5

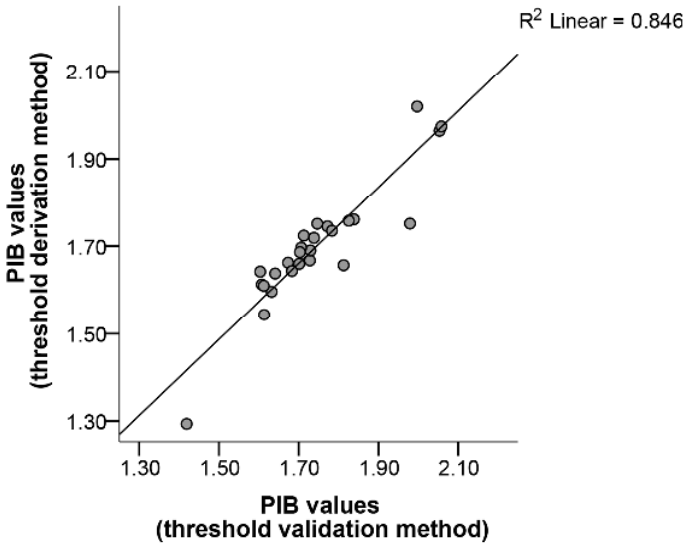

Figure legend: Shown is the correlation between the PIB index of 26 AD subjects that passed the FreeSurfer pipeline and that were further processed in template space using the threshold validation method. DVR values are presented.

REFERENCES

DeCarli C, Reed BR, Jagust W, Martinez O, Ortega M, Mungas D. Brain behavior relationships among African Americans, whites, and Hispanics. Alzheimer Dis Assoc Disord. 2008;22(4):382-91.

Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage. 2006;31(3):968-80.

Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. Journal of psychiatric research. 1975;12(3):189-98.

Hyman BT, Phelps CH, Beach TG, Bigio EH, Cairns NJ, Carrillo MC, et al. National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2012;8(1):1-13.

Marchant NL, Reed BR, DeCarli CS, Madison CM, Weiner MW, Chui HC, et al. Cerebrovascular disease, beta-amyloid, and cognition in aging. Neurobiology of aging. 2012;33(5):1006 e25-36.

Mathis CA, Wang Y, Holt DP, Huang GF, Debnath ML, Klunk WE. Synthesis and evaluation of 11C-labeled 6-substituted 2-arylbenzothiazoles as amyloid imaging agents. Journal of medicinal chemistry. 2003;46(13):2740-54.

Mormino EC, Brandel MG, Madison CM, Rabinovici GD, Marks S, Baker SL, et al. Not quite PIB-positive, not quite PIB-negative: Slight PIB elevations in elderly normal control subjects are biologically relevant. Neuroimage. 2011.

Mormino EC, Brandel MG, Madison CM, Rabinovici GD, Marks S, Baker SL, et al. Not quite PIB-positive, not quite PIB-negative: slight PIB elevations in elderly normal control subjects are biologically relevant. NeuroImage. 2012;59(2):1152-60.

Mueller SG, Laxer KD, Barakos J, Cheong I, Garcia P, Weiner MW. Widespread neocortical abnormalities in temporal lobe epilepsy with and without mesial sclerosis. Neuroimage. 2009;46(2):353-9.

Oh H, Madison C, Villeneuve S, Markley C, Jagust WJ. Association of Gray Matter Atrophy with Age, beta-Amyloid, and Cognition in Aging. Cerebral cortex. 2013.

Rabinovici GD, Furst AJ, Alkalay A, Racine CA, O'Neil JP, Janabi M, et al. Increased metabolic vulnerability in early-onset Alzheimer's disease is not related to amyloid burden. Brain : a journal of neurology. 2010;133(Pt 2):512-28.

Rabinovici GD, Rosen HJ, Alkalay A, Kornak J, Furst AJ, Agarwal N, et al. Amyloid vs FDG-PET in the differential diagnosis of AD and FTLD. Neurology. 2011;77(23):2034-42.

Reed B, Villeneuve S, Mack W, Decarli C, Chui HC, Jagust W. Associations between serum cholesterol levels and cerebral amyloidosis. JAMA Neurol. 2014;71(2):195-200.

Rosen HJ, Gorno-Tempini ML, Goldman WP, Perry RJ, Schuff N, Weiner M, et al. Patterns of brain atrophy in frontotemporal dementia and semantic dementia. Neurology. 2002;58(2):198-208.

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage. 2002;15(1):273-89.

Villeneuve S, Reed BR, Madison CM, Wirth M, Marchant NL, Kriger S, et al. Vascular risk and Abeta interact to reduce cortical thickness in AD vulnerable brain regions. Neurology. 2014;83(1):40-7.

Villeneuve S, Reed BR, Wirth M, Haase CM, Madison CM, Ayakta N, et al. Cortical thickness mediates the effect of beta-amyloid on episodic memory. Neurology. 2014;82(9):761-7.

Wirth M, Haase CM, Villeneuve S, Vogel J, Jagust WJ. Neuroprotective pathways: lifestyle activity, brain pathology, and cognition in cognitively normal older adults. Neurobiology of aging. 2014.

Wirth M, Villeneuve S, Haase CM, Madison CM, Oh H, Landau SM, et al. Associations between Alzheimer disease biomarkers, neurodegeneration, and cognition in cognitively normal older people. JAMA Neurol. 2013;70(12):1512-9.

Zhou J, Gennatas ED, Kramer JH, Miller BL, Seeley WW. Predicting regional neurodegeneration from the healthy brain functional connectome. Neuron. 2012;73(6):1216-27.