Title

Multisite study of the relationships between antemortem [11C]PIB-PET Centiloid values and postmortem measures of Alzheimer's disease neuropathology

Journal

Alzheimer's & Dementia, 15(2)

ISSN

1552-5260

Authors

La Joie, Renaud

Ayakta, Nagehan

Seeley, William W

et al.

Publication Date

2019-02-01

DOI

10.1016/j.jalz.2018.09.001

Peer reviewed

HHS Public Access

Author manuscript

Alzheimers Dement. Author manuscript; available in PMC 2020 February 01.

Published in final edited form as:

Alzheimers Dement. 2019 February ; 15(2): 205–216. doi:10.1016/j.jalz.2018.09.001.

Multi-site study of the relationships between *ante mortem* [^{11}C]PIB-PET Centiloid values and *post mortem* measures of Alzheimer's disease neuropathology.

Renaud La Joie[*,1], **Nagehan Ayakta**[*,1,2], **William W. Seeley**[1], **Ewa Borys**[3], **Adam L. Boxer**[1], **Charles DeCarli**[4], **Vincent Doré**[5], **Lea T. Grinberg**[1], **Eric Huang**[1], **Ji-Hye Hwang**[1], **Milos D. Ikonomovic**[6,7], **Clifford Jack Jr**[8], **William J. Jagust**[2], **Lee-Way Jin**[9], **William E. Klunk**[6,7,10], **Julia Kofler**[11], **Orit H. Lesman-Segev**[1], **Samuel N. Lockhart**[2,12], **Val J. Lowe**[13], **Colin L. Masters**[14], **Chester A. Mathis**[15], **Catriona L McLean**[16], **Bruce L. Miller**[1], **Daniel Mungas**[4], **James P. O'Neil**[2,17], **John M. Olichney**[4], **Joseph E. Parisi**[18,19], **Ronald C. Petersen**[19], **Howard J. Rosen**[1], **Christopher C. Rowe**[5], **Salvatore Spina**[1], **Prashanthi Vemuri**[8], **Victor L. Villemagne**[5,14], **Melissa E. Murray**[20], and **Gil D. Rabinovici**[1,2]

[1]Memory & Aging Center, Department of Neurology, University of California, San Francisco, CA, USA

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

[3]Department of Pathology, Stritch School of Medicine, Loyola University, Maywood, IL, USA

[4]Department of Neurology, University of California, Davis, CA, USA

[5]Department of Molecular Imaging & Therapy, Centre for PET, Austin Health, Heldelberg, Victoria, Australia

[6]Department of Neurology, University of Pittsburgh, PA, USA

[7]Department of Psychiatry, University of Pittsburgh, PA, USA

[8]Department of Radiology, Mayo Clinic, Rochester, MN, USA

[9]Alzheimer's Disease Center, Department of Pathology, University of California Davis, CA, USA

[10]Alzheimer's Disease Research Center, University of Pittsburgh, PA, USA

[11]Department of Pathology, University of Pittsburgh, Pennsylvania, USA

[12]Department of Internal Medicine, Division of Gerontology and Geriatric Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA

Corresponding author: Dr Renaud La Joie, Memory and Aging Center, Department of Neurology, 675 Nelson Rising Ln, Ste 190, San Francisco, CA 94158, Renaud.lajoie@ucsf.edu.

[*]should be regarded as joint first authors

Disclosures

GE Healthcare holds a license agreement with the University of Pittsburgh based on the technology described in this manuscript. Drs. Klunk and Mathis are co-inventors of PIB and, as such, have a financial interest in this license agreement. GE Healthcare provided no grant support for this study and had no role in the design or interpretation of results or preparation of this manuscript.

[13]Department of Nuclear Medicine, Mayo Clinic, Rochester, MN, USA

[14]The Florey Institute, The University of Melbourne, Melbourne, Victoria, Australia

[15]Department of Radiology, University of Pittsburgh, PA, USA

[16]Department of Anatomical Pathology, Alfred Hospital, Melbourne, Australia

[17]Biomedical Isotope Facility, MBIB Division, Lawrence Berkeley National Laboratory, CA, USA

[18]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA

[19]Department of Neurology, Mayo Clinic, Rochester, MN, USA

[20]Department of Neuroscience, Mayo Clinic, Jacksonville, FL, USA

Abstract

INTRODUCTION: We sought to establish the relationships between standard *post mortem* measures of AD neuropathology and *ante mortem* [^{11}C]PIB-PET analyzed with the Centiloid (CL) method, a standardized scale for Aβ-PET quantification.

METHODS: Four centers contributed 179 participants encompassing a broad range of clinical diagnoses, PET data and autopsy findings.

RESULTS: CL values increased with each CERAD neuritic plaque score increment (median −3 CL for no plaques, 92 CL for frequent plaques) and non-linearly with Thal Aβ phases (increases were detected starting at phase 2) with overlap between scores/phases. Findings were comparable across sites and when restricted to 56 patients who died within 2 years of PET. A threshold of 12.2 CL detected CERAD moderate-to-frequent neuritic plaques (AUC=0.910, Sensitivity=89.2%, Specificity=86.4%) while 24.4 CL identified intermediate-to-high AD neuropathological changes (AUC=0.894, Sensitivity=84.1%, Specificity=87.9%).

DISCUSSION: Our study demonstrated the robustness of a multi-site Centiloid [^{11}C]PIBPET study and established a range of pathology-based CL thresholds.

Keywords

beta-amyloid; positron emission tomography; centiloid; CERAD; Thal; Alzheimer's disease neuropathologic changes; neuropathology; harmonization; threshold; Pittsburgh Compound-B

1. Introduction.

The development of Positron Emission Tomography (PET) radiotracers with high affinity and specificity to aggregated Aβ pathology [1] was a milestone in the study of Alzheimer's disease (AD). However, the heterogeneity of Aβ-PET methods and lack of standardization have hampered the field by limiting between-study comparability and data sharing. Though multi-tracer studies suggest that quantification of Aβ with different radiotracers yields highly correlated results, the tracers differ in the recommended reference and target regions used to derive Standardized Uptake Value Ratios (SUVR) values and in the dynamic range of SUVRs in different clinical populations [2,3]. In studies applying the same radiotracer, heterogeneity arises from differences in data acquisition (scanner properties, acquisition

time window), reconstruction parameters, quantitative methods, and image preprocessing pipelines (selection of reference and target regions, use of Magnetic Resonance Imaging (MRI), preprocessing in native or template space, etc.). Finally, different approaches are used to define thresholds of AβPET positivity, i.e. to determine whether the scan shows evidence for Aβ deposition [4–6].

The Centiloid project was designed to address these issues by proposing a standardized PET processing pipeline and a method to transform resulting PET binding metrics into a common unit called "Centiloid" (CL) [7]. The Centiloid scale is anchored at 0 and 100 CL, with 0 CL representing a definitively Aβ-negative brain (originally calculated as the average value of a group of healthy subjects below age 45), and 100 CL reflecting the average signal observed in patients with typical mild-to-moderate AD dementia [8]. This harmonized method (originally designed with [^{11}C]PIB but applicable to other tracers after calibration [9,10]), has great potential to produce cohesive and comparable research results from disparate labs across the world and stimulate collaborations and data sharing.

Proper interpretation of Aβ-PET findings requires a clear understanding of the relationships between PET data and underlying Aβ neuropathology. One important application of the Centiloid method would be to provide standardized and generalizable cutoffs for Aβ-PET "positivity" based on post-mortem data. Previous studies have shown strong correlations between Aβ-PET positivity, as determined by visual interpretation or quantification, and multiple indices of Aβ neuropathology [11–16]. However, these investigations were generally small because of the difficulty of gathering large groups of patients with imaging and pathology data, and each study measured PET binding using laboratory-specific pipelines and units, limiting comparability across studies.

The goal of the present study was twofold. First, we aimed to assess the feasibility of a retrospective multi-site [^{11}C]PIB-PET study using the Centiloid approach. Second, we aimed to investigate the relationships between [^{11}C]PIB-PET imaging, as measured in CL, and standard measures of AD neuropathology, and to determine CL thresholds grounded in neuropathological standards.

2. Methods

2.1. Study design & overview

The first step of the study (Figure 1A) involved implementing the Centiloid standard pipeline and reproducing the results derived from the original dataset to calibrate the scale at the University of California, San Francisco (UCSF). This was performed using the original Centiloid dataset [7] downloaded from the Global Alzheimer's Association Interactive Network website (www.gaain.org), as described in the Supplementary methods.

We then gathered and analyzed *ante mortem* [^{11}C]PIB-PET images from patients who also had available autopsy reports (Figure 1B). Five academic sites contributed: UCSF; University of California Davis, UCD (UCSF and UCD participants were lumped as they all underwent PET at the Lawrence Berkeley National Laboratory); University of Pittsburgh, UPitt; the Mayo Clinic Rochester, Minnesota; and the Australian Imaging Biomarker and

Lifestyle (AIBL) study. [^{11}C]PIB-PET images were preprocessed using the Centiloid standard pipeline (to calculate SUVR values and convert them to CL) at the Mayo Clinic for the subset of Mayo subjects [6], and at UCSF for all other sites.

Statistical analyses were conducted to evaluate relationships between resulting CL values and three neuropathologic scales. We first considered the two indices of Aβ neuropathology: Consortium to Establish a Registry for Alzheimer's Disease (CERAD) score [17] and Thal phase [18]. Neither of these scales are meant to be quantitative measures of total Aβ load, and instead reflect different aspects of Aβ pathology. CERAD score is determined on a 4-point semi-quantitative scale reflecting the maximal neuritic plaque density observed in selected neocortical areas, while Thal is a 6-point scalecapturing the progressive stereotypic topography of Aβ (neuritic and diffuse) deposits. In addition to these Aβ-centric scales, we studied AD Neuropathologic Change (ADNC) levels, which combine information from Aβ (CERAD and Thal) and tau pathology (Braak staging [19]) into a global four-point summary scale recommended for AD neuropathologic diagnosis [20].

2.2. PET-Autopsy cohort

179 Individuals with *ante mortem* [^{11}C]PIB-PET, MRI and autopsy reports were included. This sample partly overlaps with previous papers published by each site [11,13,21–24] and included 22 cognitively normal older adults, 27 patients with Mild Cognitive Impairment, 63 with AD dementia, and 67 with non-AD dementia, who died 3.3 years after [^{11}C]PIB-PET on average. Participants' characteristics varied across sites: UCSF/UCD patients were the youngest and included a majority (62%) of cases with nonAD syndromes (primarily in the frontotemporal dementia spectrum) while most UPitt patients had a clinical diagnosis of AD (59%). Mayo participants were the oldest, and evenly distributed across diagnoses (Table 1, and Supplementary Tables 1–3 for details on clinical and neuropathologic diagnoses, respectively).

The current study being a retrospective collaborative effort, imaging and neuropathologic data were acquired by each site following their own procedures, as described in the Supplementary methods. Briefly, PET data were acquired 40-to-60 min (Mayo) or 50-to-70 min post injection (all other centers) with site-specific scanners (PET or PET-CT). Various methods were used for attenuation correction (e.g. using an external radioactive positron-emitting source or a low dose CT scan) and PET reconstruction (e.g. filtered back-projection or ordered subset expectation maximization); see center-by-center details in the Supplementary methods. T1-weighted MRIs were acquired using various 1.5, 3, or 4T scanners.

2.3. Statistical analyses

Relationships between the three neuropathologic scales and CL values were first assessed using Spearman correlations; Mann-Whitney tests were also run to conduct pairwise comparisons for each increment of the scales.

Receiver Operating Characteristic (ROC) analyses were conducted to derive pathology-based CL thresholds. Our first contrast of interest was based on discriminating between "none-to-sparse" and "moderate-to-frequent" neuritic plaques as scored by CERAD, similar

to previous studies on FDA-approved [^{18}F]-labelled tracers [14,25]. This contrast corresponds to the difference between C scores of 0–1 versus 2–3 according to the ABC neuropathologic score used for AD neuropathologic diagnosis [20].

In addition, we conducted an equivalent analysis based on Thal phase, contrasting Phases 0-to-2 and 3-to-5 (equivalent to A scores of 0–1 versus 2–3 in the ABC framework [20]). Lastly, we aimed to distinguish “intermediate-to-high” and “none-to-low” ADNC levels, as the former “should be considered adequate explanation of cognitive impairment or dementia” [20]. It is notable that intermediate-to-high ADNC levels require Braak stage ≥III and that, although [^{11}C]PIB-PET does not measure tau pathology, higher Braak stages are also usually associated with higher Aβ levels [13], suggesting that higher CL values, corresponding to higher neocortical Aβ burden, may be expected in patients in advanced versus early Braak stages.

Additional ROC analyses were conducted to explore discrimination at each increment of each of the three neuropathologic scales. The Area Under the Curve (AUC) and exact binomial 95%CI were computed and optimal thresholds were determined based on Youden’s index.

3. Results

3.1. CL values in the autopsy cohort

CL values ranged from −26 to 167 CL and showed a bi-modal distribution, with a large number of subjects centered around either 0 or 100 CL (Figure 1B). CL values varied across clinical diagnosis (Kruskal-Wallis: $p<0.001$, $\eta^2_H=0.302$, Figure 1B); patients with a clinical diagnosis of AD had a median value of 99 CL although values spanned the full range of the scale.

3.2. Relationships between CL values and neuropathologic measures

Figure 2A illustrates the distribution of CERAD scores and Thal phases, showing that our cohort encompassed a broad range of Aβ pathology levels. CL values increased with CERAD scores ($\rho=0.716$, $p<0.001$), and significant differences were found for each increment (p’s<0.02, Figure 2A). CL values increased with Thal phases ($\rho=0.768$, $p<0.001$) but pairwise comparisons showed detectable PET signal increase beginning at phase 2 (versus Phase 0: $p=0.01$, versus Phase 1: $p=0.066$; see Figure 2A). Regression analyses confirmed the nonlinear relationship between Thal phase and [^{11}C]PIB-PET: a model including a quadratic term ($Thal^2$) explained PET signal ($R^2=59.2\%$) better than a simple linear model ($R^2=57.2\%$; $\Delta R^2=1.9\%$, $p=0.007$). When entering CERAD, Thal and $Thal^2$ in a stepwise regression, the final model included CERAD and $Thal^2$ (both $p\leq0.001$) and explained 61.4% of total CL variance.

Figure 2B shows the relationships between CERAD, ADNC and CL values. CERAD and ADNC scales were correlated but not synonymous in categorizing patients as having low or high pathology, particularly when CERAD moderate-to-frequent cases had low Braak stage and therefore fell into the “low” ADNC level. [^{11}C]PIB-PET binding was correlated with

ADNC levels (ρ=0.736, p<0.001), and CL values increased with every ADNC increment (p's<0.02).

Supplementary Table 3 describes the relationships between the three neuropathologic scales and CL values in detail.

3.3. ROC analysis and CL thresholds

The results of ROC analyses are displayed in Table 2; the three contrasts of interest are discussed below.

3.3.1. CERAD: none-to-sparse versus moderate-to-frequent scores—

[^{11}C]PIB-PET distinguished individuals with moderate-to-frequent (n=120) versus none-to-sparse (n=59) CERAD scores with high accuracy (AUC=0.910, 95%CI[0.858, 0.948] see Supplementary Figure 1). A threshold of 12.2 CL separated groups with a sensitivity of 89.2% and a specificity of 86.4%, positive/negative predictive values of 93.0/79.7%, and positive/negative likelihood ratios of 6.58/0.13. It should be noted that the cutoff to distinguish none and sparse-to-frequent groups also was 12.2 CL (83.3% sensitivity, 100% specificity).

Table 3 describes the 21 individuals that were misclassified applying this threshold. Eight participants were "false positive" (i.e. CERAD none/sparse and CL>12.2, between 20 and 88), although all had detectable Aβ neuropathology (sparse neuritic plaques for all 8, with Thal phases between 2 and 4). Half (n=4) had a neuropathologic diagnosis of Lewy Body Disease, and half (n=4) had intermediate ADNC levels. The 13 "false negative" participants (i.e. CERAD moderate/frequent and CL <12.2) had CL values between −19.7 and 8.1. Three cases had maximal CERAD and Thal scores, including two patients with autosomal dominant AD (ADAD); visual inspection of these two cases (Supplementary Figure 2), showed that the low CL values were due to contamination of the cerebellar reference region rather than the absence of cortical binding. Removing the ADAD patients (n=3, including these two negative patients and a case with 168 CL) did not impact threshold calculation (Supplementary Table 4).

3.3.2. Thal: 0-to-2 versus 3-to-5—The threshold estimated to detect Thal phase 3-to-5 versus 0-to-2 was 23.5 CL (85.8% sensitivity, 95.9% specificity, 98.0%/74.6% positive/negative predictive values, and 21.03/0.15 positive/negative likelihood ratios).

3.3.3. ADNC: no-to-low versus intermediate-to-high levels—The threshold estimated to detect intermediate-to-high ADNC levels was 24.4 CL (84.1% sensitivity, 87.9% specificity, 92.2%/76.3% positive/negative predictive values, and 6.94/0.18 positive/negative likelihood ratios).

3.4. Complementary analyses

Additional analyses were conducted to extend the previous findings and to test their robustness.

3.4.1. Stability across centers and PET-to-death interval—Significant CL/CERAD associations were found in the three main centers (Figure 3A, see Supplementary Figure 3 for comparable analyses with Thal phases).

CL/CERAD correlations were unchanged when repeating the analyses in the subset of 56 participants who died within 2 years of PET (Figure 3B, see Supplementary Figure 4 for Thal analyses and other PET-to-death intervals). The ROC analysis was repeated to distinguish the 39 moderate-to-frequent from the 24 none-to-sparse CERAD cases with PET-death interval ≤2 years, resulting in AUC=0.866 [0.755, 0.939] and a threshold of 13 CL.

3.4.2. Relationship between CL and visual reads—Three clinicians (O.H.L., H.J.R., W.J.J.) read the 73 UCSF/UCD [^{11}C]PIB-PET SUVR images as negative or positive for cortical binding [26], blind to clinical information and PET quantification (Figure 3C). Among the 27 cases read as negative by all three raters, 26 had CL values below 12.2, and one case was slightly above threshold (13.8 CL; neuropathology showed CERAD frequent and Thal phase 3). Out of the 30 cases that were read as positive by consensus, all CL values were above 12.2 CL (minimum value =12.3 CL), and 27 were above 24.4 CL. Among the 16 cases with between-rater disagreement in visual interpretation, three (all with frequent CERAD scores) had CL>12.2 (14.9, 24.4, and 33.2 CL) and were read as positive by 2/3 readers. Overall, the six cases with CL values between 12.2 and 24.4, all of whom had CERAD moderate-to-frequent neuritic plaques, received positive reads 13 of 18 (6 cases x 3 readers) possible times (72%). The 30 cases above 24 CL received positive reads 89 of 90 possible times (99%); the only negative read coming from a case with 33.2 CL.

3.4.3. Participants with maximal CERAD score (frequent plaques)—CL values were extremely variable in individuals within a given CERAD score (Figure 2A), independent of center and PET-to-death interval (Figure 3). To test whether this variability reflects actual variability in Aβ load not captured by the CERAD scale, we assessed the correlation between Thal phases and CL values in individuals with maximal (frequent) CERAD score. CL values correlated with Thal stage in this sample (ρ=0.335, p=0.003) and in the subset of the sample with PET-to-death interval ≤2 years (ρ=0.654, p<0.001; Figure 3D).

4. Discussion

In this investigation of 179 individuals, we assessed relationships between AD neuropathology and [^{11}C]PIB-PET imaging quantified with the Centiloid scale. We aimed to better characterize the link between [^{11}C]PIB-PET imaging and its underlying target, and to test the robustness of the Centiloid approach to analyze multisite [^{11}C]PIB-PET data.

4.1. [^{11}C]PIB-PET-pathology relationships

CERAD score and Thal phase both contributed to global [^{11}C]PIB-PET signal, and together accounted for ~60% of the CL variance. However, this does not signify that 40% of the PET signal is unrelated to Aβ pathology: CERAD and Thal measures are ordinal, non-linear scales that do not necessarily reflect the total Aβ burden. For instance, a focal area of high

density of neuritic plaques in one sampled brain region is sufficient to classify as CERAD frequent [17] even if the rest of the cortex shows no or sparse plaques. Similarly, Thal phases represent the topography of neuritic and diffuse Aβ deposits in different regions, but do not directly reflect the quantitative Aβ burden in these regions. More quantitative, continuous measures of Aβ burden – e.g. a measure of Aβ concentration by ELISA, mass spectroscopy, or a histochemical measure that included both fluorescent intensity and surface area of Aβ plaques, each performed over a broad area similar to the Centiloid target region - would be expected to better correlate with global PET values.

The independent contribution of Thal phase is consistent with a previous study based on a partly overlapping sample [13], and could reflect [^{11}C]PIB binding to fibrillary components of non-neuritic Aβ deposits (that are sometimes “diffuse”) that are not captured by the CERAD score [11–13,22,27]. Alternatively, this relationship might be confounded by the previously highlighted correlation between Thal phases and cortical neuritic plaque density [13], reflecting the fact that Aβ pathology progression is characterized by both increasing plaque density and appearance in new areas.

Altogether, the present data indicates not only that [^{11}C]PIB-PET is able to detect the presence of Aβ pathology, but also that the intensity of PET signal is associated with more severe pathology stages, which is of high importance regarding the use of Aβ-PET imaging for monitoring disease progression and therapeutic effects in clinical trials.

Interestingly, the associations between neuropathologic measures of Aβ and [^{11}C]PIBPET were not strongly impacted by PET-to-death interval, potentially because of the slow aggregation of Aβ over time. In addition, our sample included a large group of patients with non-AD conditions, who tended to harbor no or minimal Aβ neuropathology and were at a lower risk of developing Aβ pathology within the subsequent years because of their relatively young age. The sample also included a relatively large number of AD patients who may have been near their maximal level of Aβ pathology or may have even started to show decreasing [^{11}C]PIB-PET signal [28,29].

4.2. Implementation of the Centiloid method with [^{11}C]PIB-PET data

4.2.1. Robustness of the Centiloid scale with [^{11}C]PIB-PET data—

The 0 and 100 CL anchor points are designed to represent the average [^{11}C]PIB-PET binding associated with no Aβ pathology (originally based on a group of young controls), and the typical pathology observed in patients with mild-to-moderate AD dementia. Our autopsy cohort supports the validity of these values, when looking at individuals with no or maximal levels of Aβ neuropathology on CERAD and Thal. Importantly, these patterns were observed with minor variations across centers (Figure 3A, Supplementary Figure 3), despite heterogeneity in clinical characteristics, scanners and image reconstruction methods.

4.2.2. Pathology-based thresholds—

Our analyses identified a threshold of 12.2 CL to detect moderate-to-frequent CERAD scores with 89% sensitivity and 86% specificity. This threshold also corresponded to the smallest CL value observed among scans that were read as positive by consensus of three independent visual readers. Although the six cases in the 12.2–24.4 CL range were read as positive most (72%) of the time, only 3/6 cases were

read as positive by all three raters - compared to 29/30 cases >24.4 CL read as positive by all raters. This indicates that, as expected, early signs of positivity might be less reliably identified by visual read.

Importantly, all “false positive” cases based on the 12.2 CL threshold had some level of Aβ neuropathology (CERAD sparse and Thal phase between 2 and 4, see Figure 2D and Table 3), supporting a high degree of tracer specificity: all 115 individuals with CL>12.2 had at least some neuritic plaques at autopsy. However, sensitivity was imperfect, with some individuals showing significant Aβ neuropathology at autopsy and CL<12.2, in line with studies suggesting that low levels of Aβ neuropathology detectable at autopsy are not sufficient to produce a positive PET signal [24]. Finally, technical considerations such as the suboptimal extraction from a large cortical volume of interest might also limit sensitivity (e.g. early focal binding would be diluted in a global cortical signal extraction); more restricted regional information could help detect earlier and more focal Aβ deposition [5,21,30,31].

As discussed above, [^{11}C]PIB CL values represent a continuous measure of cortical fibrillary Aβ pathology, and are not directly comparable to neuropathological scales that measure the spatial extent of Aβ deposits (Thal), maximal density of neuritic plaques (CERAD) or integration of Thal and CERAD scales with Braak staging of neurofibrillary tau pathology (ADNC). Nevertheless, such comparisons allow us to anchor CL measurements - and thresholds - to clinically interpretable measures of neuropathology. Our data suggest that the earliest detectable PIB signal occurs at approximately 12 CL, which in our study robustly distinguished CERAD none-sparse vs moderate-frequent scores and Thal phases 2–5 from 0–1. This measurement appeared independent of PET-toautopsy interval, and usually corresponded to positive visual interpretations by blinded readers. That said, CL values of 12 are lower than published PIB-CL thresholds [6,32], and may represent the limit of reliable detection of signal to noise with PIB. This threshold needs to be validated in independent cohorts. In addition, this threshold may not perform as well with ^{18}F radiotracers [9,10] that show greater variance or “noise” in Aβ-negative individuals because a threshold must account for both the ground-truth of which cases are true-positives as well as the noise inherent in the measure for which the threshold is determined. Even with PIB, higher thresholds may be appropriate in settings in which specificity should be emphasized, such as clinical trial eligibility. A more conservative threshold of 24 CL would be appropriate for identifying clinically meaningful Aβ burden in cognitively impaired patients, as this threshold best discriminated none-low from intermediate-high overall ADNC as well as Thal phases 0–2 from 3–5. Thresholds in between 12–24, such as the threshold of 19 CL derived from a “reliable worsening cut point” technique based on longitudinal [^{11}C]PET-PET data [6], may represent a reasonable compromise between the more liberal and conservative thresholds presented in this study.

4.2.3. Limitations—Our study also highlighted a number of potential limitations of applying the standard Centiloid towards processing Aβ-PET imaging. First, although the presence of negative CL values is expected (the average of young controls being 0, half of Aβ-negative individuals should have slightly negative CL values), we observed multiple individuals with values beyond what could be explained by normal variations around 0 CL.

The lowest value from the calibrating young control group was −6.8, while we observed 20 participants with lower values (Figure 1B). Further analyses of these highly negative individuals (Supplementary Table 5) showed that the vast majority were controls in the older age range or cognitively impaired patients with low or non-AD pathology, suggesting that the very negative CL values might be due to brain atrophy in participants with presumably no specific [^{11}C]PIB binding. This hypothesis is reinforced by the high prevalence of Pick's disease in this group (Supplementary Table 5), which is usually associated with severe atrophy [33–35]. Our findings highlight that any CL value should be interpreted with caution in individuals with brain atrophy. The implementation of partial volume effect correction methods might be useful in these cases [36–39] although alternative pipelines would have to be validated following Centiloid requirements.

Second, outlier analyses identified two individuals with ADAD with [^{11}C]PIB binding to both cerebral and cerebellar cortices (the latter being included in the reference region in the standard Centiloid pipeline) resulting in subthreshold CL values. This finding is consistent with previous studies suggesting that other reference regions may be preferred while studying ADAD [40–42] due to the higher burden of Aβ pathology [43,44], including amyloid angiopathy [45], in the cerebellum. Alternative methods proposed by the Centiloid project [46], for instance using the pons as a reference region, may then be preferred. It is important to note that, over and above those two cases, all 54 individuals at Thal phase 5 had - by definition - some cerebellar Aβ pathology. Yet, the structure (mostly diffuse) and/or low density of these Aβ deposits did not induce sufficient cerebellar [^{11}C]PIB binding to generate false negative CL values.

Finally, it should be highlighted that the current study only included [^{11}C]PIB-PET data and that future studies are needed to evaluate the robustness of the Centiloid approach to other radiotracers, and most importantly to combine data acquired using various radiotracers.

4.3. Conclusions

In a large cohort encompassing a broad range of clinical and neuropathologic characteristics, we showed the feasibility and robustness of a multi-site CL-based [^{11}C]PIB-PET project and derived pathology-based thresholds for PET positivity. These results contribute to a better understanding of the pathology underlying imaging findings, and should help the field towards better between-laboratory comparability and collaborations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

5. Acknowledgments

Funding

UCSF: National Institute of Health (R01-AG045611, P01-AG1972403, P50-AG023501, R01-AG032306, K24 AG053435), Tau Consortium, Consortium for Frontotemporal Dementia Research, John Douglas French Alzheimer's Foundation, Alzheimer's Association (AARF-16–443577). **UCD**: National Institute of Health (P30 AG010129, P01 AG012435, R01 AG021028, R01 AG031563). **UC Berkeley:** National Institute of Health (R01AG034570). **Mayo**: National Institute of Health (R01 NS097495, U01 AG006786, P50 AG016574/P1, R01 AG011378, R01 AG041851, R01 AG054449); the GHR Foundation, the Alexander Family Alzheimer's Disease

Research Professorship of the Mayo Foundation, Liston Award, Elsie and Marvin Dekelboum Family Foundation, Schuler Foundation. **AIBL**: NHMRC Research Fellowship 1046471, NHMRC Program Grant 1132604. **Pittsburgh:** National Institute of Health (P50 AG005133, RF1 AG025516, P01 AG025204).

Dr. Boxer receives research support from Avid, Biogen, Bristol Myers Squibb, C2N Diagnostics, Cortice Biosciences, Eli Lilly, Forum Pharmaceuticals, Genentech, and TauRx. He has served as a consultant for Asceneuron, Ipierian, Isis, Janssen, and Merck. He has stock/options in Alector and Delos.

Lea T. Grinberg receives research support from Avid Radiopharmaceuticals and Eli Lilly and served as a consultant to GlaxoSmithKline.

Dr. Ikonomovic served as a consultant and received research funding from GE Healthcare.

Dr. Jagust has served as a consultant to BioClinica, Genentech, and Novartis Pharmaceuticals.

Dr. Lowe serves is a consultant for Bayer Schering Pharma and Piramal Imaging and receives research support from GE Healthcare, Siemens Molecular Imaging, AVID Radiopharmaceuticals, the NIH (NIA, NCI), and the MN Partnership for Biotechnology and Medical Genomics.

Dr. Petersen served as a consultant for Roche, Inc, Merck, Inc., Genetech, Inc. and Biogen, Inc.

Dr. Rabinovici receives research support from Avid Radiopharmaceuticals, Eli Lilly, GE Healthcare and Piramal. He has received consulting fees from Genentech, Merck and Roche.

6. References.

[1]. 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. Ann Neurol 2004;55:306–19. doi:10.1002/ana.20009. [PubMed: 14991808]

[2]. 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;41:1398–407. doi:10.1007/s00259-014-2753-3. [PubMed: 24647577]

[3]. Johnson KA, Minoshima S, Bohnen NI, Donohoe KJ, Foster NL, Herscovitch P, et al. Appropriate use criteria for amyloid PET: a report of the Amyloid Imaging Task Force, the Society of Nuclear Medicine and Molecular Imaging, and the Alzheimer's Association. Alzheimers Dement J Alzheimers Assoc 2013;9:e-1–16. doi:10.1016/j.jalz.2013.01.002.

[4]. Villeneuve S, Rabinovici GD, Cohn-Sheehy BI, Madison C, Ayakta N, Ghosh PM, et al. Existing Pittsburgh Compound-B positron emission tomography thresholds are too high: statistical and pathological evaluation. Brain J Neurol 2015;138:2020–33. doi:10.1093/brain/awv112.

[5]. Cohen AD, Mowrey W, Weissfeld LA, Aizenstein HJ, McDade E, Mountz JM, et al. Classification of amyloid-positivity in controls: comparison of visual read and quantitative approaches. NeuroImage 2013;71:207–15. doi:10.1016/j.neuroimage.2013.01.015. [PubMed: 23353602]

[6]. Jack CR, Wiste HJ, Weigand SD, Therneau TM, Lowe VJ, Knopman DS, et al. Defining imaging biomarker cut points for brain aging and Alzheimer's disease. Alzheimers Dement J Alzheimers Assoc 2017;13:205–16. doi:10.1016/j.jalz.2016.08.005.

[7]. Klunk WE, Koeppe RA, Price JC, Benzinger TL, Devous MD, Jagust WJ, et al. The Centiloid Project: standardizing quantitative amyloid plaque estimation by PET. Alzheimers Dement J Alzheimers Assoc 2015;11:1–15.e1–4. doi:10.1016/j.jalz.2014.07.003.

[8]. Braak H, Braak E. Diagnostic criteria for neuropathologic assessment of Alzheimer's disease. Neurobiol Aging 1997;18:S85–88. [PubMed: 9330992]

[9]. Rowe CC, Jones G, Doré V, Pejoska S, Margison L, Mulligan RS, et al. Standardized Expression of 18F-NAV4694 and 11C-PiB β-Amyloid PET Results with the Centiloid Scale. J Nucl Med Off Publ Soc Nucl Med 2016;57:1233–7. doi:10.2967/jnumed.115.171595.

[10]. Rowe CC, Doré V, Jones G, Baxendale D, Mulligan RS, Bullich S, et al. 18FFlorbetaben PET beta-amyloid binding expressed in Centiloids. Eur J Nucl Med Mol Imaging 2017;44:2053–9. doi:10.1007/s00259-017-3749-6. [PubMed: 28643043]

[11]. 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 J Neurol 2008;131:1630–45. doi:10.1093/brain/awn016.

[12]. Cohen AD, Rabinovici GD, Mathis CA, Jagust WJ, Klunk WE, Ikonomovic MD. Using Pittsburgh Compound B for in vivo PET imaging of fibrillar amyloid-beta. Adv Pharmacol San Diego Calif 2012;64:27–81. doi:10.1016/B978-0-12-394816-8.00002-7.

[13]. Murray ME, Lowe VJ, Graff-Radford NR, Liesinger AM, Cannon A, Przybelski SA, et al. Clinicopathologic and 11C-Pittsburgh compound B implications of Thal amyloid phase across the Alzheimer's disease spectrum. Brain J Neurol 2015;138:1370–81. doi:10.1093/brain/awv050.

[14]. 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-β plaques: a prospective cohort study. Lancet Neurol 2012;11:669–78. doi:10.1016/S1474-4422(12)70142-4. [PubMed: 22749065]

[15]. 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;72:287–94. doi:10.1001/jamaneurol.2014.4144. [PubMed: 25622185]

[16]. Salloway S, Gamez JE, Singh U, Sadowsky CH, Villena T, Sabbagh MN, et al. Performance of [18F]flutemetamol amyloid imaging against the neuritic plaque component of CERAD and the current (2012) NIA-AA recommendations for the neuropathologic diagnosis of Alzheimer's disease. Alzheimers Dement Diagn Assess Dis Monit 2017;9:25–34. doi:10.1016/j.dadm.2017.06.001.

[17]. 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;41:479–86. [PubMed: 2011243]

[18]. Thal DR, Rüb U, Orantes M, Braak H. Phases of A beta-deposition in the human brain and its relevance for the development of AD. Neurology 2002;58:1791–800. [PubMed: 12084879]

[19]. Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol (Berl) 1991;82:239–259. [PubMed: 1759558]

[20]. 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. Alzheimers Dement J Alzheimers Assoc 2012;8:1–13. doi:10.1016/j.jalz.2011.10.007.

[21]. Villeneuve S, Rabinovici GD, Cohn-Sheehy BI, Madison C, Ayakta N, Ghosh PM, et al. Existing Pittsburgh Compound-B positron emission tomography thresholds are too high: statistical and pathological evaluation. Brain J Neurol 2015;138:2020–33. doi:10.1093/brain/awv112.

[22]. Seo SW, Ayakta N, Grinberg LT, Villeneuve S, Lehmann M, Reed B, et al. Regional correlations between [(11)C]PIB PET and post-mortem burden of amyloid-beta pathology in a diverse neuropathological cohort. NeuroImage Clin 2017;13:130–7. doi:10.1016/j.nicl.2016.11.008. [PubMed: 27981028]

[23]. Reed B, Villeneuve S, Mack W, DeCarli C, Chui HC, Jagust W. Associations between serum cholesterol levels and cerebral amyloidosis. JAMA Neurol 2014;71:195–200. doi:10.1001/jamaneurol.2013.5390. [PubMed: 24378418]

[24]. 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 (Berl) 2012;123:433–47. doi:10.1007/s00401-012-0943-2. [PubMed: 22271153]

[25]. Sojkova J, Driscoll I, Iacono D, Zhou Y, Codispoti K-E, 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;68:232–40. doi:10.1001/archneurol.2010.357. [PubMed: 21320990]

[26]. 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;77:2034–42. doi:10.1212/WNL.0b013e31823b9c5e. [PubMed: 22131541]

[27]. Kantarci K, Yang C, Schneider JA, Senjem ML, Reyes DA, Lowe VJ, et al. Antemortem amyloid imaging and β-amyloid pathology in a case with dementia with Lewy bodies. Neurobiol Aging 2012;33:878–85. doi:10.1016/j.neurobiolaging.2010.08.007. [PubMed: 20961664]

[28]. Villain N, Chételat G, Grassiot B, Bourgeat P, Jones G, Ellis KA, et al. Regional dynamics of amyloid-β deposition in healthy elderly, mild cognitive impairment and Alzheimer's disease: a voxelwise PiB-PET longitudinal study. Brain J Neurol 2012;135:2126–39. doi:10.1093/brain/aws125.

[29]. Jack CR, Wiste HJ, Lesnick TG, Weigand SD, Knopman DS, Vemuri P, et al. Brain β-amyloid load approaches a plateau. Neurology 2013;80:890–6. doi:10.1212/WNL.0b013e3182840bbe. [PubMed: 23446680]

[30]. Grothe MJ, Barthel H, Sepulcre J, Dyrba M, Sabri O, Teipel SJ, et al. In vivo staging of regional amyloid deposition. Neurology 2017;89:2031–8. doi:10.1212/WNL.0000000000004643. [PubMed: 29046362]

[31]. Palmqvist S, Schöll M, Strandberg O, Mattsson N, Stomrud E, Zetterberg H, et al. Earliest accumulation of β-amyloid occurs within the default-mode network and concurrently affects brain connectivity. Nat Commun 2017;8:1214. doi:10.1038/s41467-017-01150-x. [PubMed: 29089479]

[32]. Leuzy A, Chiotis K, Hasselbalch SG, Rinne JO, de Mendonça A, Otto M, et al. Pittsburgh compound B imaging and cerebrospinal fluid amyloid-β in a multicentre European memory clinic study. Brain J Neurol 2016;139:2540–53. doi:10.1093/brain/aww160.

[33]. Rohrer JD, Lashley T, Schott JM, Warren JE, Mead S, Isaacs AM, et al. Clinical and neuroanatomical signatures of tissue pathology in frontotemporal lobar degeneration. Brain J Neurol 2011;134:2565–81. doi:10.1093/brain/awr198.

[34]. Whitwell JL, Josephs KA, Rossor MN, Stevens JM, Revesz T, Holton JL, et al. Magnetic resonance imaging signatures of tissue pathology in frontotemporal dementia. Arch Neurol 2005;62:1402–8. doi:10.1001/archneur.62.9.1402. [PubMed: 16157747]

[35]. Perry DC, Brown JA, Possin KL, Datta S, Trujillo A, Radke A, et al. Clinicopathological correlations in behavioural variant frontotemporal dementia. Brain J Neurol 2017;140:3329–45. doi:10.1093/brain/awx254.

[36]. Su Y, Blazey TM, Snyder AZ, Raichle ME, Marcus DS, Ances BM, et al. Partial volume correction in quantitative amyloid imaging. NeuroImage 2015;107:55–64. doi:10.1016/j.neuroimage.2014.11.058. [PubMed: 25485714]

[37]. Thomas BA, Erlandsson K, Modat M, Thurfjell L, Vandenberghe R, Ourselin S, et al. The importance of appropriate partial volume correction for PET quantification in Alzheimer's disease. Eur J Nucl Med Mol Imaging 2011;38:1104–19. doi:10.1007/s00259-011-1745-9. [PubMed: 21336694]

[38]. 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 J Neurol 2010;133:512–28. doi:10.1093/brain/awp326.

[39]. Su Y, Blazey TM, Owen CJ, Christensen JJ, Friedrichsen K, Joseph-Mathurin N, et al. Quantitative Amyloid Imaging in Autosomal Dominant Alzheimer's Disease: Results from the DIAN Study Group. PLOS ONE 2016;11:e0152082. doi:10.1371/journal.pone.0152082. [PubMed: 27010959]

[40]. Fleisher AS, Chen K, Quiroz YT, Jakimovich LJ, Gomez MG, Langois CM, et al. Florbetapir PET analysis of amyloid-β deposition in the presenilin 1 E280A autosomal dominant Alzheimer's disease kindred: a cross-sectional study. Lancet Neurol 2012;11. doi:10.1016/S1474-4422(12)70227-2.

[41]. Edison P, Hinz R, Ramlackhansingh A, Thomas J, Gelosa G, Archer HA, et al. Can target-to-pons ratio be used as a reliable method for the analysis of [11C]PIB brain scans? NeuroImage 2012;60:1716–23. doi:10.1016/j.neuroimage.2012.01.099. [PubMed: 22306804]

[42]. Villemagne VL, Ataka S, Mizuno T, Brooks WS, Wada Y, Kondo M, et al. High striatal amyloid beta-peptide deposition across different autosomal Alzheimer disease mutation types. Arch Neurol 2009;66:1537–44. doi:10.1001/archneurol.2009.285. [PubMed: 20008660]

[43]. Duque-Castaño A, Roldán MI, Arango-Viana JC, Arcos-Burgos M, Cubillo H, Lopera F. [Neuropathological findings in early-onset Alzheimer's disease (E280aPS1 mutation)]. Rev Neurol 1999;29:1–6. [PubMed: 10528300]

[44]. Pro JD, Smith CH, Sumi SM. Presenile Alzheimer disease: amyloid plaques in the cerebellum. Neurology 1980;30:820–5. [PubMed: 7191064]

[45]. Mann DM, Pickering-Brown SM, Takeuchi A, Iwatsubo T, Members of the Familial Alzheimer's Disease Pathology Study Group. Amyloid angiopathy and variability in amyloid beta deposition is determined by mutation position in presenilin-1-linked Alzheimer's disease. Am J Pathol 2001;158:2165–75. [PubMed: 11395394]

[46]. Klunk WE, Koeppe RA, Price JC, Benzinger TL, Devous MD, Jagust WJ, et al. The Centiloid Project: standardizing quantitative amyloid plaque estimation by PET. Alzheimers Dement J Alzheimers Assoc 2015;11:1–15.e1–4. doi:10.1016/j.jalz.2014.07.003.

Systematic review.

Authors used PubMed and Google Scholar to review literature on i) relations between amyloid-PET and neuropathology, ii) methods used to derive amyloid-PET thresholds, and iii) use of the Centiloid method to process and quantify amyloid-PET data.

Interpretation.

In a large cohort encompassing a broad range of clinical and neuropathologic characteristics, we showed that standard measures of Aβ neuropathology (CERAD scores and Thal phases) accounted for 60% of global PIB-PET signal variance. The study also demonstrated the feasibility and robustness of the Centiloid method across multiple centers, validation of the 0 and 100 CL anchor points, and derivation of pathologically based thresholds for amyloid positivity.

Future directions.

These results should enable researchers to better understand the neuropathologic underpinning of imaging findings, and help the field towards better betweenlaboratory comparability and collaborations. The set of pathology-based thresholds estimated in the study are available for future research using the Centiloid method.

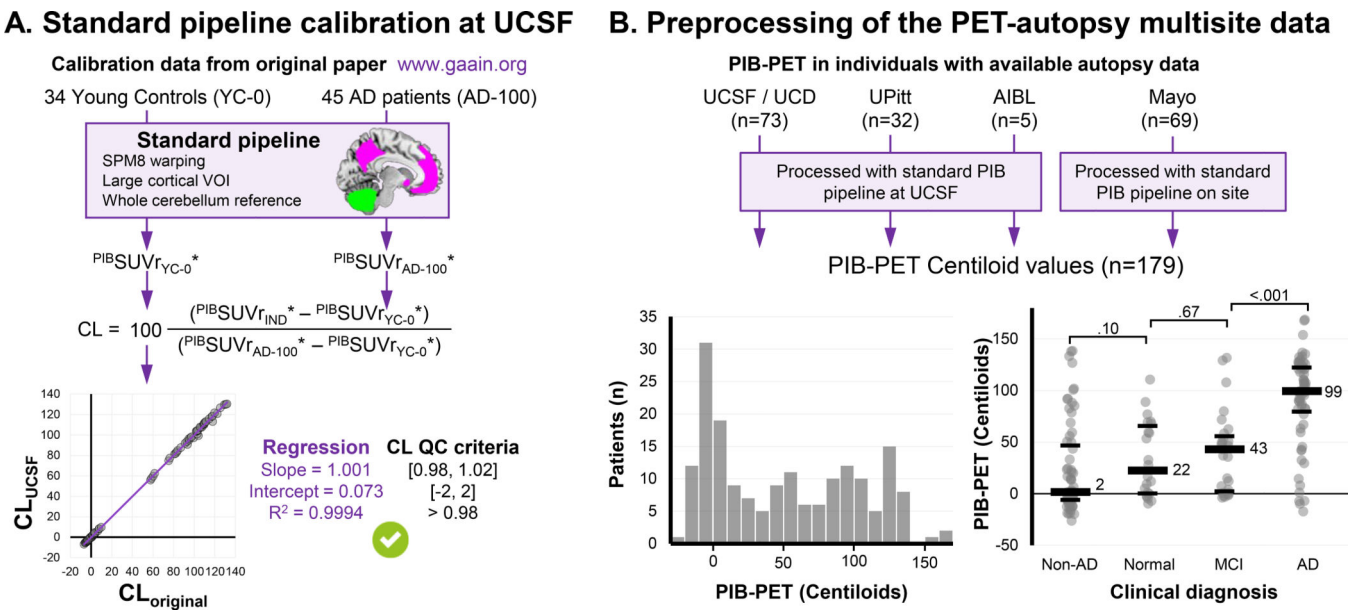

Figure 1. Overview of the present study.

A) PET data from the original article (34 young controls (YC) and 45 patients with AD) were downloaded from the GAAIN website and processed at UCSF following all the Centiloid guidelines (the * indicates that values were calculated on site based on the images downloaded from the website, in compliance with Centiloid nomenclature). PET were warped to template using SPM8, transformed into Standardized Uptake Value Ratios (SUVR) maps using the whole cerebellum as a reference region (shown in green), and average values were extracted from the cortical volume of interest (VOI, shown in purple). The average SUVR of the YC/AD groups were used to define the 0/100 CL points. For all 79 scans, CL values were calculated and compared to the CL values specified in the original paper (scatter plot) to validate the Centiloid method implementation on our site based on the CL quality control (QC) requirements (i.e. regression slope between 0.98 and 1.02, intercept between −2 and 2, R^2 above 0.98).

B) Flow diagram of the original data used in the present study showing an overview of the main analyses. The histogram shows the distribution of PIB-PET CL values in the whole group (n=179, spanning from −26 to 169) and the scatter plot illustrates the distribution of CL values (individual data points, medians, quartiles) according to primary clinical diagnosis at the time of PET. Pairwise group comparisons were conducted using Mann-Whitney tests and corresponding p-values are shown on the plot (for the sake of clarity, we only indicate the results of comparison between contiguous groups, but all the other p's < 0.05).

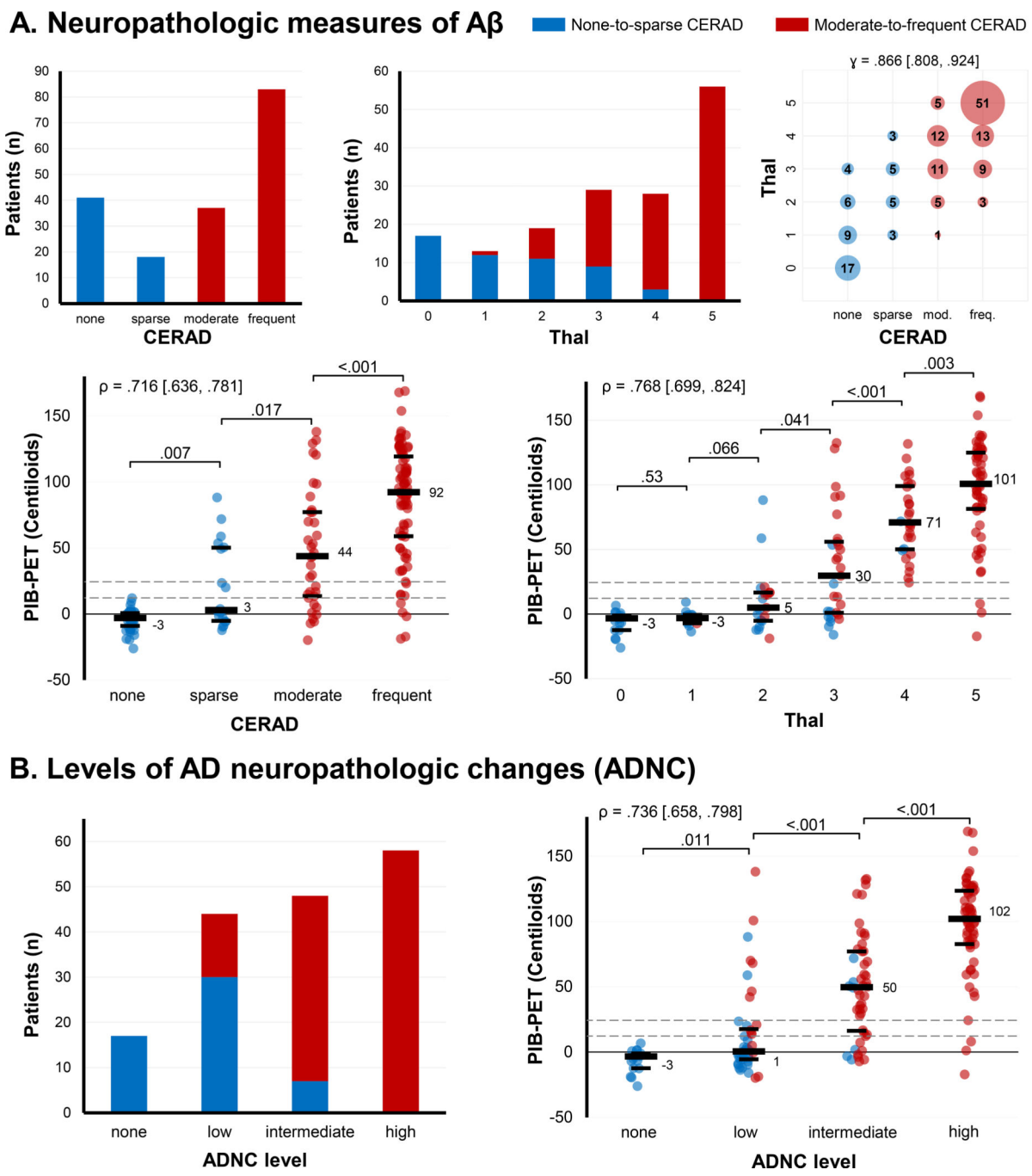

Figure 2. Relationships between CL values and *post mortem* measures of Aβ pathology (A) and levels of AD neuropathologic changes (B).

Bar graphs indicate the distribution of the neuropathologic measures in the whole group, A) CERAD, n=179 and Thal, n=162; B) ADNC, n=167. The bubble plot illustrates the relationships between the two measures of Aβ pathology (correlation between ordinal scales is shown using gamma coefficient and 95%CI). Scatter plots illustrate the distribution of CL values according to neuropathologic measures, showing individual data points (blue: none-to-sparse CERAD, red: moderate-to-frequent CERAD), medians (the actual values are also specified), and quartiles. Pairwise group comparisons were conducted using Mann-Whitney

tests and corresponding p-values are shown on the plot (we only indicate the results of comparison between contiguous groups, but all the other p's < 0.05). Spearman's ρ correlation coefficients [95%CI] are indicated.

Dotted lines illustrate the thresholds identified by the ROC analyses (see text): 12.2 to detect moderate-tofrequent CERAD scores and 24.4 to detect intermediate-to-high ADNC levels.

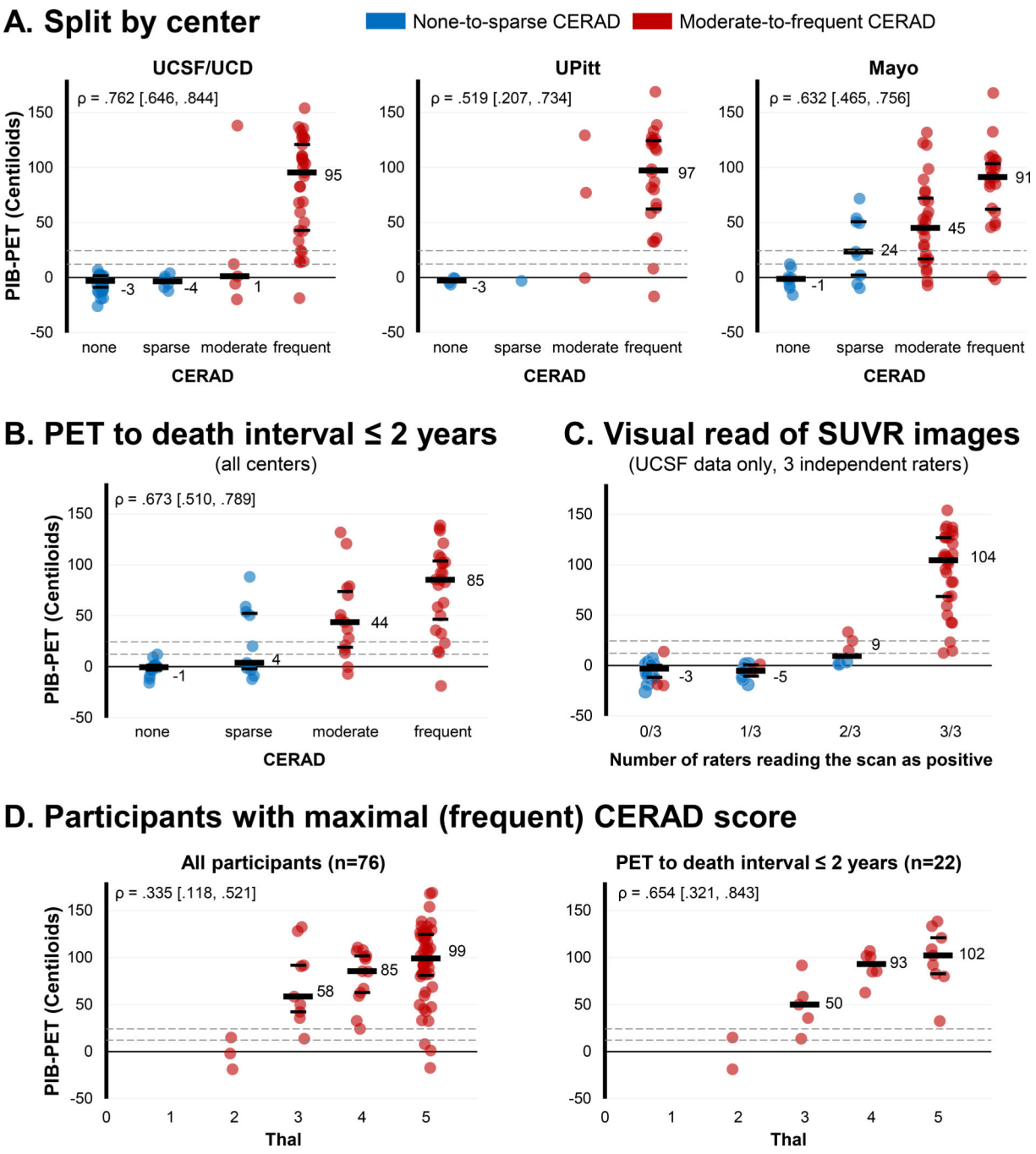

Figure 3. Complementary analyses

A) Robust associations between CL values and CERAD scores across centers. B) Main analyses (similar to Figures 2A) restricted to patients with PET to death interval ≤ 2 years (n=63 from all centers). C) Relationships between visual reads of the SUVR scans (by three independent raters blind to clinical information and CL values) and CL values in the subsample from UCSF/UCD (n=73). D) Relationships between CL values and Thal phase in individuals with maximal (frequent) CERAD scores (n=76 from all centers, including n=22 who died within 2 years of PET).

Scatter plots show individual data points (blue: none-to-sparse CERAD, red: moderate-to-frequent CERAD), medians (if n ≥ 4, the actual values are also specified), and quartiles (if n ≥ 8). Spearman's ρ correlation coefficients [95%CI] are indicated. Dotted lines illustrate the thresholds identified by the ROC analyses (see text): 12.2 to detect moderate-to-frequent CERAD scores and 24.4 to detect intermediate-to-high ADNC levels.

Table 1.

Cohort characteristics

	Total	UCSF/UCD	UPitt	Mayo	AIBL	Group comparison
n	179	73	32	69	5	
Clinical Diagnosis (%) CN-MCI-ADD-nonAD	12–15-35–37	3–8-27–62	9–16-59–16	23–23-30–23	20–0-60–20	$\chi^2_{(6)}$=43.8, p<.001
Male-Female (%)	65–35	60–40	72–28	70–30	40–60	$\chi^2_{(2)}$=2.0, p=.37
Education	15.3 (2.9)	15.8 (2.8)	15.4 (3.1)	14.8 (2.7)	12.0 (3.0)	$F_{(2,166)}$=2.3, p=.11, η^2=.03
Age at PET	73.0 (11.7)	67.0 (9.0)	74.1 (14.9)	78.5 (9.5)	75.2 (14.7)	$F_{(2,171)}$=22.1, p<.001, η^2=.21
MMSE at PET	21.3 (7.1)	22.0 (6.8)	19.8 (7.4)	21.4 (7.2)	19.4 (8.0)	$F_{(2,165)}$=1.1, p=.34, η^2=.01
CDR at PET	1.1 (0.8)	1.1 (0.8)	1.6 (0.9)	0.8 (0.7)	1.4 (1.4)	$F_{(2,161)}$=9.0, p<.001, η^2=.10
PET to death (years)	3.3 (2.1)	3.7 (2.4)	3.8 (2.1)	2.5 (1.5)	2.8 (1.4)	$F_{(2,171)}$=7.5, p<.001, η^2=.08
CERAD Score (%) None-sparse-mod-freq	23–10-21–46	40–8-7–45	13–3-9–75	12–13-41–35	0–40-20–40	$H_{(2)}$=11.4, p=.003
Thal Phase (%) 0–1-2–3-4–5	10–8-12–18-17–35	27–14-11–9-7–32	6–3-3–13-13–63	0–6-14–28-29–23	0–0-40–20-0–40	$H_{(2)}$=16.0, p<.001
ADNC levels (%) No-low-intermed-high	10–26-29–35	24–35-10–31	6–9-25–59	0–25-48–28	0–40-20–40	$H_{(2)}$=15.5, p<.001

CN: Cognitively Normal; MCI: Mild cognitive impairment; ADD: Alzheimer's disease dementia; non-AD: non Alzheimer's disease conditions (see Supplementary Table 1 for details); MMSE: Mini Mental State Examination; CDR: Clinical Dementia Rating scale; CERAD: Consortium to Establish a Registry for Alzheimer's Disease; ADNC: Alzheimer's Disease Neuropathologic Change.

For continuous variables, mean (standard deviation) are indicated. For ordinal and categorical variables, we indicated the percentage of each bin within each sample to facilitate between group comparisons. Due to the small sample size from AIBL, group comparisons focused on UCSF/UCD, UPitt and Mayo participants; ANOVAs were used for continuous variables, Kruskal-Wallis tests for ordinal variables (CERAD, Thal, ADNC), and Chi-square for categorical variables (sex, clinical diagnosis). Missing values from UCSF sample: 17 missing Thal phase, and 11 missing ADNC levels.

Alzheimers Dement. Author manuscript; available in PMC 2020 February 01.

Table 2.

ROC analyses based on multiple pathological standards of truth.

Standard of truth	N	AUROC	Threshold	Accuracy	Sensitivity	Specificity
CERAD score						
None Vs sparse-to-frequent [a]	41 Vs 138	0.919 [0.869–0.955]	12.2	87.2 [81.4–91.7]	83.3 [76.0–89.1]	100 [91.4–100]
None-sparse Vs moderate-frequent [b*]	**59 Vs 120**	**0.910 [0.858–0.948]**	**12.2**	**88.3 [82.6–92.6]**	**89.2 [82.2–94.1]**	**86.4 [75.0–94.0]**
None-to-moderate Vs frequent [c]	96 Vs 83	0.857 [0.797–0.905]	32.4	87.7 [82.0–92.1]	88.0 [79.0–94.1]	71.9 [61.8–80.6]
Thal phase						
0 Vs 1-to-5 [d]	17 Vs 157	0.891 [0.835–0.933]	7.4	78.2 [71.3–84.0]	75.8 [68.3–82.3]	100 [80.5–100]
0–1 Vs 2-to-5	30 Vs 132	0.920 [0.868–0.957]	12.0	85.2 [78.7–90.3]	81.8 [74.2–88.0]	100 [88.4–100]
0-to-2 Vs 3-to-5 [e*]	**49 Vs 113**	**0.923 [0.871–0.959]**	**23.5**	**88.9 [83.0–93.3]**	**85.8 [78.0–91.7]**	**95.9 [86.0–99.5]**
0-to-3 Vs 4–5 [f]	78 Vs 84	0.913 [0.858–0.951]	24.4	87.7 [81.6–92.3]	96.4 [89.9–99.3]	78.2 [67.4–86.8]
0-to-4 Vs 5	106 Vs 56	0.860 [0.797–0.910]	79.9	81.5 [74.6–87.1]	76.8 [63.6–87.0]	84.0 [75.6–90.4]
ADNC levels						
None Vs low-to-high	17 Vs 157	0.891 [0.835–0.933]	7.4	78.2 [71.3–84.0]	75.8 [68.3–82.3]	100 [80.5–100]
None-Low Vs intermediate-high *	**66 Vs 113**	**0.894 [0.840–0.935]**	**24.4**	**85.5 [79.5–90.3]**	**84.1 [76.0–90.3]**	**87.9 [77.5–94.6]**
None-to-intermediate Vs high	114 Vs 59	0.887 [0.830–0.930]	59.3	83.2 [76.8–88.5]	88.1 [77.1–95.1]	80.7 [72.3–87.5]

The table indicates estimates and binomial exact 95% Confidence Intervals. Thresholds were independently determined based on Youden's criteria using each contrast. The maximal number of participants was included for each contrast based on available data.

Some contrasts correspond to differences in A and C scores from the ABC score as follows

[a] C0 versus C1-to-3

[b] C0–1 versus C2–3

[c] C0-to-2 versus C3

[d] A0 versus A1-to-3

[e] A0–1 versus A2–3

[f] A0-to-2 versus A3.

[*] These contrasts were chosen as the primary criteria in our study.

AUROC: Area Under the Receiver Operating Characteristic curve. CERAD: Consortium to Establish a Registry for Alzheimer's Disease; ADNC: Alzheimer's Disease Neuropathologic Change.

Alzheimers Dement. Author manuscript; available in PMC 2020 February 01.

Table 3.

Details of misclassified cases based on a 12.2 CL threshold.

Clinical syndrome	Primary neuropathologic Diagnosis	Age at PET	CL value	PET-to-death (years)	CERAD score	Thal phase	Braak	ADNC level
False positive (CL > 12.2 and none-to-sparse CERAD)								
DLB	LBD	82	20.2	0.9	sparse	2	3	Low
Parkinson Disease	LBD	76	23.5	3.8	sparse	3	2	Low
MCI	AD	82	49.3	6.3	sparse	4	4	Intermediate
DLB	LBD	74	50.5	1.5	sparse	4	3	Intermediate
Normal	Pathological aging[#]	91	53.7	0.2	sparse	3	3	Intermediate
Normal	Lipohyalinosis of the vessels of BG and cerebellum	85	58.8	1.5	sparse	2	4	Low
MCI	AD	85	71.9	3.0	sparse	4	4	Intermediate
DLB	LBD	78	88.2	2.0	sparse	2	5	Low
False negative (CL < 12.2 and moderate-to-frequent CERAD)								
svPPA	FTLD-tau (Pick's)	58	−19.7	2.5	moderate	n/a	1	Low
nfPPA	FTLD-tau (Pick's)	72	−18.8	1.3	frequent	2	1	Low
AD dementia (ADAD)[*]	AD	42	−17.1	4.8	frequent	5	6	High
AD dementia	Pathological aging[#]	88	−7.1	0.9	moderate	1	3	Intermediate
svPPA	FTLD-TDP43	69	−5.7	2.6	moderate	n/a	4	Intermediate
MCI	Pathological aging[#]	77	−3.7	7.1	moderate	3	3	Intermediate
MCI	PSP	80	−1.8	3.5	frequent	2	4	Intermediate
DLB	LBD	77	−0.4	1.4	moderate	3	2	Low
AD dementia	AD	82	1.1	6.6	moderate	3	4	Intermediate
Normal	AD	72	1.2	4.1	frequent	5	5	High
Normal	Pathological aging[#]	80	5.0	2.1	moderate	2	1	Low
MCI	Pathological aging[#]	85	7.4	3.0	moderate	3	3	Intermediate
AD dementia (ADAD)[*]	AD	51	8.1	3.1	frequent	5	6	High

AD: Alzheimer's Disease; DLB: Dementia with Lewy Bodies; svPPA: semantic variant of Primary Progressive Aphasia; MCI: Mild Cognitive Impairment; nfPPA: non fluent Primary Progressive Aphasia; LBD: Lewy Body Disease; FTLD: Frontotemporal Lobar Degeneration; PSP: Progressive Supranuclear Palsy; BG: Basal Ganglia; CERAD: Consortium to Establish a Registry for Alzheimer's Disease; ADNC: level of Alzheimer's Disease Neuropathologic Change.

[#] "Pathological aging" is a term used at the Mayo Clinic to describe cases with evidence for Aβ pathology (sparse-moderate neuritic plaques) but low neurofibrillary tangle pathology (Braak ≤3) that is deemed insufficient to qualify for a pathological diagnosis of AD.

[*] indicates patients with Autosomal Dominant AD (ADAD) and with high cerebellum binding (see Supplementary Figure 2)