

ORIGINAL RESEARCH

Open Access



# Software compatibility analysis for quantitative measures of [<sup>18</sup>F]flutemetamol amyloid PET burden in mild cognitive impairment

Hugh G. Pemberton<sup>1,2,3\*</sup> , Christopher Buckley<sup>1</sup>, Mark Battle<sup>1,4</sup> , Ariane Bollack<sup>2</sup>, Vrajesh Patel<sup>1</sup>, Petya Tomova<sup>1</sup>, David Cooke<sup>5</sup>, Will Balhorn<sup>6</sup>, Katherine Hegedorn<sup>6</sup>, Johan Lilja<sup>7</sup>, Christine Brand<sup>1</sup> and Gill Farrar<sup>1</sup>

## Abstract

**Rationale** Amyloid- $\beta$  (A $\beta$ ) pathology is one of the earliest detectable brain changes in Alzheimer's disease pathogenesis. In clinical practice, trained readers will visually categorise positron emission tomography (PET) scans as either A $\beta$  positive or negative. However, adjunct quantitative analysis is becoming more widely available, where regulatory approved software can currently generate metrics such as standardised uptake value ratios (SUVr) and individual Z-scores. Therefore, it is of direct value to the imaging community to assess the compatibility of commercially available software packages. In this collaborative project, the compatibility of amyloid PET quantification was investigated across four regulatory approved software packages. In doing so, the intention is to increase visibility and understanding of clinically relevant quantitative methods.

**Methods** Composite SUVr using the pons as the reference region was generated from [<sup>18</sup>F]flutemetamol (GE Healthcare) PET in a retrospective cohort of 80 amnesic mild cognitive impairment (aMCI) patients (40 each male/female; mean age = 73 years, SD = 8.52). Based on previous autopsy validation work, an A $\beta$  positivity threshold of  $\geq 0.6$  SUVr<sub>pons</sub> was applied. Quantitative results from MIM Software's *MIMneuro*, Syntermed's *NeuroQ*, Hermes Medical Solutions' *BRASS* and GE Healthcare's *CortexID* were analysed using intraclass correlation coefficient (ICC), percentage agreement around the A $\beta$  positivity threshold and kappa scores.

**Results** Using an A $\beta$  positivity threshold of  $\geq 0.6$  SUVr<sub>pons</sub>, 95% agreement was achieved across the four software packages. Two patients were narrowly classed as A $\beta$  negative by one software package but positive by the others, and two patients vice versa. All kappa scores around the same A $\beta$  positivity threshold, both combined (Fleiss') and individual software pairings (Cohen's), were  $\geq 0.9$  signifying "almost perfect" inter-rater reliability. Excellent reliability was found between composite SUVr measurements for all four software packages, with an average measure ICC of 0.97 and 95% confidence interval of 0.957–0.979. Correlation coefficient analysis between the two software packages reporting composite z-scores was strong ( $r^2 = 0.98$ ).

**Conclusion** Using an optimised cortical mask, regulatory approved software packages provided highly correlated and reliable quantification of [<sup>18</sup>F]flutemetamol amyloid PET with a  $\geq 0.6$  SUVr<sub>pons</sub> positivity threshold. In particular, this work could be of interest to physicians performing routine clinical imaging rather than researchers performing more

\*Correspondence:

Hugh G. Pemberton  
h.pemberton@ucl.ac.uk

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

bespoke image analysis. Similar analysis is encouraged using other reference regions as well as the Centiloid scale, when it has been implemented by more software packages.

**Keywords** Amyloid PET, SUVR, Alzheimer's, MCI, Quantification, [<sup>18</sup>F]flutemetamol

## Introduction

The distribution of brain amyloid-beta (A $\beta$ ) can be measured using positron emission tomography (PET). Three Fluorine-18 radiolabelled A $\beta$  PET tracers have been approved for clinical use: [<sup>18</sup>F]flutemetamol (Vizamyl<sup>TM</sup>; GE Healthcare) [1], [<sup>18</sup>F]florbetaben (Neuraceq<sup>TM</sup>; Life Molecular Imaging) [2] and [<sup>18</sup>F]florbetapir (Amyvid<sup>TM</sup>) [3]. Clinical appraisal of amyloid PET imaging involves binary classification (A $\beta$  negative or positive) through visual assessment, which has been demonstrated as approximately 90% accurate in advanced clinical and end-of-life patients [1, 2, 4]. Over the last two decades, multiple studies have demonstrated the clinical utility of amyloid PET [5–15]. In addition, real-world studies have shown that an amyloid PET scan can increase diagnostic confidence [5, 9, 12, 14, 15], change etiological diagnosis in 25–44% of cases [5, 8, 9, 16] and change patient management in 37–72% of cases [8–10, 12].

However, in recent years memory clinics are increasingly assessing 'pre-dementia' patients, with ~25% of cases presenting with subjective cognitive decline (SCD) or early mild cognitive impairment (MCI) [17]. In these patients, amyloid deposition might be focal or early-stage [18], which may confound visual assessment, especially by less experienced readers [19]. In these cases, the binary classification approach may be more prone to subjectivity given the reliance on the clinician's prior experience, possibly resulting in greater inter-rater variability [3, 20–23]. Adjunct quantitative measures of A $\beta$  deposition, such as Standardized Uptake Value ratio (SUVR) [24], may also bring clinical benefit for early assessment [11, 25–27]. SUVR quantifies the ratio of tracer uptake between a reference region and a target region, when the radiotracer is estimated to have reached pseudo-equilibrium [24]. Furthermore, quantification could provide greater clinical utility alongside current dichotomous classification, such as improvements in diagnostic confidence [8, 10, 21], prediction of cognitive decline [28–31] and changes to diagnosis [16] and patient management [32–37].

Amyloid PET quantification has been used in research since the discovery of Carbon-11 labelled Pittsburgh compound B ([<sup>11</sup>C]PiB), in 2004 [24, 38]. This has resulted in several sophisticated examples of research software for processing and quantifying amyloid PET, such as PMOD, CapAIBL [39] NiftyPET [40], EvaLuation of

Brain Amyloidosis (ELBA) [41], AmyPype [42] and rPOP [43]. Concurrently, various regulatory approved (FDA 510k/CE-marked Class IIa) software packages have been designed for use in clinic, yet none are currently in widespread use for amyloid PET quantification. Clinical use of amyloid PET in the USA relies upon visual inspection, however it is interesting to note that the 2020 SNMMI Value Initiative "National Amyloid Survey" found 52% of sites (out of 176 surveyed with amyloid imaging experience) were using adjunct quantification software. The translational paucity among the other half of respondent may be down to a number of factors, but a lack of clinical validation is likely to contribute [44]. One aspect of the current work is to demonstrate the compatibility of software tools to ensure generalisability amongst users.

Therefore, this validation study aimed to investigate composite SUVR from a group of clinically relevant patients with amnesic MCI (aMCI) across four regulatory approved software packages, and measure the concordance of these quantitative results. In doing so, the secondary aim is to increase visibility and understanding of clinically available quantitative methods. The use of the composite SUVR measure was recently endorsed by the recent RSNA QIBA profile as a relevant and logistically feasible measure for amyloid quantification [45]. Each software package has unique individual features, but this collaborative project brought competing vendors together to demonstrate the concordance of results using a single relevant measure when a composite mask was implemented, and to facilitate use of quantification in routine clinical assessment. The hypothesis was that after the harmonisation exercise, all software packages under investigation will provide highly correlated quantitative results (composite SUVR) according to kappa scores, and pairwise/groupwise correlation.

## Methods

### Patient and scan information

All patients included in this analysis had previously participated in Institutional Review Board (IRB), Independent Ethics Committee (IEC)-approved studies and they (or their legally authorized representative) provided informed written consent to participate; data usage and image analysis was considered to be covered by the previous consent.

This retrospective study included two analysis phases: a pilot and a validation phase. Images for each phase were taken from patients who had participated in previous development studies for [ $^{18}\text{F}$ ]flutemetamol. The pilot data comprised of imaging sets from 11 patients from the phase II ALZ201 study [46]. The larger validation data ( $n=80$  validation data, 40 each male/female; mean age = 73 years) comprised of images from aMCI patients who had taken part in a Phase III clinical trial which determined the proportions of normal and abnormal images and the prediction of future clinical progression relative to amyloid status [47]. In both studies, patients received a single dose of approximately 185 MBq (range 166–203 MBq) of [ $^{18}\text{F}$ ]flutemetamol with image acquisition starting ~ 90 min (range 85–95 min) after injection, and collected  $6 \times 5$  min frames. Baseline clinical features of the 80 aMCI subjects included measurement of MMSE (mean = 27), CDR (0.5) and, Activities of Daily Living (mean = 74) [47].

#### Software packages

Four regulatory approved software packages (see Table 1) were used to generate composite SUVr from [ $^{18}\text{F}$ ]flutemetamol:

- MIM Software's *MIMneuro* ([https://www.mimsoftware.com/nuclear\\_medicine/mim\\_neuro](https://www.mimsoftware.com/nuclear_medicine/mim_neuro))
- Syntermed's *NeuroQ* (<https://www.syntermed.com/neuroq>)
- Hermes Medical Solutions' *BRASS* (<https://www.hermesmedical.com/neurology/>)
- GE Healthcare's *CortexID* (<https://www.gehealthcare.com/courses/aw-cortex-id>)

#### Normative database demographics

- *Cortex ID* The [ $^{18}\text{F}$ ]flutemetamol normal database of 100 amyloid negative scans spans the age range from 30 to 85 years with the majority of the subjects being 55 years old and above. Amyloid load as measured by amyloid PET in amyloid negative controls has a very weak association with subject age and thus no age correction of the normal database was deemed necessary
- *BRASS* database of 80 subjects is a subset of the 100 contained in *CortexID*
- *NeuroQ* normal data base consists of [ $^{18}\text{F}$ ]flutemetamol scans from 25 cognitively unimpaired subjects (10 < 55 years and 15 > 55 years) [46]
- *MIMneuro* The [ $^{18}\text{F}$ ]flutemetamol normal database contains 54 exams from AIBL. Exams were classified as normal according to AIBL criteria and have a

negative amyloid scan upon visual assessment. Ages span from 60 to 84 years old.

#### Image processing

All images were processed within the graphical user interfaces (GUI) of:

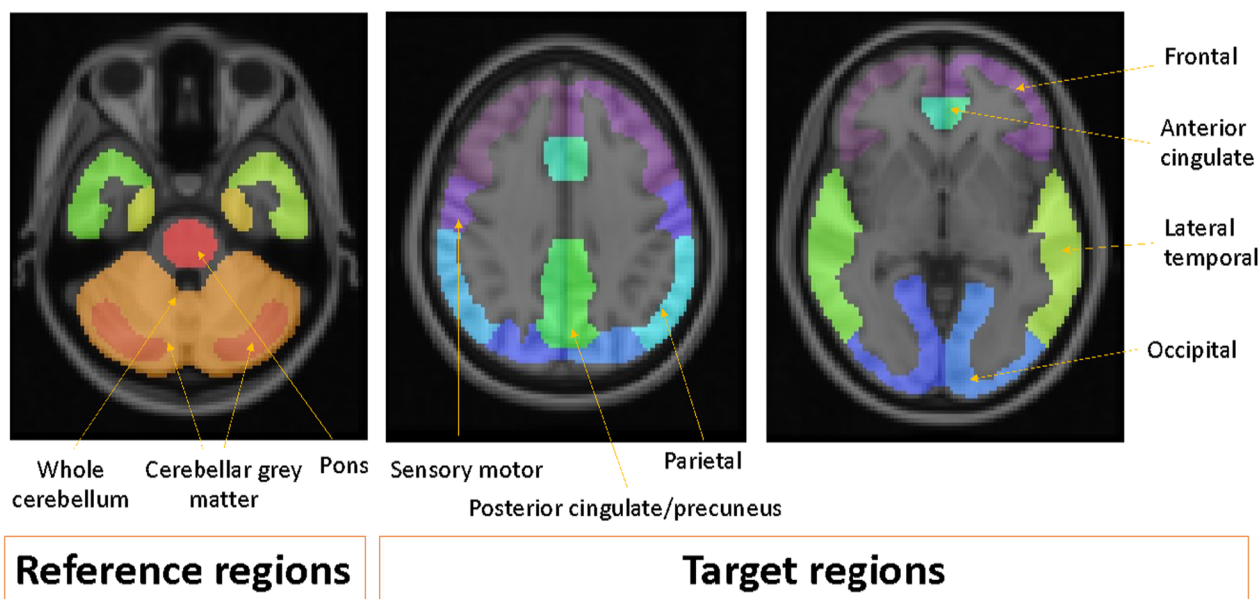
- *BRASS* using *v2.10.1.0*, by HP/PT (GE Healthcare, Amersham, UK). The *BRASS* GUI loads in DICOM folders, the correct tracer/reference region must be selected and then registration initiated. Registrations were visually checked and quantitative results were then exported for analysis.
- *CortexID* using *v2.1 Ext. 6*, by VP (GE Healthcare, Marlborough, USA). The *CortexID* GUI requires DICOM import, image registration automatically follows and quality was checked visually. SUVr results were then exported for analysis.
- *MIMneuro* using *vMIM-7.2.1 LA21-01*, by WB/KH (MIMSoftware, Ohio, USA). The *MIMneuro* GUI requires DICOM import. The tracer is detected automatically from DICOM headers, and the reference region is selected based on the tracer. Registration is both rigid and deformable. Registrations were checked visually and results exported for analysis.
- *NeuroQ* using *v3.80*, by HP/PT (GE Healthcare, Amersham, UK). The *NeuroQ* GUI requires import of patient's DICOM, and selection of reference region and appropriate tracer database. Both rigid and non-linear registration were then performed and visually quality checked. Next, composite SUVr was generated and databased.

For the purposes of both the pilot and validation phases, the pons was used as the reference region for all measures as it is an autopsy validated reference region where the positivity threshold has high concordance with a large cohort of visually inspected [ $^{18}\text{F}$ ]flutemetamol images [46, 49]. Composite SUVr was generated from cortical volumes of interest of the following regions: prefrontal, anterior cingulate, precuneus/posterior cingulate, parietal, lateral/mesial temporal, occipital, sensorimotor, see Fig. 1. This figure displays the cortical mask developed by Thurfjell et al. [49] for the purpose of an optimised quantitation of [ $^{18}\text{F}$ ]flutemetamol. All reference regions are displayed but, as previously mentioned, only the pons was used as the reference region for analyses in the present study paper; although the SUVr results for both whole cerebellum and cerebellar grey are minimally less (by 1–4%) when compared against autopsy verified images and those with visual inspection results [49].

**Table 1** Summary of four regulatory approved software packages and associated features

	MIM software's MIMneuro	Syntermed's NeuroQ	Hermes medical solutions' BRASS	GE healthcare's Cortex/D
CE/FDA status	CE Mark and FDA 510(k)	CE Mark and FDA 510(k)	FDA 510(k)	FDA 510(k)
Imaging requirements	PET, optional MRI or CT	PET, optional MR or CT	PET, optional MRI or CT	PET, optional MRI or CT
Normative database	54 controls	25 controls	80 controls [48]	> 100 controls
QC procedure	Visual QC for registration quality			
Partial volume correction	Not performed			
Spatial normalisation	Standard space (MNI)			
Attenuation correction	Required for image acquisition and processing			
Reference regions	Customizable e.g. pons, whole cerebellum, cerebellar cortex	Pons, whole cerebellum	Pons, whole cerebellum, cerebellar cortex	Pons, whole cerebellum, cerebellar cortex
Target regions available in addition to harmonised mask	Customizable e.g. Left/right hemisphere for: anterior cingulate gyrus, inferior medial frontal gyrus, lateral temporal lobe, posterior cingulate gyrus, precuneus, and superior parietal lobule	Left/right hemisphere for all (except midbrain and vermis); frontal cortex, sensorimotor cortex, Broca's region, anterior cingulate, posterior cingulate, caudate nucleus, lentiform nucleus, thalamus, parietal cortex, parietotemporal cortex, lateral temporal cortex, medial temporal cortex, primary visual cortex, associative visual cortex, midbrain, and vermis	Left/right hemisphere for all: frontal cortex, anterior cingulate, occipital cortex, parietal cortex, lateral temporal cortex, precuneus/posterior cingulate,	Left/right hemisphere for all: Prefrontal, anterior cingulate, precuneus/posterior cingulate, parietal, lateral/mesial temporal, occipital, sensorimotor
Metrics reported	SUVr, z-score, Centiloid*	SUVr, z-score	SUVr, z-score	SUVr, z-score

\*May not be approved in all jurisdictions



**Fig. 1** Representation of the cortical mask optimised for the measurement of [<sup>18</sup>F]Flutemetamol

**Pilot data processing**

Using a pilot dataset of 11 patients with varying amyloid load obtained from the aforementioned Phase II ALZ201 study [46], an estimate of concordance (percentage agreement around  $\geq 0.6$  SUV<sub>r</sub> positivity threshold) was obtained using a composite SUV<sub>r</sub> generated from the pre-existing cortical masks of each of the four software packages. In this pilot phase, varying agreement was observed across the software packages, likely due to heterogeneous cortical masks. Therefore, the cortical mask used to generate the composite cortical SUV<sub>r</sub> for *CortexID* was shared with all other software packages, with the aim of harmonising composite measures of amyloid burden. Once shared, the cortical mask was implemented into the pipelines of all software packages, and the pilot data was reassessed. Design of the mask is based on previous autopsy validated work [49] where care was taken to minimise any interference from white matter signal, which may potentially compromise any resulting SUV<sub>r</sub> measures. The *CortexID* cortical mask used for this harmonisation exercise is now available for other pipelines to use and can be obtained by contacting either GE Healthcare or the corresponding author.

**Validation data processing**

Upon completion of pilot testing it was agreed with all vendors to further assess a larger validation data set. Composite SUV<sub>r</sub> was generated with all four software packages from reconstructed and attenuation corrected images from 80 aMCI patients of varying amyloid load [47]. Images were checked visually for quality of

registration and segmentation. For the purpose of this project, the pons was used as the reference region for all measures as this has previously been shown a stable reference region [46]. Additionally, SUV<sub>r</sub> measures using the pons have been added to the [<sup>18</sup>F]flutemetamol summary of product characteristics in the EU ([https://www.ema.europa.eu/en/documents/product-information/vizamyl-epar-product-information\\_en.pdf](https://www.ema.europa.eu/en/documents/product-information/vizamyl-epar-product-information_en.pdf)). Therefore, use of the pons as a reference region allows the development of datasets which are consistent with recommended routine clinical use.

**Statistical analysis**

All 80 aMCI images were processed and quantified using four software packages. The composite SUV<sub>r</sub> values were analysed to assess compatibility. Percentage agreement was calculated across all software packages using an A $\beta$  positivity threshold of  $\geq 0.6$  SUV<sub>r</sub>, derived from previous autopsy confirmation work [49]. In order to assess reliability among the software packages, group-wise correlation (intraclass correlation coefficient, ICC) on composite SUV<sub>r</sub> was measured for all software packages combined. Kappa scores were calculated to assess inter-rater reliability of binary clinical decision between each pair of software packages (Cohen’s) and group-wise (Fleiss’). Company names have been blinded when reporting results in order to avoid any bias as this was a standardisation exercise and not competitive positioning. Finally, Bland–Altman plots were generated to visually compare the

agreement per patient between each pair of software packages. Statistics were performed using R for Mac version 1.2.5036.

**Assessment of SUVR(pons) values relative to visual inspection**

Majority visual read data (from 5 blinded readers) from a previous study [46] was available for the 80 aMCI cases examined in this study. The agreement rates between the majority and individual readers relative to the SUVR (pons) results from the 4 software tools, once the cortical masks had been implemented, is reported since this method of interpretation is also recommended in the European Summary of Product Characteristics (SmPC).

**Results**

As this study has been a collaborative and non-competitive effort to promote standardisation of [<sup>18</sup>F]flutemetamol analysis software names have been anonymised when reporting results.

**Pilot data compatibility**

Initial analysis of the composite SUVR for the 11 patients in the pilot phase showed good agreement between two of the software packages and marginally less with the other two, see Fig. 2 for graph and Table 2 further below for numeric results. The cortical masks from Softwares 1 and 3 were more heterogeneous, which lead to some

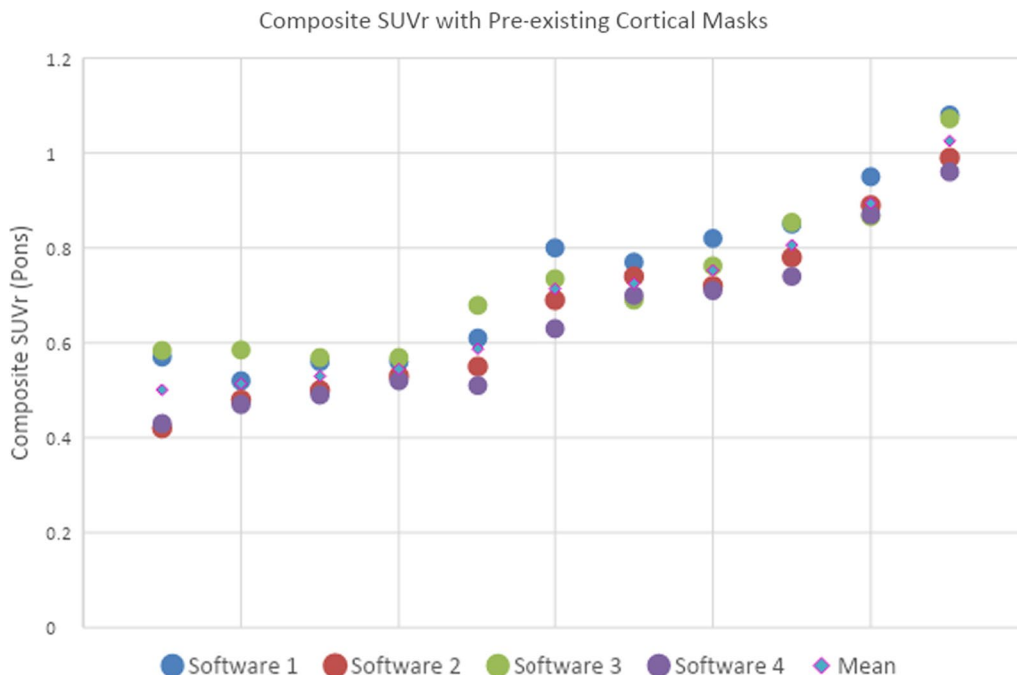
differences in the agreement between the composite SUVR, see Table 2 below. Subsequently, the cortical mask from *CortexID* was shared with all vendors, incorporated into the respective analysis pipelines and results reprocessed (see Fig. 3 and Table 2). Analysis of all four software packages found a 95% agreement around a  $\geq 0.6$  SUVR positivity threshold.

**Pilot data with harmonized cortical mask**

Figure 3 shows the agreement across the 4 software packages for the composite SUVR once the cortical masks had been harmonised across processing pipelines. An overall percentage agreement of 98% was calculated around an A $\beta$  positivity threshold  $\geq 0.6$  SUVR.

Table 2 shows the composite SUVR for all software packages using both the original and the harmonised cortical masks. In this pilot phase, good initial agreement was observed between two software packages' composite SUVR (mean SUVR difference of -0.024). However, harmonising the cortical mask across all four vendors improved consistency by reducing the difference in mean composite SUVR for:

- Software 3 vs Software 2
  - original masks = 0.061
  - harmonized masks = -0.003, a 0.064 reduction in mean composite SUVR difference

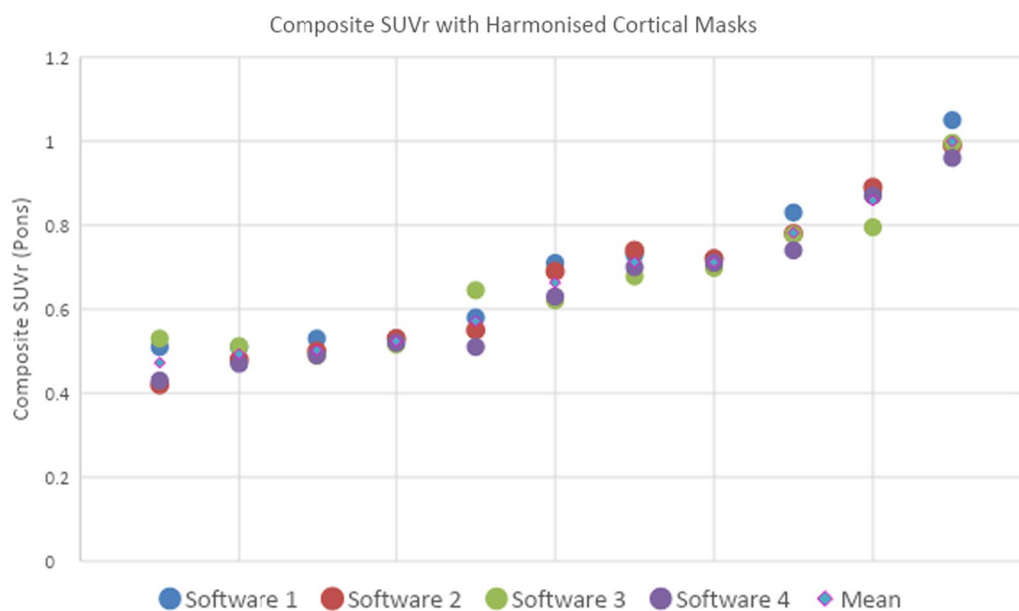


**Fig. 2** Composite SUVR of 11 patients (4 dots per patient) from pilot data using the pons as reference region with pre-existing cortical masks for all four software packages, the data is ranked by the patient's mean SUVR across the four software packages

**Table 2** Mean composite SUVr (n = 11 pilot patients) of four regulatory approved software packages with the original cortical masks, harmonised cortical masks, followed by the mean difference ( $\Delta$ ) between Software 3/Software 1 and Software 2 with original and harmonised cortical masks

	SW 2 Original	SW 4 Original	SW 4 $\Delta$	SW 2 Original	SW 3 Original	SW 3 $\Delta$	SW 3 Harmonised	SW 3 $\Delta$	SW 2 Harmonised	SW 1 Original	SW 1 $\Delta$	SW 1 Harmonised	SW 1 $\Delta$	SW 2 Harmonised
Mean composite SUVr	0.663	0.639	-0.024	0.724	0.061	0.660	-0.003	0.735	0.073	0.689	0.026			

SW software. Software names are anonymised due to proprietary nature of the software and potential commercial implications



**Fig. 3** Composite SUVR of 11 patients from pilot data using the pons as reference region with harmonised cortical mask for all four software packages, the data is ranked by the patient’s mean SUVR across the four software packages

- Software 1 vs Software 2
  - original masks = 0.073
  - harmonized masks = 0.036, a 0.047 reduction in mean composite SUVR difference

Following this improvement step, it was agreed with all vendors to further assess a larger ( $n=80$ ) validation data set (see “Validation data”).

Note, the overall the net variation of the harmonised quantitation values (Table 2) are at a similar level to that observed in test re-test studies, which was 0.9–3.8% [46].

**Validation data**

**Reliability**

The average measure ICC was 0.97, 95% confidence interval from 0.957 to 0.979, denoting ‘excellent’ reliability between composite SUVR measurements for all four software packages, see Fig. 4 showing boxplots for each software.

**Kappa scores around binary threshold classification**

All kappa scores around the  $\geq 0.6$  SUVR A $\beta$  positivity threshold, both combined (Fleiss’) and individual pairings (Cohen’s), were  $\geq 0.9$  signifying “almost perfect” inter-rater reliability. Fleiss’ Kappa score for the four software packages together = 0.95. See Table 3 for the pair-wise Cohen’s Kappa score for each of the software package combinations.

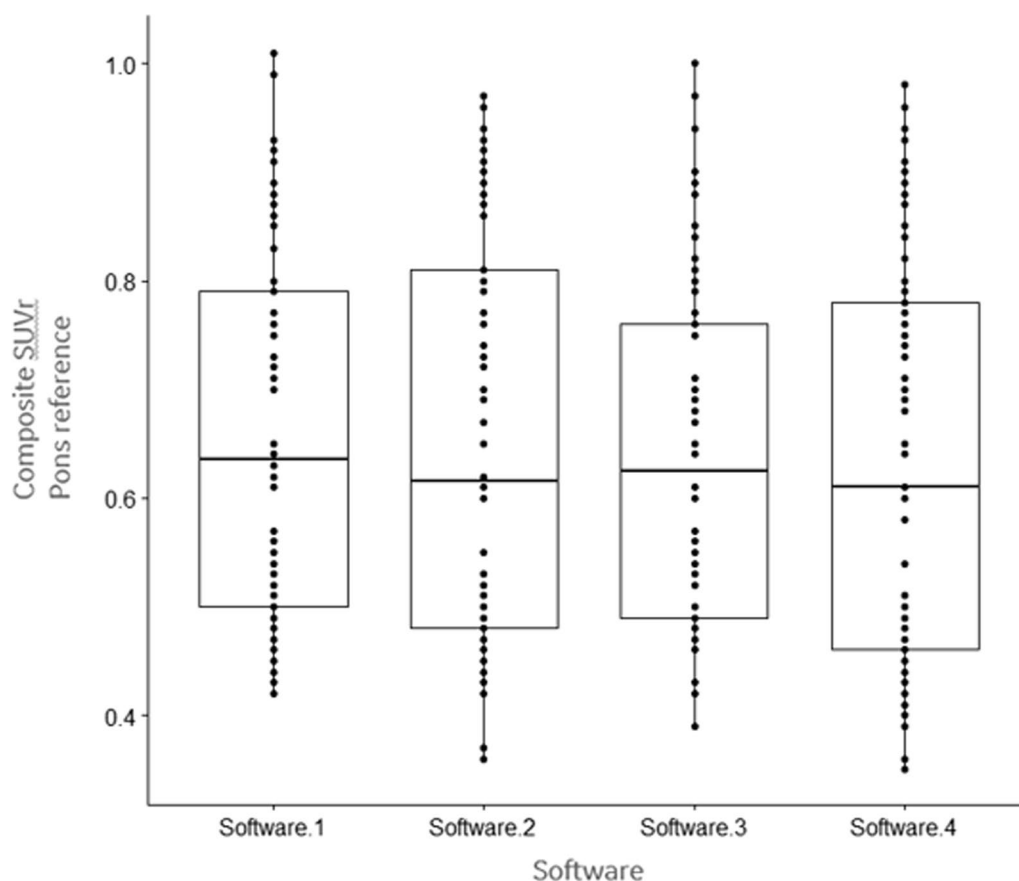
**Agreement**

Using an A $\beta$  positivity threshold of  $\geq 0.6$  SUVR [49], 95% agreement was achieved across the software packages. Two patients were narrowly classed as *negative* by one software package but *positive* by the others, and two patients vice versa, see Table 4 for composite SUVR of the 4 discordant patients.

**Assessment of SUVR(pons) values relative to visual inspection**

Majority visual read data (from 5 blinded readers) from a previous study [46] was available for the 80 aMCI cases examined in the present study. Visual read Kappa score was 0.89 (95% CI 0.82–0.96). By majority read, 43 cases were positive and 37 were negative. 72/80 cases were concordant by all 5 readers; some discordancy was observed in 8 cases (7 negatives with 2 calling positive in 3 cases and 1 calling positive in 4 cases, and a single positive case where one reader called a negative). In the present study, SUVR(pons) at a threshold of 0.6 differentiated the visual read cases in a dichotomous pattern, where only 3/80 for software 3 and  $\times 1/80$  cases for software 4 were discordant, and thus where the visual inspection and software result might have led to a discordant analysis. In cases such as these it is recommended that in clinical use there is a subsequent reassessment of the image more closely to see if there are any potential artefacts (e.g. image atrophy, lesions in the reference region or ROI) that may support whether the reader finalises their result based upon either the visual or quantitative analysis.





**Fig. 4** Boxplot showing composite SUVR using the pons as a reference region for each of the 4 software packages analysed

**Table 3** Cohen’s Kappa score for all possible software pairs

Software package pair	Cohen’s Kappa
1 vs 2	1
1 vs 3	0.92
1 vs 4	0.97
2 vs 3	0.92
2 vs 4	0.97
3 vs 4	0.90

Software names are anonymised due to proprietary nature of the software and potential commercial implications

**Table 4** Summary of four discordant patients around a  $\geq 0.6$  SUVR positivity threshold, bold values indicate results disagreeing on amyloid status with the other three software packages

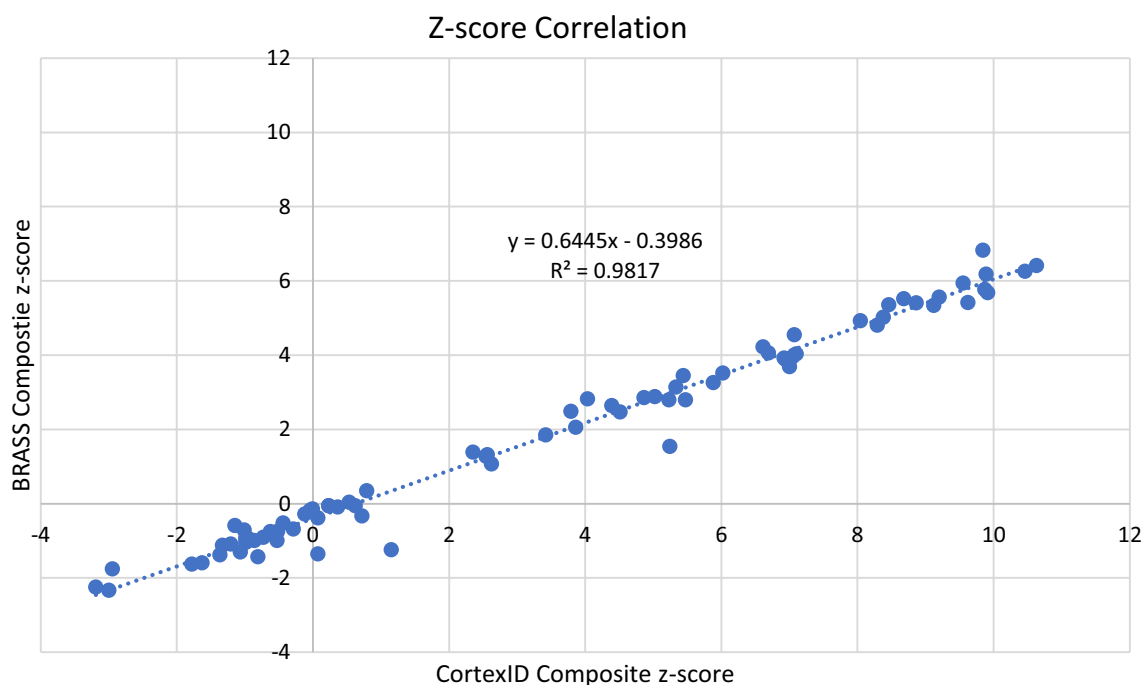
Discordant patient	Composite SUVR			
	Software 1	Software 2	Software 3	Software 4
#1	0.61	0.61	<b>0.57</b>	0.61
#2	0.51	0.48	<b>0.65</b>	0.46
#3	0.62	0.62	0.61	<b>0.58</b>
#4	0.52	0.51	<b>0.64</b>	0.51

**Z-score analysis**

Of the four software packages assessed in this paper, only Cortex ID and Hermes report a composite Z-score. The resulting correlation coefficient analysis demonstrates a strong ( $r^2=0.98$ ) linear relationship between the composite z-score analysis of the 80 aMCI subjects for the two software packages, see Fig. 5.

**Bland–Altman plots**

Bland–Altman plots display the relationship between two paired variables using the same scale, i.e. composite SUVR. The black dots show the average measurement of the 2 software packages in question, the black line shows the average difference in measurements between the two software packages and the red dotted lines show the upper and lower limits of the 95% confidence interval (CI) for the average difference between the two software packages. Figure 6 shows Bland–Altman plots for the highest (Software 1 and 2) and lowest (Software 3 and 4) agreement scores, according to Cohen’s kappa. The top plot shows tighter 95% CI and



**Fig. 5** Z-score correlation analysis between the only two software packages assessed in this paper which offer composite z-scores

smaller average difference in measurements between the two software packages.

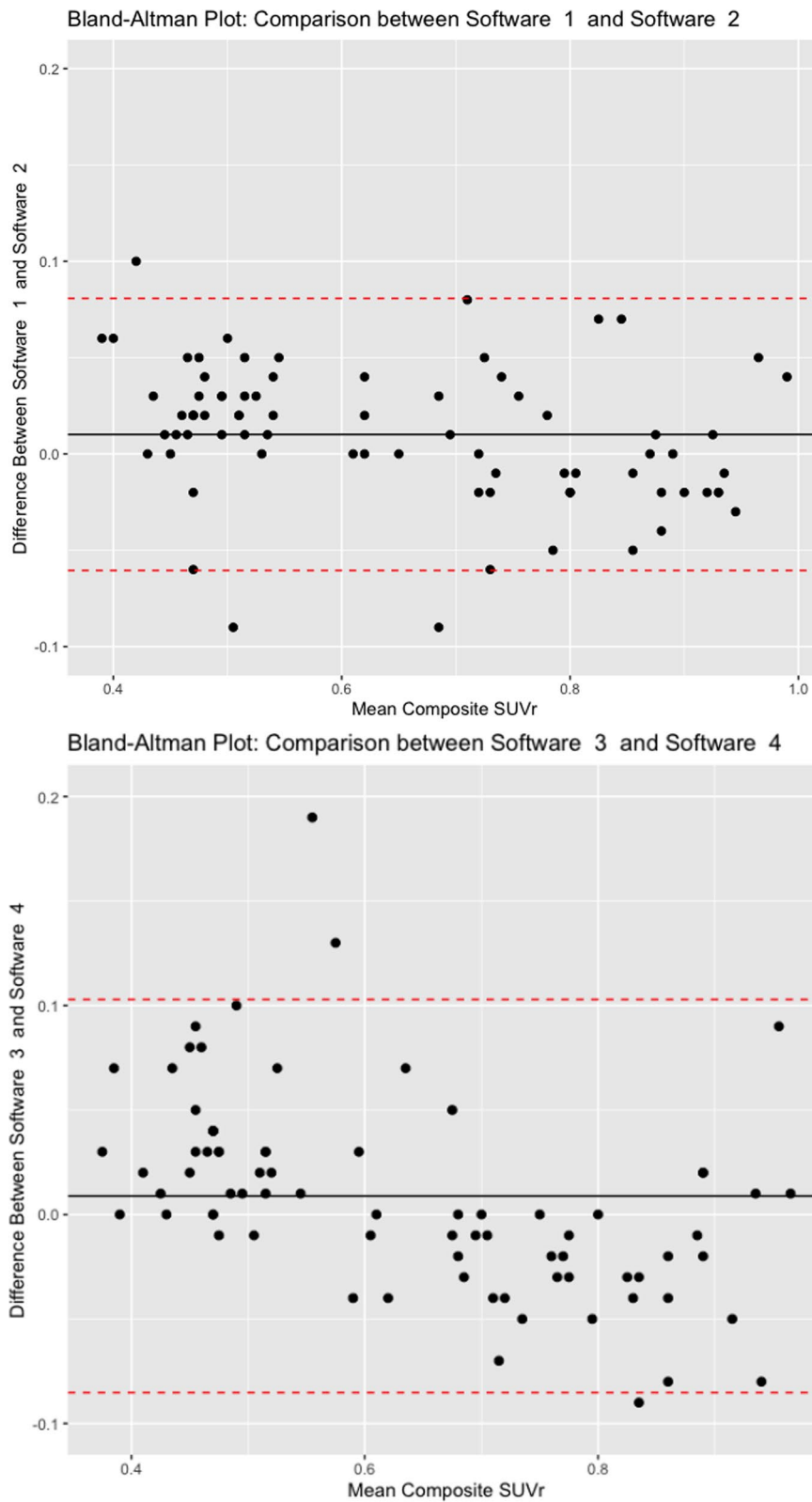
### Discussion

In this work, four regulatory approved software packages have generated highly correlated and reliable quantification (composite SUVr) of [ $^{18}\text{F}$ ]flutemetamol amyloid PET around a  $\geq 0.6$  SUVr  $\text{A}\beta$  positivity threshold [49], using the pons as a reference region. All kappa scores around this threshold, both combined (Fleiss') and individual pairings (Cohen's), were  $\geq 0.9$  signifying "almost perfect" inter-rater reliability; the average measure ICC was 0.97. Strong concordance was achieved through harmonisation of cortical masks for generating quantitative amyloid PET results. This was possible due to the collaborative efforts among competing software vendors, all of whom have now implemented to the cortical mask examined in the present study.

The clinical standard of binary classification through visual assessment has been demonstrated as approximately 90% accurate in advanced clinical and end-of-life patients [1, 2, 4]. This provides useful stratification of amyloid status for research, clinical trials and routine clinical practice. However, visual assessment can be challenged in a heterogeneous clinical population. For example, cortical thinning or atrophy can be compounded by partial volume effects, which subsequently raises the question of performing partial volume correction (PVC),

or not. There is no consensus in the field regarding this issue; some recent evidence suggests an increase in sensitivity for detecting early stage cerebral amyloidosis when using PVC [50]. However, other studies comparing various methods have proven inconclusive [51, 52]. It is worth noting none of the software packages in the study currently perform PVC. Comorbidities further complicating visual assessments include normal pressure hydrocephalus [53] or other neurodegenerative disorders [13, 37, 54–56]. Adjunct quantitative measures of  $\text{A}\beta$  deposition, such as those examined in this study, can provide greater clinical utility in addition to current dichotomous classification and contribute to improvements in diagnostic confidence [8, 10, 21], prediction of cognitive decline [28–31] and changes to diagnosis [16] and patient management [32–37].

Previous work has been carried out comparing regulatory approved software packages (Hermes and Syngo.via) on 225 subjects (probable AD, MCI, controls), showing high sensitivity and specificity for both [47]. Most recently, Syngo.via, CortexID and PMOD were assessed on 195 patients with cognitive impairment, marginally different positivity thresholds were noted along with very high correlation between different software and normalization methods [57]. The present study has compared twice as many software packages approved for clinical use (i.e. FDA 510(k) cleared and/or CE-marked and demonstrated very strong concordance across all four



**Fig. 6** Bland–Altman plots for the highest (Software 1 and 2) and lowest (Software 3 and 4) agreement scores, according to Cohen’s kappa. Y-axis shows the difference between the composite SUVr for each software package and the x-axis shows the mean composite SUVr of the two software packages

applications. The focus for the larger validation set was on clinically relevant subjects (i.e. aMCI patients), in line with the amyloid PET appropriate use criteria [58], i.e. patients with “persistent or progressive unexplained MCI”, and perhaps more likely to benefit from amyloid PET in the earlier stages of cognitive impairment. It is worth noting that visual inspection, even after quantification, is still recommended in order to assess cases which may have atrophy and where quantification may potentially be compromised.

As noted in Table 1, the software packages have a variety of features and image processing differences. While the results were highly compatible across all four software packages following harmonisation of the cortical mask, there was disagreement in 5% of cases ( $n=4$ ) where two patients were narrowly classed as positive by one software package but negative by the others, and two patients vice versa. Possible reasons for these minor discrepancies are differences in spatial normalization and registration steps carried out by each vendor. In addition, despite harmonisation of the cortical regions, the pons reference region may not be equally harmonized, and thus could contribute to the minor discrepancies observed. However, as a high-uptake region, the pons is likely to be more robust to such variations than other reference regions. However, the reliability results from this validation exercise (Table 2) demonstrate that the variation is approximately comparable to that observed in test–retest [46] and that measurement of [ $^{18}\text{F}$ ]flutemetamol using a harmonised cortical mask would not influence analysis above that observed for test–rest.

Only Cortex ID and Hermes report a composite Z-score, the correlation coefficient analysis was strong (see Fig. 5). It is worth noting that the threshold between negative and positive scans for the two packages differs: Cortex ID has a threshold of approximately 2 whilst that of Hermes is lower at approximately 1.5. These differences are likely due to the composition and size of the normative databases, as shown in Table 1.

### Limitations

Ideally, all four software packages would have been installed on the same workstation and results independently generated. However, due to the proprietary nature of the software this was not possible and three of the software packages were installed on different workstations at GE Healthcare with the final vendor generating their own results before sharing for group-wise analysis. In addition, only composite SUVr with the pons as reference region was assessed; use of other reference regions and quantitative metrics was beyond the scope of planned work.

Pons was used as the reference region in this analysis since the aim was to generate data to support the language in the European Summary of Product Characteristics (SmPC), which has recently been updated to add quantitation as an adjunct to visual inspection of [ $^{18}\text{F}$ ]flutemetamol images. Stated thresholds using pons as the reference region in the SMPC are quoted to be 0.59–0.61 and are derived from autopsy validated images using CERAD pathology as the standard of truth. Other reference regions can be used and in addition to pons but data showing the concordance between quantitation and visual inspection was 1–2% less than that derived from the pons [49].

It is also appreciated that other quantitative metrics, such as z-score, may be useful alternatives/adjuncts to the SUVr measure. A cortical z-score of over 2 is normally used to indicate whether the composite uptake of PET amyloid is abnormal [49]. Possibly more relevant for the z-score measure is to use this metric to assess early regional amyloid uptake when the composite measure is close to, or at, the threshold [59].

### Future directions

The Centiloid scale is a cross-tracer SUVr transformation which produces a single-figure for amyloid burden measure, which is expected to be consistent across tracers. While the method currently provides global rather than regional amyloid measures, clinical use of the Centiloid scale is increasing [59]. Therefore, as more software packages begin to offer this metric as part of the suite, similar compatibility analysis is encouraged. Further analysis of the more subtle differences between the processing pipelines of the four software packages in question was beyond the scope of this paper. However, additional analysis would be of interest to further elucidate the root cause of the minor discrepancies observed. It would also be of note to assess the compatibility of regional SUVr, in addition to the composite measures in the current paper. The z-score analysis in the paper encourages further investigation into the potential value of a consolidated normative database assessing the impact of database size and composition on the z-score threshold value.

### Conclusions

Regulatory approved and/or cleared software packages provide highly correlated and reliable quantification of [ $^{18}\text{F}$ ]flutemetamol amyloid PET based around a  $\geq 0.6$  SUVr positivity threshold, when using pons as a reference region. This concordance was achieved through collaboration between competing vendors, and the harmonisation of cortical masks for generating quantitative

amyloid PET results. Where possible, harmonisation of image processing steps is encouraged in order to facilitate clinical validation and widen adoption of clinically relevant quantitative measures, with the ultimate aim of enhancing consistency of image interpretation leading to accurate diagnosis and management decisions in patients with AD pathology.

#### Acknowledgements

The authors would like to thank the investigators and patients from which the images were sourced for this retrospective analysis.

#### Author contributions

HP, CBU, MB, VP, PT, DC, WB, KH, JL, CBr and GF contributed to the study conception and design. Material and data preparation were performed by HP, CBU, MB, VP, PT, DC, WB, KH, JL, CBr, AB and GF. The first draft of the manuscript was written by HP and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

#### Funding

No specific funding.

#### Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

#### Declarations

##### Ethics approval and consent to participate

All patients included in this analysis had previously participated in Institutional Review Board (IRB), Independent Ethics Committee (IEC)-approved studies (National Clinical Trial numbers—NCT01028053/NCT00785759/NCT01672827) and they (or their legally authorized representative) provided informed written consent to participate; data usage was considered to be covered by the previous consent. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

##### Consent for publication

Written informed consent was obtained from the patient for publication of this study and accompanying images.

##### Competing interests

HP, CBr, MB, VP, CBU, PT and GF are all full-time employees of GE Healthcare. DC is a full-time employee of Syntermed. WB and KH are full time employees of MIM Software Inc. JL is a full-time employee of HERMES Medical Solutions. AB declares that they have no competing interests.

##### Author details

<sup>1</sup>GE Healthcare, Pollards Wood, Chalfont St Giles, Amersham HP8 4SP, UK. <sup>2</sup>Centre for Medical Image Computing (CMIC), Department of Medical Physics and Bioengineering, University College London, London, UK. <sup>3</sup>UCL Queen Square Institute of Neurology, University College London, London, UK. <sup>4</sup>Clinical Memory Research Unit, Department of Clinical Sciences, Lund University, Malmö, Sweden. <sup>5</sup>Syntermed, Atlanta, GA, USA. <sup>6</sup>MIM Software, Cleveland, OH, USA. <sup>7</sup>HERMES Medical Solutions, Stockholm, Sweden.

Received: 26 September 2022 Accepted: 5 May 2023

Published online: 24 May 2023

#### References

- Salloway S, Gamez JE, Singh U, 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. *Alzheimer's Dement Diagnosis, Assess Dis Monit.* 2017;9:25–34. <https://doi.org/10.1016/j.dadm.2017.06.001>.
- Sabri O, Sabbagh MN, Seibyl J, et al. Florbetaben PET imaging to detect amyloid beta plaques in Alzheimer's disease: Phase 3 study. *Alzheimer's Dement.* 2015;11:964–74. <https://doi.org/10.1016/j.jalz.2015.02.004>.
- Clark CM, Pontecorvo MJ, Beach TG, et al. Cerebral PET with florbetapir compared with neuropathology at autopsy for detection of neuritic amyloid- $\beta$  plaques: a prospective cohort study. *Lancet Neurol.* 2012;11:669–78. [https://doi.org/10.1016/S1474-4422\(12\)70142-4](https://doi.org/10.1016/S1474-4422(12)70142-4).
- Buckley CJ, Sherwin PF, Smith APL, et al. Validation of an electronic image reader training programme for interpretation of [<sup>18</sup>F]flutemetamol  $\beta$ -amyloid PET brain images. *Nucl Med Commun.* 2017;38:234–41. <https://doi.org/10.1097/MNM.0000000000000633>.
- De Wilde A, Van Der Flier WM, Pelkmans W, et al. Association of amyloid positron emission tomography with changes in diagnosis and patient treatment in an unselected memory clinic cohort: the ABIDE project. *JAMA Neurol.* 2018;75:1062–70. <https://doi.org/10.1001/jamaneurol.2018.1346>.
- Rabinovici GD, Gatsonis C, Apgar C, et al. Association of amyloid positron emission tomography with subsequent change in clinical management among medicare beneficiaries with mild cognitive impairment or dementia. *J Am Med Assoc: JAMA.* 2019;321:1286–94. <https://doi.org/10.1001/jama.2019.2000>.
- Chiotis K, Saint-Aubert L, Boccardi M, et al. Clinical validity of increased cortical uptake of amyloid ligands on PET as a biomarker for Alzheimer's disease in the context of a structured 5-phase development framework. *Neurobiol Aging.* 2017;52:214–27. <https://doi.org/10.1016/j.neurobiolaging.2016.07.012>.
- Fantoni ER, Chalkidou A, O'Brien JT, et al. A systematic review and aggregated analysis on the impact of amyloid PET Brain imaging on the diagnosis, diagnostic confidence, and management of patients being evaluated for Alzheimer's disease. *J Alzheimer's Dis.* 2018;63:783–96. <https://doi.org/10.3233/JAD-171093>.
- Barthel H, Sabri O. Clinical use and utility of amyloid imaging. *J Nucl Med.* 2017;58:1711–7. <https://doi.org/10.2967/jnumed.116.185017>.
- Grundman M, Johnson KA, Lu M, et al. Effect of amyloid imaging on the diagnosis and management of patients with cognitive decline: impact of appropriate use criteria. *Dement Geriatr Cogn Disord.* 2016;41:80–92. <https://doi.org/10.1159/000441139>.
- Collij LE, Salvadó G, Shekari M, et al. Visual assessment of [<sup>18</sup>F]flutemetamol PET images can detect early amyloid pathology and grade its extent. *Eur J Nucl Med Mol Imaging.* 2021;48:2169–82. <https://doi.org/10.1007/s00259-020-05174-2>.
- Zwan MD, Bouwman FH, Konijnenberg E, et al. Diagnostic impact of [<sup>18</sup>F]flutemetamol PET in early-onset dementia. *Alzheimer's Res Ther.* 2017. <https://doi.org/10.1186/s13195-016-0228-4>.
- Pontecorvo MJ, Siderowf A, Dubois B, et al. Effectiveness of florbetapir PET imaging in changing patient management. *Dement Geriatr Cogn Disord.* 2017;44:129–43. <https://doi.org/10.1159/000478007>.
- Schipke CG, Peters O, Heuser I, et al. Impact of beta-amyloid-specific florbetaben pet imaging on confidence in early diagnosis of Alzheimer's disease. *Dement Geriatr Cogn Disord.* 2012;33:416–22. <https://doi.org/10.1159/000339367>.
- Zannas AS, Doraiswamy PM, Shpanskaya KS, et al. Impact of 18F-florbetapir PET imaging of  $\beta$ -amyloid neuritic plaque density on clinical decision-making. *Neurocase.* 2014;20:466–73. <https://doi.org/10.1080/13554794.2013.791867>.
- Leuzy A, Savitcheva I, Chiotis K, et al. Clinical impact of [<sup>18</sup>F]flutemetamol PET among memory clinic patients with an unclear diagnosis. *Eur J Nucl Med Mol Imaging.* 2019. <https://doi.org/10.1007/s00259-019-04297-5>.
- Van Der Flier WM, Scheltens P. Amsterdam dementia cohort: performing research to optimize care. *J Alzheimer's Dis.* 2018;62:1091–111. <https://doi.org/10.3233/JAD-170850>.
- Fantoni E, Collij L, Alves IL, et al. The spatial-temporal ordering of amyloid pathology and opportunities for PET imaging. *J Nucl Med.* 2020;61:166–71. <https://doi.org/10.2967/jnumed.119.235879>.
- Pontecorvo MJ, Arora AK, Devine M, et al. Quantitation of PET signal as an adjunct to visual interpretation of florbetapir imaging. *Eur*

- J Nucl Med Mol Imaging. 2017;44:825–37. <https://doi.org/10.1007/s00259-016-3601-4>.
20. Joshi AD, Pontecorvo MJ, Clark CM, et al. Performance characteristics of amyloid PET with florbetapir F 18 in patients with Alzheimer's disease and cognitively normal subjects. *J Nucl Med*. 2012;53:378–84. <https://doi.org/10.2967/jnumed.111.090340>.
  21. Bucci M, Savitcheva I, Farrar G, et al. A multisite analysis of the concordance between visual image interpretation and quantitative analysis of [18F]flutemetamol amyloid PET images. *Eur J Nucl Med Mol Imaging*. 2021;48:2183–99. <https://doi.org/10.1007/s00259-021-05311-5>.
  22. Paghera B, Altomare D, Peli A, et al. Comparison of visual criteria for amyloid-pet reading: could criteria merging reduce inter-rater variability? *Q J Nucl Med Mol Imaging*. 2021;64:414–21. <https://doi.org/10.23736/51824-4785.19.03124-8>.
  23. Yamane T, Ishii K, Sakata M, et al. Inter-rater variability of visual interpretation and comparison with quantitative evaluation of 11C-PiB PET amyloid images of the Japanese Alzheimer's Disease Neuroimaging Initiative (J-ADNI) multicenter study. *Eur J Nucl Med Mol Imaging*. 2017;44:850–7. <https://doi.org/10.1007/s00259-016-3591-2>.
  24. Lopresti BJ, Klunk WE, Mathis CA, et al. Simplified quantification of Pittsburgh Compound B amyloid imaging PET studies: a comparative analysis. *J Nucl Med*. 2005;46:1959–72.
  25. Aisen PS, Cummings J, Doody R, et al. The future of anti-amyloid trials. *J Prev Alzheimer's Dis*. 2020;7:146–51. <https://doi.org/10.14283/jpad.2020.24>.
  26. Blennow K, Zetterberg H. Amyloid and Tau biomarkers in CSF. *J Prev Alzheimer's Dis*. 2015;2:1–5. <https://doi.org/10.14283/jpad.2015.41>.
  27. Milà-Alomà M, Salvadó G, Shekari M, et al. Comparative analysis of different definitions of amyloid- $\beta$  positivity to detect early downstream pathophysiological alterations in preclinical Alzheimer. *J Prev Alzheimer's Dis*. 2021;8:68–77. <https://doi.org/10.14283/jpad.2020.51>.
  28. Farrell ME, Jiang S, Schultz AP, et al. Defining the lowest threshold for amyloid-PET to predict future cognitive decline and amyloid accumulation. *Neurology*. 2021;96:e619–31. <https://doi.org/10.1212/WNL.00000000000011214>.
  29. Farrell ME, Chen X, Rundle MM, et al. Regional amyloid accumulation and cognitive decline in initially amyloid-negative adults. *Neurology*. 2018;91:E1809–21. <https://doi.org/10.1212/WNL.0000000000006469>.
  30. van der Kall LM, Truong T, Burnham SC, et al. Association of  $\beta$ -amyloid level, clinical progression, and longitudinal cognitive change in normal older individuals. *Neurology*. 2021;96:e662–70. <https://doi.org/10.1212/WNL.00000000000011222>.
  31. Hanseeuw BJ, Malotau V, Dricot L, et al. Defining a centiloid scale threshold predicting long-term progression to dementia in patients attending the memory clinic: an [18F] flutemetamol amyloid PET study. *Eur J Nucl Med Mol Imaging*. 2021;48:302–10. <https://doi.org/10.1007/s00259-020-04942-4>.
  32. Camus V, Payoux P, Barré L, et al. Using PET with 18F-AV-45 (florbetapir) to quantify brain amyloid load in a clinical environment. *Eur J Nucl Med Mol Imaging*. 2012;39:621–31. <https://doi.org/10.1007/s00259-011-2021-8>.
  33. Guerra UP, Nobili FM, Padovani A, et al. Recommendations from the Italian Interdisciplinary Working Group (AIMN, AIP, SINDEM) for the utilization of amyloid imaging in clinical practice. *Neurol Sci*. 2015;36:1075–81. <https://doi.org/10.1007/s10072-015-2079-3>.
  34. Kobylecki C, Langheinrich T, Hinz R, et al. 18F-florbetapir PET in patients with frontotemporal dementia and Alzheimer disease. *J Nucl Med*. 2015;56:386–91. <https://doi.org/10.2967/jnumed.114.147454>.
  35. Daniela P, Orazio S, Alessandro P, et al. A survey of FDG- and amyloid-PET imaging in dementia and grade analysis. *Biomed Res Int*. 2014. <https://doi.org/10.1155/2014/785039>.
  36. Klunk WE, Koeppe RA, Price JC, et al. The Centiloid project: standardizing quantitative amyloid plaque estimation by PET. *Alzheimer's Dement*. 2015;11:1-15.e4. <https://doi.org/10.1016/j.jalz.2014.07.003>.
  37. Akamatsu G, Ikari Y, Ohnishi A, et al. Voxel-based statistical analysis and quantification of amyloid PET in the Japanese Alzheimer's disease neuroimaging initiative (J-ADNI) multi-center study. *EJNMMI Res*. 2019. <https://doi.org/10.1186/s13550-019-0561-2>.
  38. Klunk WE, Engler H, Nordberg A, et al. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. *Ann Neurol*. 2004;55:306–19. <https://doi.org/10.1002/ana.20009>.
  39. Bourgeat P, Dore V, Fripp J, et al. Computational analysis of PET by AIBL (CapAIBL): a cloud-based processing pipeline for the quantification of PET images. *Med Imaging 2015 Image Process*. 2015;9413:94132. <https://doi.org/10.1117/12.2082492>.
  40. Markiewicz PJ, Ehrhardt MJ, Erlandsson K, et al. NiftyPET: a high-throughput software platform for high quantitative accuracy and precision PET imaging and analysis. *Neuroinformatics*. 2018;16:95–115. <https://doi.org/10.1007/s12021-017-9352-y>.
  41. Chincarini A, Sensi F, Rei L, et al. Standardized uptake value ratio-independent evaluation of brain amyloidosis. *J Alzheimer's Dis*. 2016;54:1437–57. <https://doi.org/10.3233/JAD-160232>.
  42. Buckley CJ, Foley C, Battle M, et al. AmyPype: an automated system to quantify AMYPAD's [18F]flutemetamol and [18F]florbetaben images including regional SUVR and Centiloid analysis. *Eur J Nucl Med Mol Imaging*. 2019;46:5323–4.
  43. Iaccarino L, La Joie R, Koeppe R, et al. rPOP: robust PET-only processing of community acquired heterogeneous amyloid-PET data. *Neuroimage*. 2022;246:118775. <https://doi.org/10.1016/j.neuroimage.2021.118775>.
  44. Pemberton HG, Zaki LAM, Goodkin O, et al. Technical and clinical validation of commercial automated volumetric MRI tools for dementia diagnosis—a systematic review. *Neuroradiology*. 2021;63:1773–89. <https://doi.org/10.1007/s00234-021-02746-3>.
  45. Smith AM, Obuchowski NA, Foster NL, et al. The RSNA QIBA profile for amyloid PET as an imaging biomarker for cerebral amyloid quantification. *J Nucl Med*. 2023;64:294–303. <https://doi.org/10.2967/jnumed.122.264031>.
  46. Vandenberghe R, Van Laere K, Ivanou A, et al. 18F-flutemetamol amyloid imaging in Alzheimer disease and mild cognitive impairment: a phase 2 trial. *Ann Neurol*. 2010;68:319–29. <https://doi.org/10.1002/ana.22068>.
  47. Wolk DA, Sadowsky C, Safirstein B, et al. Use of flutemetamol F 18-labeled positron emission tomography and other biomarkers to assess risk of clinical progression in patients with amnesic mild cognitive impairment. *JAMA Neurol*. 2018;75:1114–23. <https://doi.org/10.1001/JAMANEUROL.2018.0894>.
  48. Curry S, Patel N, Fakhry-Darian D, et al. Advances in neurodegenerative and psychiatric imaging special feature: Full paper: Quantitative evaluation of beta-amyloid brain PET imaging in dementia: a comparison between two commercial software packages and the clinical report. *Br J Radiol*. 2019. <https://doi.org/10.1259/bjr.20181025>.
  49. Thurfjell L, Lilja J, Lundqvist R, et al. 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. <https://doi.org/10.2967/jnumed.114.142109>.
  50. Teipel SJ, Dyrba M, Vergallo A, et al. Partial Volume correction increases the sensitivity of 18F-florbetapir-positron emission tomography for the detection of early stage amyloidosis. *Front Aging Neurosci*. 2021;13:846. <https://doi.org/10.3389/fnagi.2021.748198>.
  51. Shidahara M, Thomas BA, Okamura N, et al. A comparison of five partial volume correction methods for Tau and Amyloid PET imaging with [18F] THK5351 and [11C]PIB. *Ann Nucl Med*. 2017;31:563–9. <https://doi.org/10.1007/s12149-017-1185-0>.
  52. Schwarz CG, Gunter JL, Lowe VJ, et al. A comparison of partial volume correction techniques for measuring change in serial amyloid PET SUVR. *J Alzheimer's Dis*. 2019;67:181–95. <https://doi.org/10.3233/JAD-180749>.
  53. Rinne JO, Wong DF, Wolk DA, et al. Flutemetamol PET imaging and cortical biopsy histopathology for fibrillar amyloid  $\beta$  detection in living subjects with normal pressure hydrocephalus: pooled analysis of four studies. *Acta Neuropathol*. 2012;124:833–45. <https://doi.org/10.1007/s00401-012-1051-z>.
  54. Matsuda H, Ito K, Ishii K, et al. Quantitative evaluation of 18F-flutemetamol PET in patients with cognitive impairment and suspected Alzheimer's disease: a multicenter study. *Front Neurol*. 2021. <https://doi.org/10.3389/fneur.2020.578753>.
  55. Chételat G, Arbizu J, Barthel H, et al. Amyloid-PET and 18F-FDG-PET in the diagnostic investigation of Alzheimer's disease and other dementias. *Lancet Neurol*. 2020;19:951–62. [https://doi.org/10.1016/S1474-4422\(20\)30314-8](https://doi.org/10.1016/S1474-4422(20)30314-8).
  56. Ossenkoppele R, Jansen WJ, Rabinovici GD, et al. Prevalence of amyloid PET positivity in dementia syndromes: a meta-analysis. *JAMA*. 2015;313:1939–49. <https://doi.org/10.1001/jama.2015.4669>.
  57. Müller EG, Stokke C, Stokmo HL, et al. Evaluation of semi-quantitative measures of 18 F-flutemetamol PET for the clinical diagnosis of Alzheimer's disease. *Quant Imaging Med Surg*. 2022;12:493–509. <https://doi.org/10.21037/QIMS-21-188>.

58. Johnson KA, Minoshima S, Bohnen NI, 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*. 2013. <https://doi.org/10.1016/j.jalz.2013.01.002>.
59. Pemberton HG, Collij LE, Heeman F, et al. Quantification of amyloid PET for future clinical use: a state-of-the-art review. *Eur J Nucl Med Mol Imaging*. 2022;5:1–21. <https://doi.org/10.1007/S00259-022-05784-Y>.

### **Publisher's Note**

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

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)

---