



# Evaluation of semi-quantitative measures of $^{18}\text{F}$ -flutemetamol PET for the clinical diagnosis of Alzheimer's disease

Ebba Gløersen Müller<sup>1,2^</sup>, Caroline Stokke<sup>3,4^</sup>, Henning Langen Stokmo<sup>1,2^</sup>, Trine Holt Edwin<sup>5,6^</sup>, Anne-Brita Knapskog<sup>6^</sup>, Mona-Elisabeth Revheim<sup>1,2^</sup>

<sup>1</sup>Division of Radiology and Nuclear Medicine, Department of Nuclear Medicine, Oslo University Hospital, Oslo, Norway; <sup>2</sup>Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway; <sup>3</sup>Division of Radiology and Nuclear Medicine, Department of Diagnostic Physics, Oslo University Hospital, Oslo, Norway; <sup>4</sup>Department of Physics, University of Oslo, Oslo, Norway; <sup>5</sup>Institute of Health and Society, Faculty of Medicine, University of Oslo, Oslo, Norway; <sup>6</sup>Department of Geriatric Medicine, The memory clinic, Oslo University Hospital, Oslo, Norway

*Contributions:* (I) Conception and design: ME Revheim, C Stokke; (II) Administrative support: ME Revheim; (III) Provision of study materials or patients: TH Edwin, AB Knapskog, ME Revheim; (IV) Collection and assembly of data: EG Müller, HL Stokmo, TH Edwin; (V) Data analysis and interpretation: EG Müller, C Stokke, HL Stokmo, ME Revheim; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

*Correspondence to:* Ebba Gløersen Müller. Oslo University Hospital, Division of Radiology and Nuclear Medicine, Department of Nuclear Medicine, Ullevål, Postboks 4956 Nydalen 0424 Oslo, Norway. Email: ebbagm@gmail.com.

**Background:**  $^{18}\text{F}$ -flutemetamol positron emission tomography (PET) is used to assess cortical amyloid- $\beta$  burden in patients with cognitive impairment to support a clinical diagnosis. Visual classification is the most widely used method in clinical practice although semi-quantification is beneficial to obtain an objective and continuous measure of the A $\beta$  burden. The aims were: first to evaluate the correspondence between standardized uptake value ratios (SUVRs) from three different software, Centiloids and visual classification, second to estimate thresholds for supporting visual classification and last to assess differences in semi-quantitative measures between clinical diagnoses.

**Methods:** This observational study included 195 patients with cognitive impairment who underwent  $^{18}\text{F}$ -flutemetamol PET. PET images were semi-quantified with SyngoVia, CortexID suite, and PMOD. Receiver operating characteristics curves were used to compare visual classification with composite SUVR normalized to pons (SUVRpons) and cerebellar cortex (SUVRcer), and Centiloids. We explored correlations and differences between semi-quantitative measures as well as differences in SUVR between two clinical diagnosis groups: Alzheimer's disease-group and non-Alzheimer's disease-group.

**Results:** PET images from 191 patients were semi-quantified with SyngoVia and CortexID and 86 PET-magnetic resonance imaging pairs with PMOD. All receiver operating characteristics curves showed a high area under the curve (>0.98). Thresholds for a visually positive PET was for SUVRcer: 1.87 (SyngoVia) and 1.64 (CortexID) and for SUVRpons: 0.54 (SyngoVia) and 0.55 (CortexID). The threshold on the Centiloid scale was 39.6 Centiloids. All semi-quantitative measures showed a very high correlation between different software and normalization methods. Composite SUVRcer was significantly different between SyngoVia and PMOD, SyngoVia and CortexID but not between PMOD and CortexID. Composite SUVRpons were significantly different between all three software. There were significant differences in the mean rank of SUVRpons, SUVRcer, and Centiloid between Alzheimer's disease-group and non-Alzheimer's disease-group.

<sup>^</sup> ORCID: Ebba Gløersen Müller, 0000-0003-3964-6055; Caroline Stokke, 0000-0003-4465-9635; Henning Langen Stokmo, 0000-0002-2861-0340; Trine Holt Holt Edwin, 0000-0002-0530-8771; Anne-Brita Knapskog, 0000-0003-4867-551X; Mona-Elisabeth Revheim, 0000-0003-3300-7420.

**Conclusions:** SUVR from different software performed equally well in discriminating visually positive and negative  $^{18}\text{F}$ -flutemetamol PET images. Thresholds should be considered software-specific and cautiously be applied across software without preceding validation to categorize scans as positive or negative. SUVR and Centiloid may be used alongside a thorough clinical evaluation to support a clinical diagnosis.

**Keywords:** Flutemetamol; quantification; amyloid-beta; positron emission tomography (PET); Centiloid

Submitted Feb 15, 2021. Accepted for publication Jul 06, 2021.

doi: 10.21037/qims-21-188

**View this article at:** <https://dx.doi.org/10.21037/qims-21-188>

## Introduction

Positron emission tomography (PET) with amyloid-beta ( $\text{A}\beta$ ) tracers, such as  $^{18}\text{F}$ -flutemetamol, is useful to evaluate cortical  $\text{A}\beta$  burden in patients with cognitive impairment (1,2).  $\text{A}\beta$ -PET is acknowledged as a biomarker for Alzheimer's disease (AD), used together with other biomarkers to support a clinical diagnosis (3,4) and to classify patients in the AD continuum (5).  $^{18}\text{F}$ -flutemetamol PET is validated for binary visual classification (6) which is used in the clinical setting for diagnostic purposes. Semi-quantification of tracer uptake is frequently used in research (5,7) and may be used to support a visual classification in clinical practice, which is particularly helpful in equivocal cases. The clinical usefulness of thresholds is still under evaluation. Other strengths of semi-quantification are that it offers a more objective image evaluation compared to visual classification, and it functions as a continuous measure of the  $\text{A}\beta$  burden. This might facilitate the detection of changes in  $\text{A}\beta$  burden over time or in response to an intervention. However, different  $\text{A}\beta$  tracers, imaging protocols, PET scanners, and software for semi-quantification are used between centers. These factors represent a challenge when results are compared across centers and when thresholds are applied. Different commercially available software are used to semi-quantify  $\text{A}\beta$ -PET images and many centers have additionally developed their own semi-quantification methods. SyngoVia (Siemens Healthineers) has been validated for the semi-quantification of  $^{18}\text{F}$ -florbetapir (8). CortexID suite (GE Healthcare), hereafter referred to as CortexID, has been validated for clinical use to support the visual classification of  $^{18}\text{F}$ -flutemetamol PET images (9-12). SyngoVia and CortexID do not enable magnetic resonance imaging (MRI) to be used and utilize only PET or PET together with a low dose computed tomography (CT). PMOD (PMOD Technologies LLC) is a research software comprising a tool for post-processing brain PET images

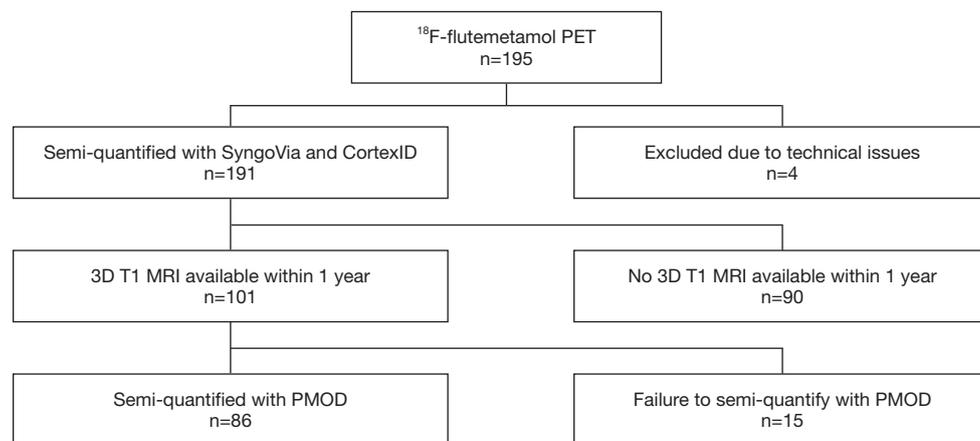
(PNEURO). PNEURO facilitates the use of MRI during post-processing of PET to enable the segmentation of MRI images into grey and white matter. This is beneficial for a subsequent correct assessment of cortical  $^{18}\text{F}$ -flutemetamol and partial volume correction (PVC), to correct errors due to limited spatial resolution of PET images. "The Centiloid Project" was recently published with the purpose to standardize the semi-quantification of  $\text{A}\beta$ -PET (7). To our knowledge, no studies have compared standardized uptake value ratios (SUVR) from commercially available software for semi-quantification of  $^{18}\text{F}$ -flutemetamol PET, although methods of semi-quantification have been explored quite extensively (9,13-16). Hence, it is of interest to explore differences in SUVR obtained from different software and compare SUVR and Centiloid (CL) with validated visual classification to facilitate appropriate use of semi-quantification and thresholds in both research and clinical practice.

We compare  $^{18}\text{F}$ -flutemetamol SUVRs obtained with two commonly used PET-only software (SyngoVia and CortexID) and PMOD using two different atlases (Automatic anatomical labeling (AAL)-merged atlas and CL atlas). The aims are: (I) to evaluate how SUVR<sub>pons</sub> and SUVR<sub>cer</sub> from CortexID, SyngoVia, and PMOD correspond to each other and visual classification, as well as how CL corresponds to visual classification; (II) to estimate semi-quantitative thresholds that may be applied to support visual classification in a memory clinic population; (III) to assess differences in SUVR between patients with AD and non-AD diagnoses.

## Methods

### Study population

This observational cross-sectional study includes 195 patients with cognitive complaints, who had



**Figure 1** Flowchart displaying number of patients with PET and MRI images eligible for semi-quantification with the different software. PET, positron emission tomography; n, number of patients; 3D T1 MRI, volumetric T1-weighted magnetic resonance imaging.

<sup>18</sup>F-flutemetamol PET performed between February 2015 and October 2018. All patients were referred from the Memory Clinic at Oslo University Hospital (OUH), Ullevål. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Informed consent was obtained from all participants in the study by inclusion in the Norwegian registry of persons assessed for cognitive symptoms (NorCog). At the time of inclusion in NorCog, all patients were deemed to have sufficient cognitive capacity to consent. The Regional Ethics Committee for medical research in the South-East of Norway (REC 2017/1929) and the Data Protection Officer at our institution approved the study.

#### ***<sup>18</sup>F-flutemetamol PET and MRI acquisition***

All patients were examined using the same PET/CT scanner, Siemens Biograph40 mCT (Siemens Healthineers, Erlangen, Germany). A standard protocol for <sup>18</sup>F-flutemetamol PET/CT was used for all patients. Image acquisition started approximately 90 minutes (mean 91.2, SD 6.6) after a bolus injection of 185 MBq (mean 187.4, SD 7.9) <sup>18</sup>F-flutemetamol. Low-dose CT was performed for attenuation correction and anatomic information, without contrast enhancement (120 kV, 70 mAs, and with a slice thickness of 3 mm). PET data were acquired for 20 minutes (four frames of 5 minutes each). 3D dynamic emission data were reconstructed using a resolution recovery algorithm with time of flight (TrueX with 2 iterations, 21 subsets, and a Gaussian filter with a full width half maximum (FWHM) of 2 mm and a matrix size

of 400×400). The slice thickness of the reconstructed image series was 2 mm.

Most patients had MRI performed within 1 year of PET, where 101 patients had volumetric T1-weighted images available. MRI was performed as part of the diagnostic evaluation in the clinical routine on scanners at different referring hospitals and clinics.

#### ***Semi-quantification of <sup>18</sup>F-flutemetamol uptake***

Image selection for the different software is shown in *Figure 1*. All PET images were semi-quantified using the “Ratio” function (in later versions named “cortex”) in MI-neurology application from SyngoVia (version VB20) and CortexID (AW server, version 3.2). In SyngoVia the PET volume is spatially normalized to a template built on MNI space, using the AAL-atlas and a VOI set that has been adopted from Fleisher *et al.* (8). CortexID uses an adaptive template method as described by Lundqvist *et al.* (12) to spatially normalize the PET volume to a template similar to “the narrow VOI set” previously described (9,12). SyngoVia provides six regions (anterior cingulate, frontal, parietal, posterior cingulate, precuneus, temporal) covering both hemispheres while CortexID provides eight cortical regions (prefrontal, anterior cingulate, precuneus/posterior cingulate, parietal, lateral temporal, occipital, sensorimotor, temporal mesial) in each hemisphere. Both SyngoVia and CortexID allow pons and cerebellar cortex to be used as reference region. Volumes for each region are not provided, consequently, the composite SUVRs are not volume-weighted but represent the mean of SUVR across

all regions. All PET images with volumetric T1-weighted MRI available within 1 year were semi-quantified with PMOD (17). PMOD uses a co-registration of the PET volume to the patients' MRI. Two atlases were applied in two separate processing pipelines, namely the AAL-merged atlas and Centiloid atlas. All images in the AAL pipeline were processed manually, each volume of interest (VOI) was based on the segmented grey matter from the MRI, and region-based voxel-wise PVC was applied (18). Smaller VOIs were combined to create larger volume-weighted VOIs. For further details of volume merging please refer to the [Appendix 1](#). All segmented grey matter was included in the composite VOI, except for regions covering subcortical grey matter (caudate, putamen, thalamus). The settings for the Centiloid pipeline were applied as described by the manufacturer (19). Visual inspection of VOI placement was used for quality control. SUVR normalized to the whole cerebellum was converted to CL using the equation from the recently published research article validating PMOD for obtaining CL (14).

$$CL = (115.24 \times SUVR_{Flut}) - 107.86 \quad [1]$$

Details of the PMOD processing steps that were performed for this study are outlined in [Appendix 1](#) and are based on previous experience as well as the user manuals (17,19). An overview of differences between software may be inspected in [Appendix 1](#).

### *Visual classification of <sup>18</sup>F-flutemetamol uptake*

All <sup>18</sup>F-flutemetamol PET images were visually classified, blinded from all clinical information, by one nuclear medicine physician (EGM) with more than seven years of experience with PET neuroimaging. This visual classification was compared to the image report in the patient's medical record, which was a consensus read performed by at least two nuclear medicine physicians with experience interpreting <sup>18</sup>F-flutemetamol PET. If there was discrepancy between the study-specific classification and the classification in the image report another reader (MER) classified the images, and the majority read was used for further analyses. SyngoVia was used for image interpretation and all nuclear medicine physicians interpreting <sup>18</sup>F-flutemetamol PET had completed the electronic image reader training program (6). With this validated method, an <sup>18</sup>F-flutemetamol PET scan is regarded as positive when at least one of the 5

bilateral regions (frontal, lateral temporal, lateral parietal, posterior cingulate/precuneus, and striatum) show increased cortical uptake, using pons as reference. If no regions show increased cortical uptake the image is regarded as negative. The regional visual classification was performed by assessing each of the five regions separately.

### *Clinical evaluation*

As described previously (20), all patients were assessed according to a standardized and comprehensive protocol including detailed information from the patients and the caregivers about symptoms, previous disorders, use of medication, and demographic information (21). The cognitive function was assessed by several cognitive tests, including the Mini-Mental State Examination (MMSE) (22), the Consortium to Establish a Registry of Alzheimer's Disease (CERAD) 10-item word list and constructional praxis exercise (23), the Clock Drawing Test (CDT) (24), the Trail Making Tests A and B (TMT A and B) (25), the animal-naming test, the Controlled Oral Word Association Test (COWAT-FAS test) (26-28) and the 15-word short form of the Boston Naming Test (BNT) (29). For this study, the severity of the cognitive impairment was scored by an experienced rater (THE) using the Clinical Dementia Rating scale (CDR) (30). Diagnosis and staging of cognitive impairment were made by two experienced memory clinic physicians (THE and ABK). Diagnoses were made retrospectively based on all available information in medical records, in time-proximity to the PET examination. A third experienced physician was consulted in equivocal cases. All patients were assessed for clinical etiology, hereafter referred to as clinical diagnosis, and stage (subjective cognitive decline (SCD), mild cognitive impairment (MCI), or dementia). SCD was diagnosed using the criteria from the Subjective Cognitive Decline Initiative (31). MCI or dementia, as well as clinical AD diagnoses, were based on the National Institute of Aging and the Alzheimer's Association (NIA-AA)-criteria from 2011 (3,4). According to the NIA-AA 2011 criteria (3), AD mixed is used as a clinical diagnosis when a patient meets all core clinical criteria of AD but has evidence of concomitant disease (cerebrovascular disease, dementia with Lewy bodies, neurological disease, etc.). AD mixed was diagnosed, when appropriate, to all stages of cognitive impairment and not limited to diagnose dementia patients. All other diagnoses were made using the diagnostic criteria as follows: vascular dementia (32), frontotemporal dementia (33), primary

progressive aphasia (34), Parkinson's disease dementia (35), dementia with Lewy bodies (36) and other Parkinson plus disorders (37-40). Patients that did not fit the criteria for any of the above-mentioned disorders or had clear evidence of other diagnoses, were classified as having a non-neurodegenerative disease.

### Statistical analyses

Statistical analyses were performed with IBM SPSS®, version 26 (IBM Corp., Armonk, NY, USA), STATA (version 16.1, StataCorp LLC, Texas, USA), and selected plots were created using Python version 3.6.8 (Python Software Foundation, Beaverton, OR, USA) using the following packages: Jupyter Notebook (v6.0.3), Pandas (v1.0.3), Matplotlib (v3.2.1) and Seaborn (v0.10.1). Excel 2016, version 16.34 (Microsoft Corp., Redmond, WA, USA) was used to calculate Youden's indices.

Descriptive statistics were performed on all variables to assess mean, standard deviation (SD), and confidence interval. The distribution of parameters was assessed with histograms, QQ-plots, and Shapiro-Wilk tests. Cohen's  $\kappa$  was calculated for assessing interrater agreement of visual  $^{18}\text{F}$ -flutemetamol classifications. Receiver operating characteristics (ROC) curves were used to compare composite SUVRs and CL to visual classification and to compare regional SUVRs to regional visual classifications. Wilcoxon signed-rank test was performed to assess differences between composite SUVRpons and SUVRcer from SyngoVia, CortexID, and PMOD. Related samples sign test was performed to assess differences between regional SUVRpons and SUVRcer from SyngoVia, CortexID, and PMOD. Youden's indices were calculated and used to inspect optimal thresholds. Correlations between CL, SUVRpons, and SUVRcer across software were assessed with Pearson correlation coefficients.

False negatives (FN) were used to describe patients with visually positive PET and SUVR below the derived thresholds and false positives (FP) were used to describe patients with visually negative PET and SUVR above the derived thresholds. Patients with clinical diagnoses with expected amyloid pathology, such as in AD, AD mixed, and logopenic variant of primary progressive aphasia were included in the AD group (41,42). The remaining patients were included in the non-AD-group. Mann-Whitney U test was performed to assess for differences in median SUVRcer, SUVRpons, and CL between the AD-group and the non-AD-group. Binomial logistic regression was

performed to assess the likelihood of AD diagnosis from the different semi-quantitative measures, with adjustment for age and sex.

All analyses performed with CL, SUVRpons, and SUVRcer from PMOD were performed on the subgroup of 86 patients. All analyses performed with SUVRpons and SUVRcer from CortexID and SyngoVia were performed on the whole group of 191 patients. Results were considered significant if  $P < 0.05$ .

## Results

Patient characteristics are summarized in *Table 1*. There was a mean of 67.1 days (SD 151.9) between PET and MRI. Interrater agreement between the visual classification performed for this study and the visual classification from the clinical report ( $n=191$ ) gave a Cohen's  $\kappa$  of 0.94 (SE 0.26) with disagreement in image classification in six patients. The SUVRpons ranged from 0.31 to 1.09 in SyngoVia ( $n=191$ ), from 0.32 to 1.09 in CortexID ( $n=191$ ) and from 0.42 to 1.08 in PMOD ( $n=86$ ). The SUVRcer ranged from 1.11 to 3.69 in SyngoVia ( $n=191$ ), from 1.04 to 2.9 in CortexID ( $n=191$ ) and from 1.11 to 3.04 in PMOD ( $n=86$ ). CL ranged from -16.30 to 153.26 ( $n=86$ ). Distribution of SUVRs and CL for patients eligible for semi-quantification with all three software ( $n=86$ ) are displayed in *Figure 2*.

### SUVR against visual classification

ROC curves with composite SUVRpons, composite SUVRcer, and CL against visual classification showed high AUC ( $>0.98$ ) for all software (*Table 2*). ROC curves with regional SUVRpons and SUVRcer (frontal, temporal, parietal, and posterior cingulate/precuneus) from SyngoVia, Cortex ID, and PMOD against regional visual classification showed high AUC ( $>0.96$ ) for all software. Please refer to the [Appendix 1](#) for further details of AUCs from regional classifications.

### Thresholds from different software

Thresholds corresponding to the highest Youden's index showed different values for each software and each normalization region, although SyngoVia and CortexID showed similar numerical SUVRpons thresholds (*Table 2*). Only PMOD with SUVRpons and CL was able to maintain specificities above 80% (PMOD 91% and CL 88%) with a

**Table 1** Patient characteristics for all 191 patients

Patient characteristics	Values	Stage of disease		
		SCI, n	MCI, n	Dementia, n
Females, n	84	–	–	–
Males, n	107	–	–	–
Age (years), mean (min–max, SD)	67.2 (42–82, 7.9)	–	–	–
Age (years) females, mean (SD)	67.0 (7.9)	–	–	–
Age (years) males, mean (SD)	67.3 (8.0)	–	–	–
Visual classification <sup>18</sup> F-flutemetamol PET				
Positive, n	108	–	–	–
Negative, n	83	–	–	–
Cognitive tests				
MMSE, mean (SD) <sup>†</sup>	25 (4.5)	–	–	–
CDR-SOB, mean (SD) <sup>‡</sup>	3.5 (2.7)	–	–	–
Clinical diagnoses, n (%)				
Alzheimer's disease	105 (55.0)	1	25	79
Vascular dementia	13 (6.8)	0	3	10
Frontotemporal lobar degeneration	4 (2.1)	0	0	4
Primary progressive aphasia	9 (4.7)	0	3	6
Lewy body dementia	7 (3.7)	0	0	7
Parkinson disease dementia	1 (0.5)	0	0	1
Other Parkinson plus neurodegeneration	3 (1.6)	0	1	2
Non-neurodegenerative disease	49 (25.6)	17	26	6

<sup>†</sup>, missing in 24 patients within 1 year from PET; <sup>‡</sup>, missing in 7 patients within 1 year from PET; <sup>§</sup>, including Alzheimer's disease and mixed Alzheimer's disease. n, number of patients; y, years; SD, standard deviation; MMSE, mini mental status examination; CDR-SOB, clinical dementia rating - sum of boxes; SCI, subjective cognitive impairment; MCI, mild cognitive impairment.

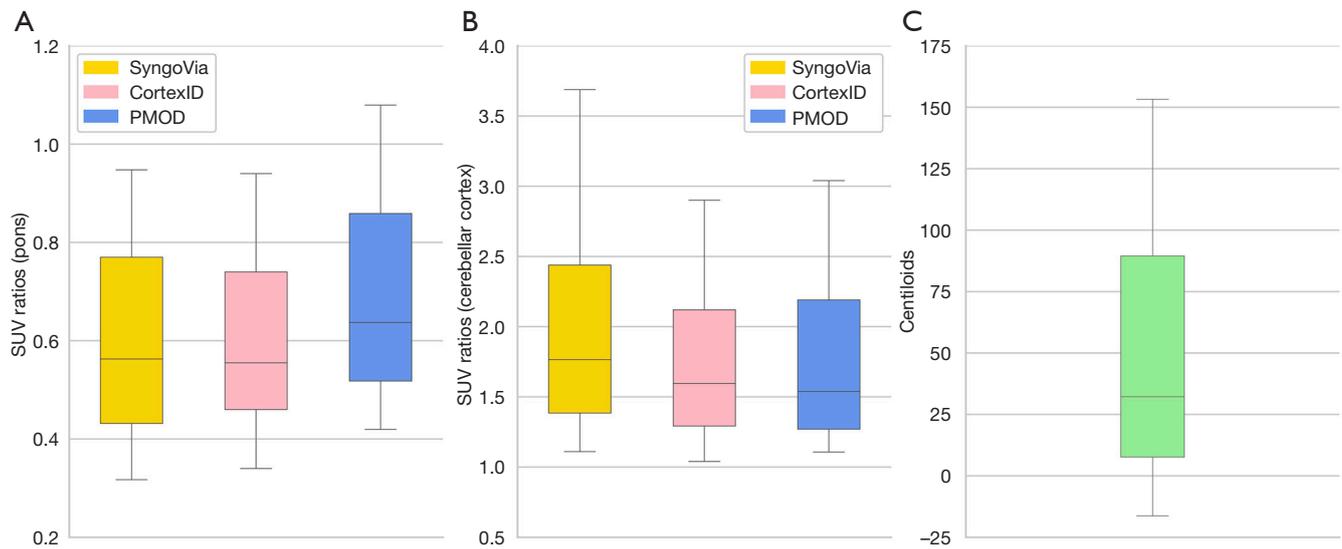
fixed 100% sensitivity.

### Comparison of SUVR between software

Wilcoxon signed-rank test showed no significant median differences between composite SUVR<sub>cer</sub> from CortexID and PMOD. For all other composite SUVR, there were significant median differences between the software (Table 3). SUVRs and CL were not normally distributed. The correlation was assessed with scatterplots and there were linear relationships between SUVR<sub>pons</sub> and SUVR<sub>cer</sub> between the three software as well as between CL and SUVRs from SyngoVia and CortexID (Figures 3,4). Pearson correlation coefficients (r) showed a very strong correlation

( $P < 0.001$ ) between SUVR<sub>pons</sub> and SUVR<sub>cer</sub> from all three software, and between CL and the SUVRs from SyngoVia and CortexID.

Sign tests for exploring differences in regional SUVR showed no significant median differences in SUVR<sub>cer</sub> for the anterior cingulate region ( $P = 0.332$ ) and frontal region ( $P = 0.332$ ) between SyngoVia and PMOD and no significant differences in SUVR<sub>cer</sub> for the parietal region ( $P = 0.332$ ) and the temporal region ( $P = 0.914$ ) between CortexID and PMOD. SUVR<sub>cer</sub> for the posterior cingulate/precuneal region as well as SUVR<sub>pons</sub> for all regions showed significant differences ( $P < 0.001$ ) between the three software. For complete results of the related samples sign tests please refer to the Appendix 1. Figure 5 gives an overview of the



**Figure 2** Distribution of semi-quantitative measures (n=86). Composite SUV ratios normalized to pons for each software (A). Composite SUV ratios normalized to cerebellar cortex for each software (B). Centiloids obtained with PMOD (C). SUV, standardized uptake value ratio.

**Table 2** Thresholds for each software

Semi-quantitative measure	Software	n	Threshold	Sensitivity	Specificity	FN/FP	AUC
SUV <sub>Rpons</sub>	SyngoVia	191	0.537	98%	100%	2/0	0.992
			0.418	100%	43%		
	CortexID	191	0.550	96%	100%	4/0	0.990
			0.425	100%	37%		
	PMOD	86	0.637	96%	98%	2/1	0.995
			0.583	100%	91%		
SUV <sub>Rcer</sub>	SyngoVia	191	1.869	94%	100%	6/0	0.989
			1.513	100%	69%		
	CortexID	191	1.640	93%	100%	8/0	0.992
			1.395	100%	78%		
	PMOD	86	1.556	93%	98%	3/1	0.986
			1.366	100%	71%		
Centiloids	PMOD	86	39.578	93%	100%	3/0	0.995
			23.332	100%	88%		

n, number of patients; FN, false negatives; FP, false positives; AUC, area under the curve; SUV<sub>Rpons</sub>, standardized uptake value ratio normalized to pons; SUV<sub>Rcer</sub>, standardized uptake value ratio normalized to cerebellar cortex.

**Table 3** Differences in composite standardized uptake value ratios (SUVR) from each software

Composite ratios	Software compared	n	Negative diff.	Positive diff.	Ties	P value
SUVRpons	SyngoVia - PMOD	86	84	2	0	<0.001
	CortexID - PMOD	86	81	5	0	<0.001
	SyngoVia - CortexID	86	52	34	0	0.024
		191	111	80	0	0.006
SUVRcer	SyngoVia - PMOD	86	8	78	0	<0.001
	CortexID - PMOD	86	48	38	0	0.933
	SyngoVia - CortexID	86	6	80	0	<0.001
		191	16	175	0	<0.001

P values were obtained through Wilcoxon signed rank tests. n, number of patients; diff, difference; SUVRpons, SUVR normalized to pons; SUVRcer, SUVR normalized to cerebellar cortex.

distribution of regional SUVRs from the different software.

### *Differences in SUVR and CL between clinical diagnoses*

Mann Whitney U tests were run to determine if there were differences in SUVR and CL between patients with expected amyloid pathology (AD-group) and patients without expected amyloid pathology (non-AD-group). The distribution of SUVR and CL was different, as assessed by visual inspection. The mean rank of SUVRcer and SUVRpons from all three software and CL from PMOD was significantly different between the AD-group and the non-AD-group ( $P < 0.001$ ) (Figure 6). Multivariate binomial logistic regression models showed that higher SUVRpons and SUVRcer from all software, and CL, were significantly associated with AD diagnosis ( $P > 0.001$ ), also after adjustment for age and sex. An overview of mean SUVRs and CL across each diagnosis groups is included in the Appendix 1.

## Discussion

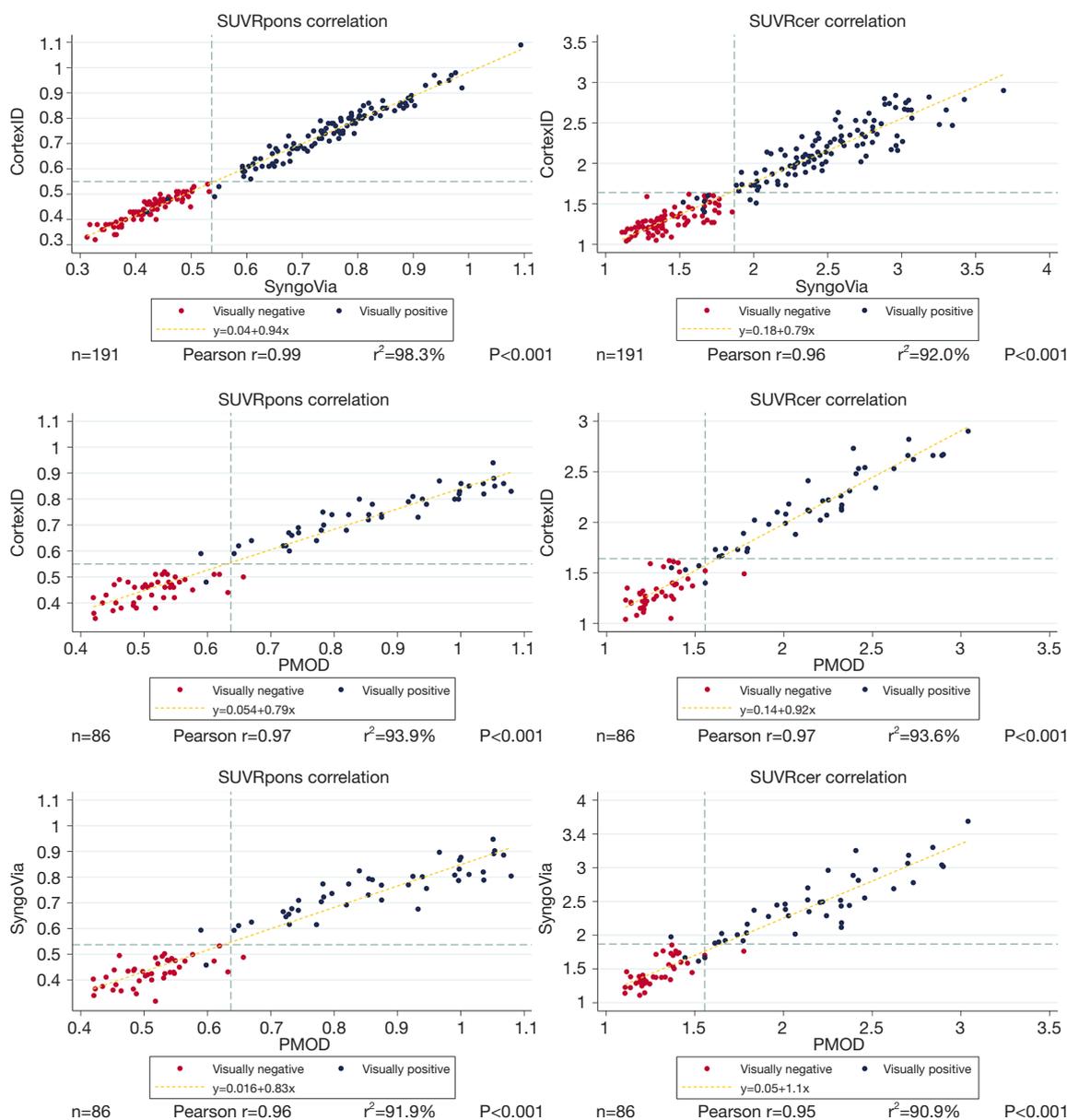
Our study demonstrates that SUVR derived from SyngoVia, CortexID, and PMOD, as well as CL, were equally good for discriminating visually positive and negative  $^{18}\text{F}$ -flutemetamol PET images. There were significant median differences in composite SUVRs obtained from the different software, except for SUVRcer from CortexID and PMOD which showed no significant differences. CortexID gave slightly higher SUVRpons compared to SyngoVia, and SyngoVia gave higher SUVRcer compared to CortexID. PMOD gave higher SUVRpons compared to CortexID

and SyngoVia, and slightly higher SUVRcer compared to SyngoVia.

We found a strong correlation between SUVR from all software as well as CL, demonstrated by the high Pearson correlation coefficients. Thresholds from each software are provided and may be used to aid visual classification in equivocal cases. The mean rank of SUVR and CL was significantly different between patients with expected amyloid pathology and patients without expected amyloid pathology. Logistic regression analyses showed that higher SUVRs and CL were significantly associated with AD-diagnosis, also after adjusting for age and sex.

We used default settings for obtaining SUVR from SyngoVia and CortexID, and a recently validated method for obtaining Centiloid from PMOD (14). The PMOD processing with MRI using the AAL-merged atlas is not a standardized method. Different choices during the image processing might give different results, which is undesirable in clinical routine. PMOD semi-quantification was consequently included for comparison as this software offers a more detailed delineation of cortical structures as well as facilitating PVC. The present study indicates, however, that this detailed semi-quantification is not necessary for supporting the visual classification of positive or negative  $^{18}\text{F}$ -flutemetamol PET. On the contrary, when monitoring the amyloid burden over time or in response to an intervention, it is desirable with a more precise and detailed approach.

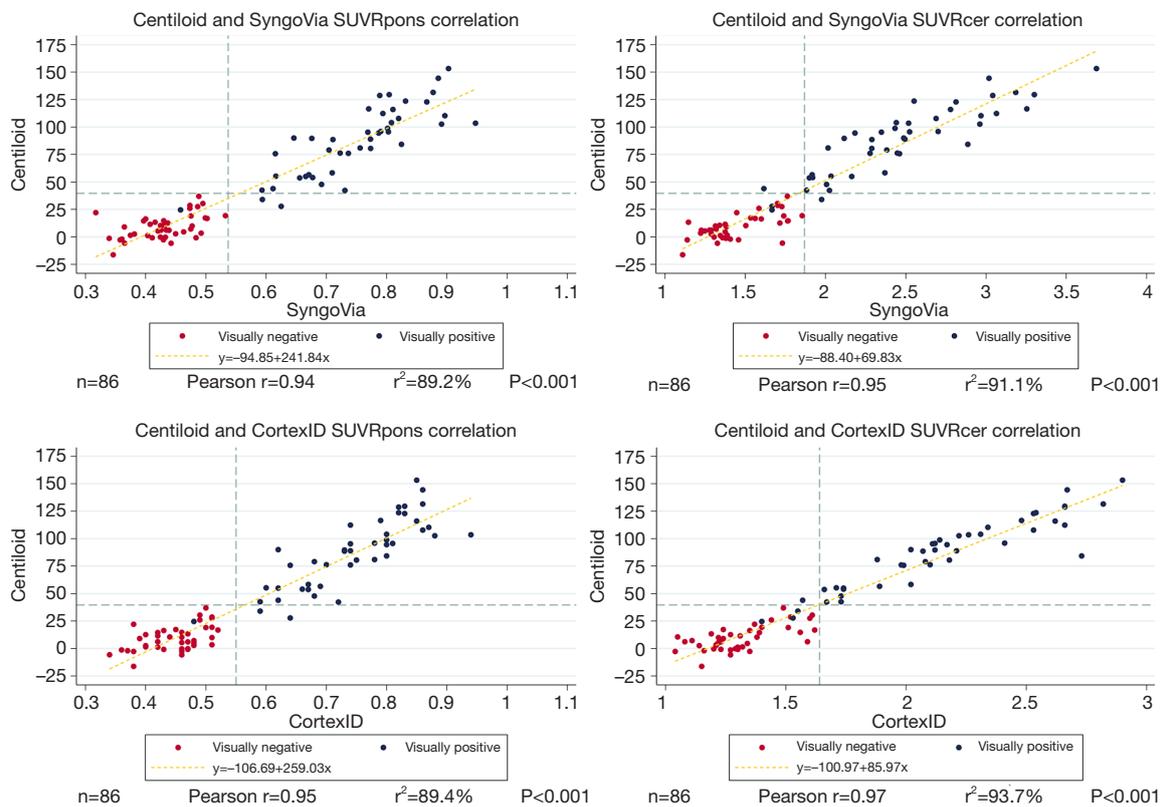
Our results support previous notions that SUVR thresholds are specific for each software and tracer used (43,44). Using a threshold derived from software A to classify images semi-quantified with software B, is generally



**Figure 3** Scatterplots showing the correlation between SUVRpons and SUVRcer. Top row: CortexID and SyngoVia. Middle row: CortexID and PMOD. Bottom row: SyngoVia and PMOD. Dashed lines (light blue) represent the thresholds derived from visual classification for each software, for each normalization method. Dashed line (yellow) represents the linear regression line. SUVRpons, standardized uptake value ratios with normalization to pons; SUVRcer, standardized uptake value ratios with normalization to cerebellar cortex; n, number of patients.

inappropriate. This is an important issue to be aware of. If thresholds are applied across software a preceding validation is important. The CL scale, however, is meant to represent a common quantitative output value across different tracers (13). Different staging systems based on

amyloid PET have been proposed to obtain an estimate of the stage of histopathological changes. Staging systems based on SUVR including striatal regions (45–47) and not including striatal regions (48) as well as staging based on the number of positive regions by visual classification (49)



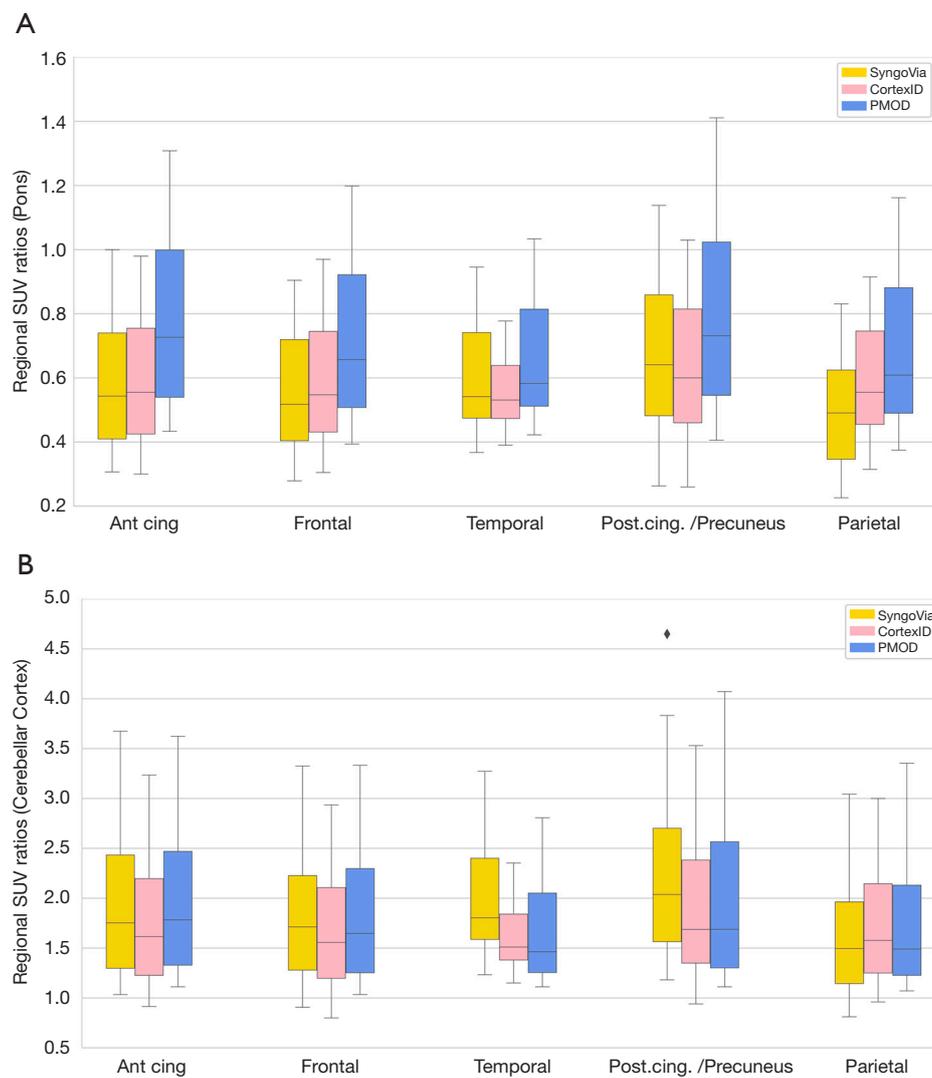
**Figure 4** Scatterplots showing the correlation between Centiloid, SUVRpons and SUVRcer. Top row: Centiloid from PMOD, SUVRpons and SUVRcer from SyngoVia. Bottom row: Centiloid from PMOD, SUVRpons and SUVRcer from CortexID. Dashed lines (light blue) represent the thresholds derived from visual classification for each software, for each normalization method. Dashed line (yellow) represents the linear regression line. SUVRpons, standardized uptake value ratios with normalization to pons; SUVRcer, standardized uptake value ratios with normalization to cerebellar cortex; n, number of patients.

have been evaluated. Subcortical structures are not included in SyngoVia or CortexID, however, the anterior striatum is included in the CL atlas. Visual classification, which includes striatum, is still the gold standard for evaluating amyloid PET. One may argue that inclusion of striatum in semi-quantification would be valuable, however, this is still under investigation. Striatum's role in AD has been suggested to provide a more accurate clinicopathological diagnosis (50) as well as serving as the earliest localization for A $\beta$  deposition in individuals with autosomal dominant AD mutations (51). Nevertheless, a standardized approach for semi-quantification, preferably including striatum, and implemented in commonly used software, would facilitate further exploration of optimal staging systems.

Previously published SUVRpons thresholds from CortexID by Thurfjell *et al.* varies from 0.56 to 0.62, depending on the thickness of VOIs applied in the software (9).

The narrow VOI set, which corresponds to the template applied in CortexID, gave a threshold of 0.58 (9,12), very close to the CortexID threshold of 0.55 in the current study. Interestingly, our CortexID threshold is very similar to the derived SyngoVia SUVRpons threshold of 0.54. In the study by Thurfjell *et al.* (9), the histopathological assessment postmortem was used as the standard of truth as opposed to visual classification in the present study. Cortical <sup>18</sup>F-flutemetamol PET signal has shown to be highly correlated to neuritic plaque burden assessed postmortem (1,6,52,53), making thresholds based on visual classification a reliable approach in a memory clinic population.

Our threshold of 39.6 CL is higher than previously published, ranging from CL 12.2 to 24.4 (54-57), where histopathology postmortem was used as the standard of truth. In a recent study by Collij *et al.* a threshold of CL 17 was found to be optimal compared to visual classification



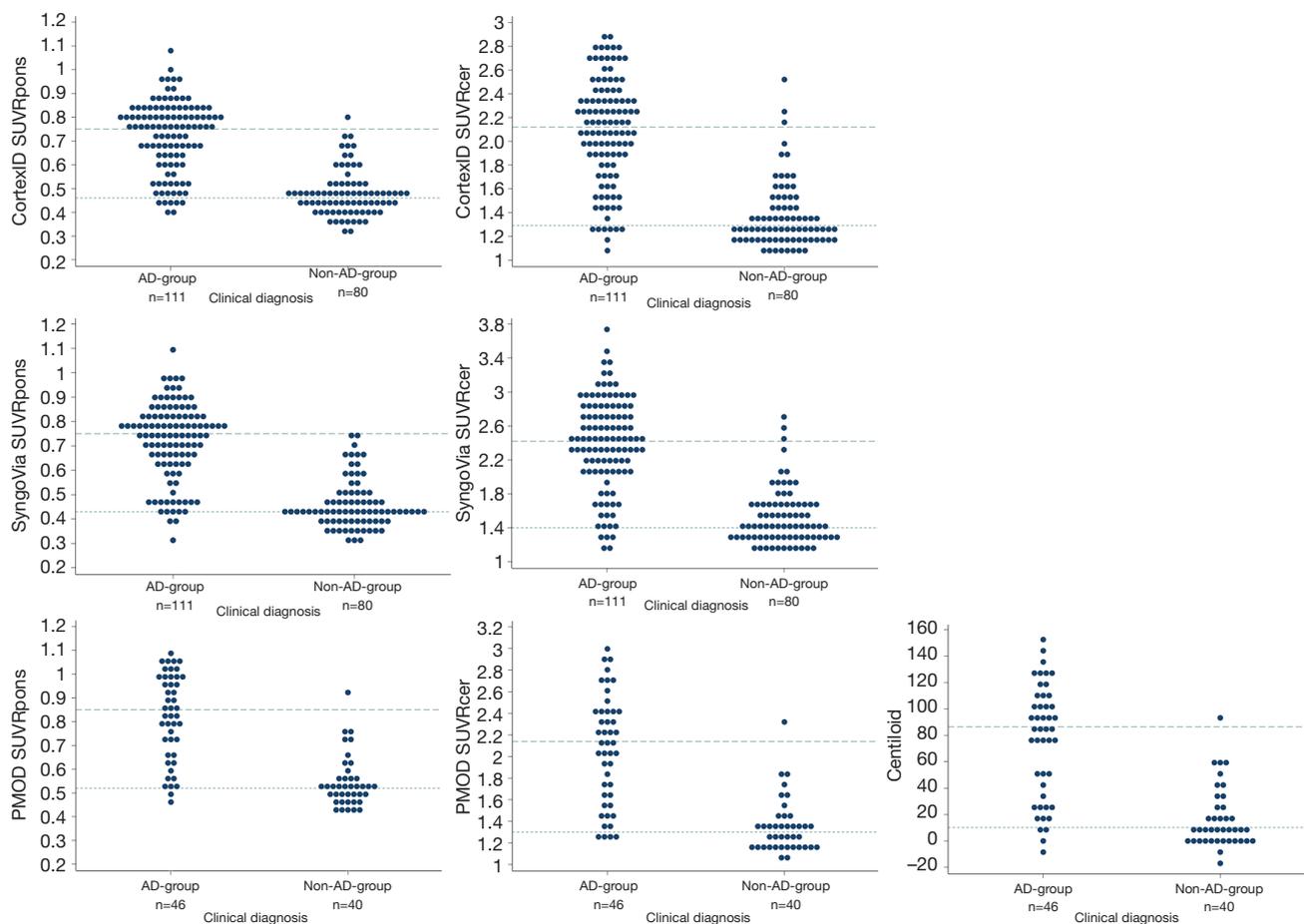
**Figure 5** Distribution of regional semi-quantitative measures (n=86). Distribution of regional standardized uptake value ratios with normalization to pons for each software (A) and regional standardized uptake value ratios with normalization to cerebellar cortex for each software (B). SUV, standardized uptake value.

in a group of 497 patients, in which 352 patients were cognitively unimpaired individuals with increased risk of AD (49). We believe the difference in CL thresholds between this study and ours is mainly due to different populations. However, the presented CL threshold is similar to a recently published threshold of 40 CL with visual classification as the standard of truth (58).

No strict consensus has been reached regarding which region to use for reference when calculating SUVRs. The choice of reference region impacts the numerical value of SUVRs due to different degrees of white and grey matter in

different reference regions.  $^{18}\text{F}$ -flutemetamol has high non-specific binding to white matter. Using pons as a reference will give the lowest SUVR, in our material, ranging from 0.3 to 1.1. Using the whole cerebellum will give a SUVR somewhat higher, due to the mix of white and grey matter. Using cerebellar cortex will give the highest SUVR due to low  $^{18}\text{F}$ -flutemetamol uptake, ranging in our dataset from 1.0 to 3.7. Both pons and cerebellar cortex may exhibit amyloid plaques, however, normally not until the latest phases of amyloid deposition (47,59).

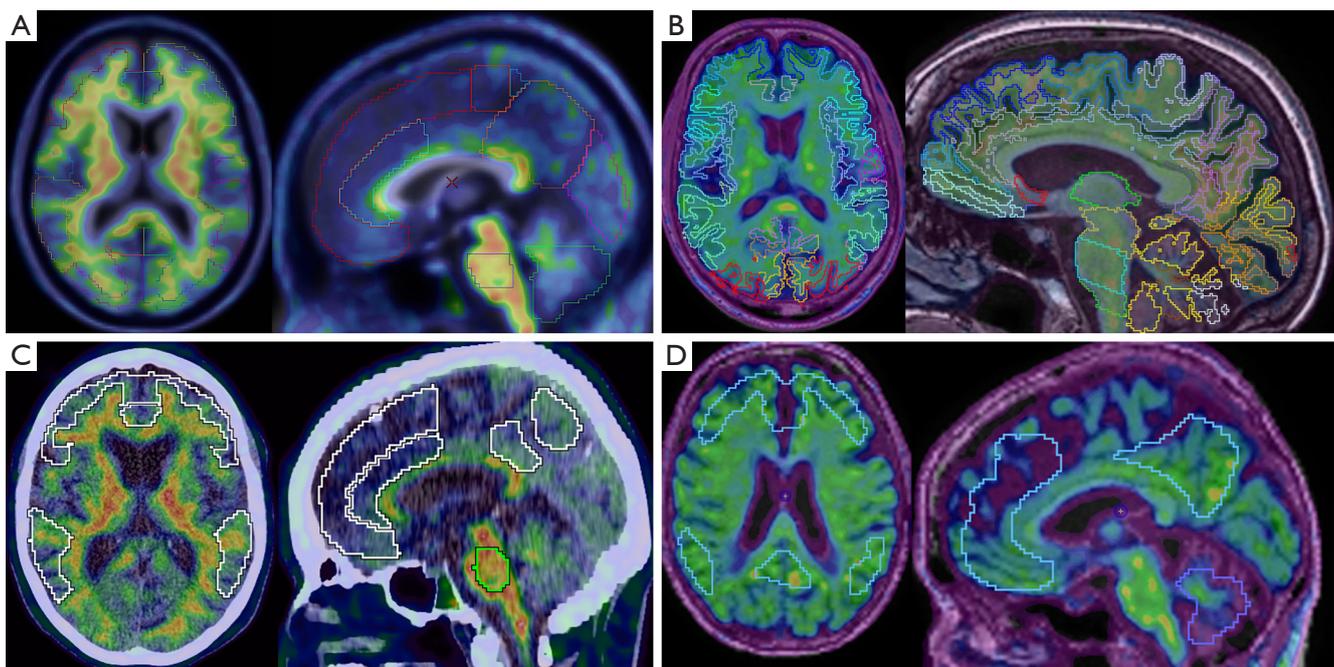
The variation in SUVR between software is likely



**Figure 6** Dot plots showing the distribution of SUVRpons, SUVRcer and Centiloid in the AD-group and the non-AD-group. The long-dashed line represents the median of SUVR and CL in the AD-group and the short-dashed line represent the median of SUVR and CL in the non-AD-group. SUVRpons, standardized uptake value ratios with normalization to pons; SUVRcer, standardized uptake value ratios with normalization to cerebellar cortex; AD-group, Alzheimer's disease group; non-AD group, non-Alzheimer's disease group; n; number of patients.

multifactorial. The atlases and VOIs applied in the three software are different (Figure 7), as described in the methods. SyngoVia uses the AAL-atlas and applies a VOI set that has been adopted from Fleisher *et al.* (8). CortexID uses an adaptive template method using a VOI template previously described (9,12). SUVRs from SyngoVia and CortexID are not volume weighted. In PMOD we used the AAL atlas and each VOI is based on the segmented grey matter from the MRI. We applied both normalization to the cerebellar cortex and pons in SyngoVia, CortexID, and PMOD. To obtain SUVRs for subsequent calculation of CL, another atlas is used, with obligatory normalization against the whole cerebellum. The composite SUVRs are consequently based on different volumes for each software.

In PMOD, only the segmented grey matter is included in the VOI set. For the other software, some degree of white matter will occasionally be included in the VOIs, due to individual differences in gyral pattern. Furthermore, PVC was only applied in the AAL pipeline in PMOD. Nevertheless, the explored software were equally good in discriminating visually positive from negative images in a memory clinic population. As many centers exhibit more than one software for semi-quantification, this finding may be useful. Whether this statement is true also for individuals in the earliest pre-symptomatic stages of the AD continuum should be further explored. Asymptomatic patients are rarely encountered in memory clinics and are thus not referred for amyloid imaging in clinical routine (60),



**Figure 7** Overview of volumes of interest from: CortexID (A), PMOD with automatic anatomical labelling atlas (AAL) merged atlas (B), SyngoVia (C), PMOD with the Centiloid atlas (D). Images are taken from different patients with default angling of images according to each software.

which makes these study questions difficult to answer when investigating a naturalistic memory clinic cohort.

We found significant differences in the mean rank of  $SUVR_{pons}$ ,  $SUVR_{cer}$ , and CL between the AD-group and the non-AD group. Problems of circularity may be suspected when considering that the results of the  $^{18}F$ -flutemetamol PET may have influenced the clinical diagnosis, however, the 2011 diagnostic criteria applied in the present study advise that biomarkers are used for support only.

Limitations of this study are the lack of a control group and the lack of histopathology as the standard of truth for PET, to ensure that the thresholds reliably reflect amyloid deposition. The deviation in image delay time is acknowledged as a limitation, possibly violating equilibrium conditions. All patients were imaged with the same scanner during clinical routine, with a relatively low variation in the remaining examination parameters. The MRIs were performed on different MRI scanners, reflecting the clinical situation with images obtained at different referring hospitals. The quality of MRI was manually assessed by visual inspection and post-processing in PMOD was canceled if the segmentation results appeared distorted. We

acknowledge this as a weakness although we emphasize that MRI was used only for segmentation in PMOD.

## Conclusions

$SUVR$  from different software performed equally well in discriminating visually positive from negative  $^{18}F$ -flutemetamol PET images. Although semi-quantitative measures from different software showed a very high correlation, there were statistically significant differences in  $SUVR$  between them, and thresholds should be considered software specific. Thresholds may be useful for supporting visual classification in clinical routine when semi-quantification is performed with the same software as the threshold is derived from. We demonstrate significant differences in the mean rank of  $SUVR$ s and CL between clinical diagnoses further supporting the usefulness of amyloid PET in a memory clinic setting.

## Acknowledgments

We would like to acknowledge the Norwegian registry of persons assessed for cognitive symptoms (NorCog) for

providing access to the patient data.

*Funding:* This work was supported by Civitan Norway Research Foundation for Alzheimer's disease and Oslo University Hospital.

## Footnote

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/qims-21-188>). HLS has received consultancy fees from Siemens Healthineers. THE has been a rater in two clinical trials (Roche BN29553 and Boehringer-Ingelheim 1346.0023) outside the submitted work. ABK has been a national coordinator and principal investigator in three clinical trials (Roche BN29553, Boehringer-Ingelheim 1346.0023 and Novo Nordisk, EVOKE NN6535-4730) outside the submitted work. The other authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted following the Declaration of Helsinki (as revised in 2013). The Regional Ethics Committee for medical research in the South-East of Norway (REC 2017/1929) and the Data Protection Officer at our institution approved the study. Informed consent was obtained from all individual participants and all participants were deemed to have sufficient cognitive capacity to consent at the time of inclusion.

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

1. 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.
2. Minoshima S, Drzezga AE, Barthel H, Bohnen N, Djekidel M, Lewis DH, Mathis CA, McConathy J, Nordberg A, Sabri O, Seibyl JP, Stokes MK, Van Laere K. SNMMI Procedure Standard/EANM Practice Guideline for Amyloid PET Imaging of the Brain 1.0. *J Nucl Med* 2016;57:1316-22.
3. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R, Mohs RC, Morris JC, Rossor MN, Scheltens P, Carrillo MC, Thies B, Weintraub S, Phelps CH. 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;7:263-9.
4. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, Snyder PJ, Carrillo MC, Thies B, Phelps CH. 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;7:270-9.
5. Jack CR Jr, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement* 2018;14:535-62.
6. Buckley CJ, Sherwin PF, Smith AP, Wolber J, Weick SM, Brooks DJ. Validation of an electronic image reader training programme for interpretation of 18Fflutemetamol  $\beta$ -amyloid PET brain images. *Nucl Med Commun* 2017;38:234-41.
7. Klunk WE, Koeppe RA, Price JC, Benzinger TL, Devous MD, Sr., Jagust WJ, Johnson KA, Mathis CA, Minhas D, Pontecorvo MJ, Rowe CC, Skovronsky DM, Mintun MA. The Centiloid Project: standardizing quantitative amyloid plaque estimation by PET. *Alzheimers Dement* 2015;11:1-15.e1-4.
8. Fleisher AS, Chen K, Liu X, Roontiva A, Thiyyagura P, Ayutyanont N, Joshi AD, Clark CM, Mintun MA, Pontecorvo MJ, Doraiswamy PM, Johnson KA, Skovronsky DM, Reiman EM. Using positron emission tomography and florbetapir F18 to image cortical amyloid in patients with mild cognitive impairment or dementia due to Alzheimer disease. *Arch Neurol* 2011;68:1404-11.
9. Thurfjell L, Lilja J, Lundqvist R, Buckley C, Smith A, Vandenberghe R, Sherwin P. Automated quantification of 18F-flutemetamol PET activity for categorizing scans as

- negative or positive for brain amyloid: concordance with visual image reads. *J Nucl Med* 2014;55:1623-8.
10. Lilja J, Thurfjell L, Sörensen J. Visualization and Quantification of 3-Dimensional Stereotactic Surface Projections for 18F-flutemetamol PET Using Variable Depth. *J Nucl Med* 2016;57:1078-83.
  11. Lilja J, Leuzu A, Chiotis K, Savitcheva I, Sörensen J, Nordberg A. Spatial Normalization of 18F-flutemetamol PET Images Using an Adaptive Principal-Component Template *J Nucl Med* 2019;60:285-91.
  12. Lundqvist R, Lilja J, Thomas BA, Lötjönen J, Villemagne VL, Rowe CC, Thurfjell L. Implementation and validation of an adaptive template registration method for 18F-flutemetamol imaging data. *J Nucl Med* 2013;54:1472-8.
  13. Cho SH, Choe YS, Kim HJ, Jang H, Kim Y, Kim SE, Kim SJ, Kim JP, Jung YH, Kim BC, Baker SL, Lockhart SN, Na DL, Park S, Seo SW. A new Centiloid method for 18F-florbetaben and 18F-flutemetamol PET without conversion to PiB. *Eur J Nucl Med Mol Imaging* 2020;47:1938-48.
  14. Battle MR, Pillay LC, Lowe VJ, Knopman D, Kemp B, Rowe CC, Doré V, Villemagne VL, Buckley CJ. Centiloid scaling for quantification of brain amyloid with 18F-flutemetamol using multiple processing methods. *EJNMMI Res* 2018;8:107.
  15. Hammers DB, Atkinson TJ, Dalley BCA, Suhrie KR, Horn KP, Rasmussen KM, Beardmore BE, Burrell LD, Duff K, Hoffman JM. Amyloid Positivity Using 18F-flutemetamol-PET and Cognitive Deficits in Nondemented Community-Dwelling Older Adults. *Am J Alzheimers Dis Other Dement* 2017;32:320-8.
  16. Chincarini A, Peira E, Morbelli S, Pardini M, Bauckneht M, Arbizu J, et al. Semi-quantification and grading of amyloid PET: A project of the European Alzheimer's Disease Consortium (EADC). *Neuroimage Clin* 2019;23:101846.
  17. PMOD Technologies LLC. PMOD Neuro Tool (PNEURO), User manual Version 4.0. In: User Manual Version 4.0. Available online: [https://www.pmod.com/web/?page\\_id=702](https://www.pmod.com/web/?page_id=702)
  18. Thomas BA, Erlandsson K, Modat M, Thurfjell L, Vandenberghe R, Ourselin S, Hutton BF. The importance of appropriate partial volume correction for PET quantification in Alzheimer's disease. *Eur J Nucl Med Mol Imaging* 2011;38:1104-19.
  19. PMOD Technologies LLC. Centiloid Analysis. In: PMOD application note, version 3.9. PMOD Technologies LLC. Available online: [https://www.pmod.com/files/pdf/applications/Centiloid\\_AppNote.pdf](https://www.pmod.com/files/pdf/applications/Centiloid_AppNote.pdf)
  20. Müller EG, Edwin TH, Stokke C, Navelsaker SS, Babovic A, Bogdanovic N, Knapskog AB, Revheim ME. Amyloid- $\beta$  PET-Correlation with cerebrospinal fluid biomarkers and prediction of Alzheimer's disease diagnosis in a memory clinic. *PLoS One* 2019;14:e0221365.
  21. Brækhus A, Ulstein I, Wyller TB, Engedal K. The Memory Clinic--outpatient assessment when dementia is suspected. *Tidsskr Nor Laegeforen* 2011;131:2254-7.
  22. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189-98.
  23. Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, Mellits ED, Clark C. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* 1989;39:1159-65.
  24. Shulman KI. Clock-drawing: is it the ideal cognitive screening test? *Int J Geriatr Psychiatry* 2000;15:548-61.
  25. Reitan RM. Validity of the Trail Making Test as an indicator of organic brain damage. *Percept Mot Skills* 1958;8:271-6.
  26. Benton A, Hamsher K. Multilingual aphasia examination manual. Iowa City: University of Iowa, 1978.
  27. Lezak MD, Howieson DB, Loring DW, Fischer JS. Neuropsychological assessment. USA: Oxford University Press, 2004.
  28. Malek-Ahmadi M, Small BJ, Raj A. The diagnostic value of controlled oral word association test-FAS and category fluency in single-domain amnesic mild cognitive impairment. *Dement Geriatr Cogn Disord* 2011;32:235-40.
  29. Goodglass H, Kaplan E, Weintraub S. Boston naming test. Philadelphia, PA: Lea & Febiger, 1983.
  30. Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. A new clinical scale for the staging of dementia. *Br J Psychiatry* 1982;140:566-72.
  31. Jessen F, Amariglio RE, van Boxtel M, Breteler M, Ceccaldi M, Chételat G, et al. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimers Dement* 2014;10:844-52.
  32. Sachdev P, Kalaria R, O'Brien J, Skoog I, Alladi S, Black SE, Blacker D, Blazer DG, Chen C, Chui H, Ganguli M, Jellinger K, Jeste DV, Pasquier F, Paulsen J, Prins N, Rockwood K, Roman G, Scheltens P; International Society for Vascular Behavioral and Cognitive Disorders. Diagnostic criteria for vascular cognitive disorders:

- a VASCOG statement. *Alzheimer Dis Assoc Disord* 2014;28:206-18.
33. 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;134:2456-77.
  34. Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, Ogar JM, Rohrer JD, Black S, Boeve BF, Manes F, Dronkers NF, Vandenberghe R, Rascovsky K, Patterson K, Miller BL, Knopman DS, Hodges JR, Mesulam MM, Grossman M. Classification of primary progressive aphasia and its variants. *Neurology* 2011;76:1006-14.
  35. Emre M, Aarsland D, Brown R, Burn DJ, Duyckaerts C, Mizuno Y, et al. Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Mov Disord* 2007;22:1689-707; quiz 1837.
  36. McKeith IG, Boeve BF, Dickson DW, Halliday G, Taylor JP, Weintraub D, et al. Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB Consortium. *Neurology* 2017;89:88-100.
  37. Armstrong MJ, Litvan I, Lang AE, Bak TH, Bhatia KP, Borroni B, Boxer AL, Dickson DW, Grossman M, Hallett M, Josephs KA, Kertesz A, Lee SE, Miller BL, Reich SG, Riley DE, Tolosa E, Tröster AI, Vidailhet M, Weiner WJ. Criteria for the diagnosis of corticobasal degeneration. *Neurology* 2013;80:496-503.
  38. Gilman S, Wenning GK, Low PA, Brooks DJ, Mathias CJ, Trojanowski JQ, Wood NW, Colosimo C, Dürr A, Fowler CJ, Kaufmann H, Klockgether T, Lees A, Poewe W, Quinn N, Revesz T, Robertson D, Sandroni P, Seppi K, Vidailhet M. Second consensus statement on the diagnosis of multiple system atrophy. *Neurology* 2008;71:670-6.
  39. Litvan I. Diagnosis and management of progressive supranuclear palsy. *Semin Neurol* 2001;21:41-8.
  40. Litvan I, Agid Y, Calne D, Campbell G, Dubois B, Duvoisin RC, Goetz CG, Golbe LI, Grafman J, Growdon JH, Hallett M, Jankovic J, Quinn NP, Tolosa E, Zee DS. Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome): report of the NINDS-SPSP international workshop. *Neurology* 1996;47:1-9.
  41. Mesulam M, Wicklund A, Johnson N, Rogalski E, Léger GC, Rademaker A, Weintraub S, Bigio EH. Alzheimer and frontotemporal pathology in subsets of primary progressive aphasia. *Ann Neurol* 2008;63:709-19.
  42. Santos-Santos MA, Rabinovici GD, Iaccarino L, Ayakta N, Tammewar G, Lobach I, et al. Rates of Amyloid Imaging Positivity in Patients With Primary Progressive Aphasia. *JAMA Neurol* 2018;75:342-52.
  43. Jack CR Jr, Wiste HJ, Weigand SD, Therneau TM, Lowe VJ, Knopman DS, Gunter JL, Senjem ML, Jones DT, Kantarci K, Machulda MM, Mielke MM, Roberts RO, Vemuri P, Reyes DA, Petersen RC. Defining imaging biomarker cut points for brain aging and Alzheimer's disease. *Alzheimers Dement* 2017;13:205-16.
  44. Su Y, Flores S, Wang G, Hornbeck RC, Speidel B, Joseph-Mathurin N, et al. Comparison of Pittsburgh compound B and florbetapir in cross-sectional and longitudinal studies. *Alzheimers Dement (Amst)* 2019;11:180-90.
  45. Grothe MJ, Barthel H, Sepulcre J, Dyrba M, Sabri O, Teipel SJ; Alzheimer's Disease Neuroimaging Initiative. In vivo staging of regional amyloid deposition. *Neurology* 2017;89:2031-8.
  46. Hanseeuw BJ, Betensky RA, Mormino EC, Schultz AP, Sepulcre J, Becker JA, Jacobs HIL, Buckley RF, LaPoint MR, Vannini P, Donovan NJ, Chhatwal JP, Marshall GA, Papp KV, Amariglio RE, Rentz DM, Sperling RA, Johnson KA; Harvard Aging Brain Study. PET staging of amyloidosis using striatum. *Alzheimers Dement* 2018;14:1281-92.
  47. Thal DR, Beach TG, Zanette M, Lilja J, Heurling K, Chakrabarty A, Ismail A, Farrar G, Buckley C, Smith APL. Estimation of amyloid distribution by 18Fflutemetamol PET predicts the neuropathological phase of amyloid  $\beta$ -protein deposition. *Acta Neuropathol* 2018;136:557-67.
  48. Mattsson N, Palmqvist S, Stomrud E, Vogel J, Hansson O. Staging  $\beta$ -Amyloid Pathology With Amyloid Positron Emission Tomography. *JAMA Neurol* 2019;76:1319-29.
  49. Collij LE, Salvadó G, Shekari M, Lopes Alves I, Reimand J, Wink AM, Zwan M, Niñerola-Baizán A, Perissinotti A, Scheltens P, Ikonovic MD, Smith APL, Farrar G, Molinuevo JL, Barkhof F, Buckley CJ, van Berckel BNM, Gispert JD; AMYPAD consortium. Visual assessment of 18Fflutemetamol PET images can detect early amyloid pathology and grade its extent. *Eur J Nucl Med Mol Imaging* 2021;48:2169-82.
  50. Beach TG, Thal DR, Zanette M, Smith A, Buckley C. Detection of Striatal Amyloid Plaques with 18Fflutemetamol: Validation with Postmortem Histopathology. *J Alzheimers Dis* 2016;52:863-73.
  51. Cohen AD, Landau SM, Snitz BE, Klunk WE, Blennow K, Zetterberg H. Fluid and PET biomarkers for amyloid pathology in Alzheimer's disease. *Mol Cell Neurosci* 2019;97:3-17.
  52. Ikonovic MD, Buckley CJ, Heurling K, Sherwin P,

- Jones PA, Zanette M, Mathis CA, Klunk WE, Chakrabarty A, Ironside J, Ismail A, Smith C, Thal DR, Beach TG, Farrar G, Smith AP. Post-mortem histopathology underlying  $\beta$ -amyloid PET imaging following flutemetamol F 18 injection. *Acta Neuropathol Commun* 2016;4:130.
53. Ikonomic MD, Fantoni ER, Farrar G, Salloway S. Infrequent false positive  $^{18}\text{F}$ flutemetamol PET signal is resolved by combined histological assessment of neuritic and diffuse plaques. *Alzheimers Res Ther* 2018;10:60.
  54. Amadoru S, Doré V, McLean CA, Hinton F, Shepherd CE, Halliday GM, Leyton CE, Yates PA, Hodges JR, Masters CL, Villemagne VL, Rowe CC. Comparison of amyloid PET measured in Centiloid units with neuropathological findings in Alzheimer's disease. *Alzheimers Res Ther* 2020;12:22.
  55. La Joie R, Ayakta N, Seeley WW, Borys E, Boxer AL, DeCarli C, et al. Multisite study of the relationships between antemortem  $^{11}\text{C}$ PIB-PET Centiloid values and postmortem measures of Alzheimer's disease neuropathology. *Alzheimers Dement* 2019;15:205-16.
  56. Navitsky M, Joshi AD, Kennedy I, Klunk WE, Rowe CC, Wong DF, Pontecorvo MJ, Mintun MA, Devous MD Sr. Standardization of amyloid quantitation with florbetapir standardized uptake value ratios to the Centiloid scale. *Alzheimers Dement* 2018;14:1565-71.
  57. Doré V, Bullich S, Rowe CC, Bourgeat P, Konate S, Sabri O, Stephens AW, Barthel H, Fripp J, Masters CL, Dinkelborg L, Salvado O, Villemagne VL, De Santi S. Comparison of  $^{18}\text{F}$ -florbetaben quantification results using the standard Centiloid, MR-based, and MR-less CapAIBL® approaches: Validation against histopathology. *Alzheimers Dement* 2019;15:807-16.
  58. Hanseeuw BJ, Malotau V, Dricot L, Quenon L, Sznajder Y, Cerman J, Woodard JL, Buckley C, Farrar G, Ivanoiu A, Lhommel R. Defining a Centiloid scale threshold predicting long-term progression to dementia in patients attending the memory clinic: an  $^{18}\text{F}$  flutemetamol amyloid PET study. *Eur J Nucl Med Mol Imaging* 2021;48:302-10.
  59. 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.
  60. Johnson KA, Minoshima S, Bohnen NI, Donohoe KJ, Foster NL, Herscovitch P, Karlawish JH, Rowe CC, Carrillo MC, Hartley DM, Hedrick S, Pappas V, Thies WH; Amyloid Imaging Taskforce. 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* 2013;9:e-1-16.

**Cite this article as:** Müller EG, Stokke C, Stokmo HL, Edwin TH, Knapskog AB, Revheim ME. Evaluation of semi-quantitative measures of  $^{18}\text{F}$ -flutemetamol PET for the clinical diagnosis of Alzheimer's disease. *Quant Imaging Med Surg* 2022;12(1):493-509. doi: 10.21037/qims-21-188

### Overview of region merging in PMOD (AAL-merged)

Frontal region: Precentral\_l, Precentral\_r, Rolandic\_Oper\_l, Rolandic\_Oper\_r, Supp\_Motor\_Area\_l, Supp\_Motor\_Area\_r, Olfactory\_l, Olfactory\_r, Frontal\_Sup\_l, Frontal\_Sup\_r, Frontal\_Mid\_l, Frontal\_Mid\_r, Frontal\_Inf\_l, Frontal\_Inf\_r, Rectus\_l, Rectus\_r, Paracentral\_Lobule\_l, Paracentral\_Lobule\_r

Insular region: Insula\_l, Insula\_r

Anterior cingulate region: Cingulum\_Ant\_l, Cingulum\_Ant\_r

Posterior cingulate region: Cingulum\_Post\_l, Cingulum\_Post\_r

Temporal region: Hippo\_Parahippo\_l, Hippo\_Parahippo\_r, Amygdala\_l, Amygdala\_r, Fusiform\_l, Fusiform\_r, Heschl\_l, Heschl\_r, Temporal\_l, Temporal\_r

Parietal region: Postcentral\_l, Postcentral\_r, SupraMarginal\_l, SupraMarginal\_r, Angular\_l, Angular\_r, Parietal\_l, Parietal\_r

Precuneal region: Precuneus\_l, Precuneus\_r

Occipital region: Calcarine\_l, Calcarine\_r, Cuneus\_l, Cuneus\_r, Lingual\_l, Lingual\_r, Occipital\_l, Occipital\_r

Pons: Pons

Cerebellar cortex: Vermis, Cerebellum\_Crus\_l, Cerebellum\_Crus\_r, Cerebellum\_l, Cerebellum\_r

Composite region: all of the above in addition to Cingulum\_Mid\_l and Cingulum\_Mid\_r

Regions not included: CaudateNucl\_l, CaudateNucl\_r, Putamen\_l, Putamen\_r, Pallidum\_l, Pallidum\_r, Thalamus\_l, Thalamus\_r, Medulla, Midbrain.

### Overview of region merging in CortexID

1. Prefrontal left + prefrontal right/2 = frontal region
  2. Anterior cingulate left + anterior cingulate right /2 = anterior cingulate region
  3. Precuneus/Posterior cingulate left + Precuneus/Posterior cingulate right/2 = Prec.PCC region
  4. Parietal left + parietal right / 2 = parietal region
  5. Temporal lateral left + temporal lateral right + temporal mesial left + temporal mesial right /4 = temporal region
- The occipital and sensorimotor regions from CortexID was included in the composite score but not included in regional analyses.

### Overview of region merging in SyngoVia

1. Posterior cingulate + precuneus / 2 = Precuneus / Posterior cingulate cortex region.
- All other SyngoVia regions were used as they are displayed in the software.

### PMOD, PNEURO v. 4.0 processing steps for AAL-merged pipeline

Subheadings follow the user manual for PNEURO version 4.0 from PMOD (page 64–83) (17).

#### *PET image loading and time averaging*

Crop PET. No motion correction or image denoising was applied.

#### *MR image loading and segmentation:*

Crop MR. Varied between autocrop and manual crop, depending on the results from segmentation. Default settings were applied: Denoising: low, Segmentation: 3 probability maps, sampling: 3 pix, bias regularization: light 60, cleanup: light, Affine regularization: European brains.

### ***PET to MR matching***

Segmentation was manually inspected. “PET-MR Matching required” was chosen. Matching sampling 3 mm was applied.

### ***MR-based normalization***

Matching was manually verified. Probability Maps transformation was applied. Split brain was chosen.

### ***Brain segments calculation***

Normalization was verified manually. AAL-merged atlas was chosen if not already selected. Inspection of «validate split» was performed, although not relevant for this study. White matter parcellation was not performed.

### ***Outlining of brain structures***

In Result space – “Input”: “MR” was chosen and «Mask by» «Probability». “Individual” was used with the following thresholds: GM → 0.3 and CSF → 0.5. “Mask non-cortical regions” was not applied.

### ***Brain VOI editing and statistics calculation***

If VOIs included non-brain structures (most often meninges) the eraser function was used to correct the VOIs. Otherwise, no changes were applied. PVC was applied using the “Region based voxel-wise, using a resolution of 2 mm×2 mm×2 mm (based on measurements from the scanner used).

All edited VOIs were saved after editing. All protocols were saved.

## **PMOD, PNEURO v.4.0 processing steps for Centiloid pipeline**

All processing steps were applied according to the PMOD Application note version 3.9 for Centiloid Analysis (19).

PET: CROP 20×30×20 as autocrop size.

MR: CROP 20×30×20 as autocrop size.

Denoising: medium, Segmentation: 3 prob maps, Sampling: 3 mm, Bias regularization: Medium, 60 kernel, Cleanup: light, Affine registration: European brains

→Segment MR

PET-MR matching required, PET: 4.0 x 4.0 x 4.0, Matching sample: 2 mm

→ Match PET to MR

Inspect, No split brain, Inspect

→Segment brain

Result space: Atlas, Mask by: Probability – Individual, Mask by GM: 0.0, No CSF mask, Mask non-cortical regions

→Outline

No partial volume correction was applied, QC (quality control) (yes)

→Statistics

Relative to WC (whole cerebellum). Save protocol. Inspect images from quality control

**Table S1** Overview of differences in software

	SyngoVia	CortexID	PMOD AAL-merged	PMOD Centiloid
Reference region/s	Pons, cerebellum whole, cerebellar cortex	Pons, cerebellum whole, cerebellar cortex	Optional	Cerebellum whole
Number of regions	6 cortical	16 cortical	71 cortical and white matter	1 cortical composite
Uni/bilateral regions	Bilateral	Unilateral	Unilateral	Bilateral
3D MRI obligatory	No	No	Yes	Yes
PVC applied	No	No	Yes	No
SUV for each region	Yes	No	Yes	Optional
Volume for each region	No	No	Yes	Optional
SUVR displayed	Yes	Yes	No	Optional

3D MRI, volumetric magnetic resonance imaging; PVC, partial volume correction; SUV, standardized uptake value; SUVR, standardized uptake value ratio.

**Table S2** Results from ROC curves of regional uptake against regional visual classification

Region	Area under the curve		
	PMOD, n=86	CortexID, n=191	SyngoVia, n=191
Frontal SUVR <sub>pons</sub>	0.990	0.997	0.996
Frontal SUVR <sub>cer</sub>	0.986	0.988	0.990
Temporal SUVR <sub>pons</sub>	0.988	0.995	0.989
Temporal SUVR <sub>cer</sub>	0.978	0.972	0.966
Parietal SUVR <sub>pons</sub>	0.980	0.982	0.976
Parietal SUVR <sub>cer</sub>	0.979	0.974	0.984
Posterior cingulate/Precuneus SUVR <sub>pons</sub>	0.994	0.996	0.993
Posterior cingulate/Precuneus SUVR <sub>cer</sub>	0.988	0.984	0.970

ROC, receiver operating characteristic; SUVR, Standardized uptake value ratio; SUVR<sub>pons</sub>, SUVR normalized to pons; SUVR<sub>cer</sub>, SUVR normalized to cerebellar cortex.

**Table S3** Related samples sign test of regional SUVRpons (n=86)

Regional SUVR	Software tested	Negative diff.	Positive diff.	Ties	P value
Frontal SUVRpons	SyngoVia – PMOD	86	0	0	<0.001
	CortexID – PMOD	84	2	0	<0.001
	SyngoVia – CortexID	75	11	0	<0.001
Ant.cing SUVRpons	SyngoVia – PMOD	86	0	0	<0.001
	CortexID – PMOD	85	1	0	<0.001
	SyngoVia – CortexID	58	26	2	0.001
Temporal SUVRpons	SyngoVia – PMOD	76	10	0	<0.001
	CortexID – PMOD	79	7	0	<0.001
	SyngoVia – CortexID	22	64	0	<0.001
Parietal SUVRpons	SyngoVia – PMOD	86	0	0	<0.001
	CortexID – PMOD	77	9	0	<0.001
	SyngoVia – CortexID	85	1	0	<0.001
Post.cing/precuneus SUVRpons	SyngoVia – PMOD	79	7	0	<0.001
	CortexID – PMOD	86	0	0	<0.001
	SyngoVia – CortexID	16	70	0	<0.001

SUVR, standardized uptake value ratio; SUVRpons, SUVR normalized to pons; diff, difference; Ant.cing, anterior cingulate cortex; Post.cing, posterior cingulate cortex.

**Table S4** Related samples sign test of regional SUVRcer (n=86)

Regional SUVR	Software tested	Negative diff.	Positive diff.	Ties	P value
Frontal SUVRcer	SyngoVia - PMOD	38	48	0	0.332
	CortexID - PMOD	64	22	0	<0.001
	SyngoVia - CortexID	13	73	0	<0.001
Ant.cing SUVRcer	SyngoVia - PMOD	48	38	0	0.332
	CortexID - PMOD	78	8	0	<0.001
	SyngoVia - CortexID	5	81	0	<0.001
Temporal SUVRcer	SyngoVia - PMOD	86	0	0	<0.001
	CortexID - PMOD	42	44	0	0.914
	SyngoVia - CortexID	2	84	0	<0.001
Parietal SUVRcer	SyngoVia - PMOD	58	28	0	0.002
	CortexID - PMOD	48	38	0	0.332
	SyngoVia - CortexID	67	19	0	<0.001
Post.cing/precuneus SUVRcer	SyngoVia - PMOD	12	74	0	<0.001
	CortexID - PMOD	63	23	0	<0.001
	SyngoVia - CortexID	1	85	0	<0.001

SUVR, standardized uptake value ratio; SUVRcer, SUVR normalized to cerebellar grey matter; diff, difference; Ant.cing, anterior cingulate cortex; Post.cing, posterior cingulate cortex.

**Table S5** Semi-quantitative measures per diagnosis group

	n	SyngoVia		CortexID		n	PMOD		
		SUVRpons	SUVRcer	SUVRpons	SUVRcer		SUVRpons	SUVRcer	CL
AD	76	0.75 (0.15)	2.46 (0.48)	0.75 (0.14)	2.17 (0.41)	32	0.83 (0.18)	2.08 (0.47)	79 (40.0)
AD mixed	29	0.63 (0.17)	2.19 (0.65)	0.64 (0.16)	1.86 (0.49)	11	0.79 (0.21)	1.96 (0.64)	65 (53.0)
VaD	13	0.45 (0.09)	1.55 (0.24)	0.47 (0.10)	1.38 (0.27)	3	0.66 (0.24)	1.68 (0.56)	40 (43.0)
FTD	4	0.42 (0.06)	1.36 (0.18)	0.43 (0.04)	1.22 (0.12)	2	0.55 (0.03)	1.38 (0.00)	14 (5.0)
PPA <sup>†</sup>	9	0.63 (0.19)	2.09 (0.65)	0.63 (0.19)	1.79 (0.51)	5	0.73 (0.25)	1.84 (0.69)	52 (57.0)
DLB	7	0.58 (0.12)	1.96 (0.46)	0.58 (0.94)	1.72 (0.32)	4	0.63 (0.11)	1.52 (0.26)	35 (25.0)
PDD	1	0.41 (0.00)	1.35 (0.00)	0.42 (0.00)	1.22 (0.00)	0	–	–	–
Park plus	3	0.50 (0.15)	1.84 (0.17)	0.51 (0.16)	1.60 (0.22)	1	0.64 (0.00)	1.61 (0.00)	42 (0.0)
Other	49	0.45 (0.09)	1.47 (0.31)	0.46 (0.09)	1.22 (0.26)	28	0.52 (0.09)	1.29 (0.18)	12 (17.0)

†, including patients with logopenic variant of primary progressive aphasia. Values are displayed as mean (standard deviation). n, number of patients; SUVRpons, standardized uptake value ratio with normalization to pons; SUVRcer, standardized uptake value ratio with normalization to cerebellar cortex; CL, centiloids; AD, Alzheimer's disease; VaD, Vascular dementia; FTD, Frontotemporal lobar degeneration; PPA, Primary progressive aphasia; DLB, Dementia with Lewy Bodies; PDD, Parkinson disease dementia; Park plus, Parkinson plus disorders.