Journal of Nuclear Medicine, published on October 12, 2018 as doi:10.2967/jnumed.118.211532

Assessing Amyloid Pathology in Cognitively Normal Subjects using [^{18}F]Flutemetamol PET: Comparing Visual Reads and Quantitative Methods

*Lyduine E. Collij, MSc[1], *Elles Konijnenberg, MD[2]; Juhan Reimand, MD[1,3,4], Mara ten Kate, MD[2]; Anouk den Braber, PhD[2,5]; Isadora Lopes Alves, PhD[1]; Marissa Zwan, PhD[2]; Maqsood Yaqub, PhD[1]; Daniëlle M.E. van Assema, MD, PhD[6]; Alle Meije Wink, PhD[1]; Adriaan A. Lammertsma, PhD.[1]; Philip Scheltens. MD, PhD[2]; Pieter Jelle Visser, MD, PhD[2]; Frederik Barkhof, MD, PhD[1,7] & Bart N.M. van Berckel, MD, PhD.[1]

[1]Dept. of Radiology and Nuclear Medicine, VU University Medical Center, Amsterdam, The Netherlands; [2]Alzheimer Center and Dept. of Neurology, VU University Medical Center, Amsterdam, The Netherlands; [3]Centre of Radiology, North Estonia Medical Centre, Tallinn, Estonia; [4]Department of Health Technologies, Tallinn University of Technology, Tallinn, Estonia; [5]Dept. of Biological Psychology, VU University Amsterdam, The Netherlands. [6]Dept. of Radiology and Nuclear medicine, Erasmus Medical Center, Rotterdam, the Netherlands; [7]Institute of Neurology and Healthcare Engineering, University College London, London, United Kingdom.

***Authors contributed equally**

Corresponding author:

Lyduine E. Collij
Department of Radiology and Nuclear Medicine, VU University Medical Center
De Boelelaan 1117, 1108 HV Amsterdam, The Netherlands
l.collij@vumc.nl
+31 620222419

Total words: 4993
Words in abstract: 235
Character count title: 135
Character count running title: 40
Number of references: 38
Number of tables: 1
Number of figures: 3

This project has received funding from the the EU/EFPIA Innovative Medicines Initiative (IMI) Joint Undertaking EMIF grant agreement (grant No115372) and the EU-EFPIA IMI-2 Joint Undertaking (grant No 115952). This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA. Supported by the NIHR UCLH biomedical research centre. This work also received in kind sponsoring of the PET-tracer from GE Healthcare.

ABSTRACT

Objective: Determine the optimal approach for assessing amyloid pathology in a cognitively normal elderly population.

Methods: Dynamic 18F-Flutemetamol PET scans acquired using a coffee-break protocol (0-30 and 90-110 min. scan) from 190 cognitively normal elderly (mean age 70.4 years, 60% female) were included. Parametric images were generated from standard uptake value ratio (SUVr) and non-displaceable binding potential (BP_{ND}) methods, with cerebellar grey matter as a reference region and were visually assessed by three trained readers. Inter-reader agreement was calculated using Kappa statistics and (semi-)quantitative values were obtained. Global cut-offs were calculated for both SUVr and BP_{ND} using a ROC analysis and the Youden Index. Visual assessment was related to (semi-)quantitative classifications.

Results: Inter-reader agreement in visual assessment was moderate for SUVr ($\kappa = 0.57$) and good for BP_{ND} images ($\kappa = 0.77$). There was discordance between readers for 35 cases (18%) using SUVr and for 15 cases (8%) using BP_{ND}, with 9 overlapping cases. For the total cohort, the mean (±SD) SUVr and BP_{ND} values were 1.33 (± 0.21) and 0.16 (± 0.12), respectively. Most of the 35 cases (91%) where SUVr image assessment was discordant between readers, were classified as negative based on (semi-) quantitative measurements.

Conclusion: The use of parametric BP_{ND} images for visual assessment of 18F-Flutemetamol in a population with low amyloid burden improves inter-reader agreement. Implementing semi-quantification in addition to visual assessment of SUVr images can reduce false-positive classification in this population.

Key words: 1) 18F-Flutemetamol PET, **2)** amyloid pathology, **3)** visual assessment, **4)** preclinical Alzheimer's disease.

Introduction

Alzheimer's disease (AD) is the most common cause of dementia, accounting for 60-80% of cases above 65 years of age.(*1*) Its pathological hallmark is the accumulation of the amyloid-β (Aβ) peptide, thought to start years before cognitive impairment.(*2*) In fact, abnormal Aβ levels are seen in 20-40% of cognitively normal subjects between the ages of 60 and 90 years.(*3*) These subjects are considered to be in the preclinical stage of AD,(*4,5*) which provides a unique opportunity for secondary prevention studies and is gaining increasing research focus.(*6*) To this end, reliable identification of amyloid pathology *in vivo* using Positron Emission Tomography (PET) is of the utmost importance in this population.

The identification of amyloid burden by means of visual interpretation of summed late images or of semi-quantitative standardized uptake value ratio (SUVr) images is currently suggested to be sufficient. Previous studies have shown a high inter-reader agreement for the visual assessment of SUVr images and a high imaging-pathology correlation in clinical populations and end-of-life subjects.(*7-9*) It has been shown, however, that SUVr overestimates amyloid burden compared with quantitative non-displaceable binding potential values (BP_{ND}).(*10*) As such, quantitative BP_{ND} images may be more reliable also for visual interpretation. In a memory clinic population, Zwan and colleagues showed that visual assessment of parametric BP_{ND} 11C-PiB images resulted in a higher inter-reader agreement than the frequently used SUV and SUVr images.(*11*) To date, it remains to be determined whether these findings translate to the increasingly available 18F-labelled Aβ targeting tracers, such as 18F-Flutemetamol, and, more importantly, to the challenging population of cognitively normal elderly participants who have generally a minimal amyloid load.

The purpose of this study was to compare two parametric imaging methods (SUVr vs. BP_{ND}) to determine the optimal approach for assessment of early amyloid pathology. To this end, we investigated the agreement in visual assessment of SUVr and BP_{ND} images between three readers and its relationship to (semi-)quantitative measures.

Materials and Methods

Project

The data used in this study originate from the Innovative Medicine Initiative European Information Framework for AD (EMIF-AD) project (http://www.emif.eu/). The overall aim of the EMIF-AD project is to discover and validate diagnostic markers, prognostic markers and risk factors for AD in non-demented subjects.

Subjects

A total of 199 subjects from the PreclinAD cohort were included at the VU University Medical Center. Inclusion criteria were age ≥ 60 years and normal cognition according to a delayed recall score of > -1.5 standard deviation (SD) of demographically adjusted normative data on the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) 10 word list,(*12*) a Telephone Interview for Cognitive Status modified (TICS-m) score of 23 or higher,(*13*) a 15-item Geriatric Depression Scale (GDS) score of <11,(*14*) and a Clinical Dementia Rating (CDR) scale of 0.(*15*) Exclusion criteria were any physical, neurological or psychiatric condition that interfere with normal cognition. PET-acquisition failed in 3 subjects and 6 BP_{ND} images had missing visual assessment, resulting in a visual assessment for both SUVr and BP_{ND} images of 190 subjects in the present study. PET quantification failed in 5 subjects, thus a total of 185 subjects were used for the quantitative analysis. Written informed consent was obtained from all subjects and the study was approved by the Medical Ethics Review Committee of the VU University Medical Center.

Positron Emission Tomography

PET scans were obtained using a Philips Ingenuity TF PET-MRI camera (Philips Healthcare, Cleveland, USA). Thirty minutes scans were acquired immediately following a manual injection of 18F-Flutemetamol (191 ± 20 MBq).(*16*) After an interval of 60 minutes, in which the patient was taken from the scanner bed, a second scan of 20 minutes was acquired, starting 90 minutes after injection.(*17*) Immediately prior to each part of the PET scan, a T1-weighted gradient echo pulse MRI scan was acquired for attenuation correction (AC) of the PET data. The first emission scan was reconstructed into

18 frames of increasing length (6x5, 3x10, 4x60, 2x150, 2x300, 1x600 s.) using the standard LOR-RAMLA reconstruction algorithm for the brain. The second scan was reconstructed with the same algorithm into 4 frames of 300 seconds each. Fist, Vinci Software 2.56 (Max Planck Institute for Neurological Research, Cologne, Germany) was used to combine the two PET scans into a single multi-frame image. Next, each individual's T1 was co-registered to the dynamic PET using the generic multimodality setting of Vinci with a linear rigid-body schema and normalized mutual information as the similarity measure. Parametric BP_{ND} images were generated from the entire image set using the receptor parametric mapping (RPM) implementation in PPET.(*18-20*) SUVr images were generated based on the 90- to 110-minutes scan interval. Next, T1-based VOIs using the Hammers atlas implemented in PVElab software were projected onto the PET images to extract regional values.(*21*) Cerebellar grey matter was used as reference tissue for both analyses.(*22*) Finally, global values were computed based on the average of frontal (volume-weighted average of superior, middle, and inferior frontal gyrus), parietal (volume weighted average of posterior cingulate, superior parietal gyrus, postcentral gyrus, and inferolateral remainder of parietal lobe), and temporal (volume-weighted average of parahippocampal gyrus, hippocampus, medial temporal lobe, superior, middle, and inferior temporal gyrus) regions.(*23,24*)

Magnetic Resonance Imaging

Whole-brain scans were obtained using a 3T Philips Achieva scanner (Philips Healthcare, Cleveland, USA) of the PET-MRI system described above equipped with an 8-channel head coil. Isotropic structural 3D T1-weighted images were acquired using a sagittal turbo field echo (TFE) sequence with the following settings: 1.00 x 1.00 x 1.00 mm^3 voxels, repetition time (TR) = 7.9 ms, echo time (TE) = 4.5 ms, flip angle (FA) = 8º. A 3D sagittal fat-saturated fluid-attenuated inversion recovery (FLAIR) sequence was acquired using the following settings: 1.12 x 1.12 x 1.12 mm^3 voxels, TR = 4800 ms, TE=279 ms, inversion time = 1650 ms. The structural 3D-T1 and 3D FLAIR images were used for assessment of global cortical atrophy (GCA),(*25*) average medial temporal atrophy (MTA),(*26*) and Fazekas score for white matter hyperintensities (WMH).(*27,28*)

Visual Assessment of PET Images

Three trained readers, blinded to clinical information, first assessed all SUVr images and subsequently all BP_{ND} images, in a randomized order. Images deemed dubious by the reader were reassessed at a separate occasion. Images were scaled to 90% of the pons signal using rainbow colour scaling and transverse, sagittal, and coronal views were displayed using the software package Vinci 2.56. Images were rated as either *positive* (binding in one or more cortical brain region or striatum unilaterally) or *negative* (predominantly white matter uptake) according to criteria defined by the manufacturer (GE Healthcare). PET images were assessed together with a T1-weighted MR scan to limit the influence of atrophy on the visual assessment.

The level of experience in visual assessment of 18F-Flutemetamol images differed among readers; a nuclear medicine physician with considerable experience, a nuclear medicine physician trainee with basic experience, and a radiologist in training with 6 months of experience in nuclear medicine. All readers completed the 18F-Flutemetamol reader training provided by GE Healthcare.

Statistical Analyses

Baseline demographics were assessed using simple descriptive statistical analyses. Kappa statistics were used to asses inter-reader agreement among the three readers, intra-reader agreement between the two methods and agreement between visual read and (semi-)quantitative classifications. Agreement was considered poor if κ was less than 0.20, satisfactory if κ was 0.21–0.40, moderate if κ was 0.41– 0.60, good if κ was 0.61–0.80, and excellent if κ was more than 0.80. Differences in MRI measurements between PET- and PET+ cases were assessed using a Mann-Whitney *U* analysis. The correlation between (semi-)quantitative SUVr and BP_{ND} measurements was assessed using Spearman's rho. Cut-off values were calculated for both SUVr and BP_{ND} using a ROC analysis and the Youden Index. Possible overestimation of amyloid burden using semi-quantitative SUVr was investigated by calculating the difference between SUVr -1 and BP_{ND} values. Differences in global overestimation between PET- and PET+ cases were assessed using a Mann-Whitney *U* analysis. Regional differences in binding and overestimation were assessed using a Wilcoxon paired test. Amyloid status resulting from quantitative

assessment was considered as the “true” amyloid status for all analysis, in the absence of post-mortem confirmation.

Results

Baseline demographics are provided in Table 1.

Visual Reads

Inter-reader agreement in visual assessment was moderate for SUVr images (κ = 0.57) and good for BP_{ND} images (κ = 0.77). There was discordance between readers for 35 cases (18%) using SUVr and for 15 cases (8%) using BP_{ND}, with 9 overlapping cases. **Fig. 1** shows examples of agreement and disagreement in visual interpretation of 18F-Flutemetamol images. On average, the rating was positive in 35 (18%) of the SUVr images and in 26 (14%) of the BP_{ND} images. The reader with the least experience classified 55 (29%) SUVr images as positive compared to 21 (11%) and 29 (15%) by the intermediate and most experienced reader, respectively.

Intra-reader agreement (i.e. within reader, between SUVr and BP_{ND}) differed among readers, with moderate agreement (κ = 0.52) between methods seen in the reader with least experience, excellent agreement (κ = 0.97) seen in the reader with moderate experience, and good agreement in the reader with most experience (κ = 0.76).

When applying ‘majority rules’ (i.e. 2 out of 3 readers agreed on a scan being either positive or negative), positivity was assigned to 27 (14%) cases based on SUVr and to 25 (13%) cases based on the BP_{ND}, with 22 overlapping cases. Thus, 8 cases showed inter-method discordance; i.e. 5 cases were rated positive on SUVr, but negative on BP_{ND} and 3 cases were rated positive on BP_{ND} and negative on SUVr. The remaining 160 cases were classified as negative on both images (**Fig. 2A**).

Visual Reads Related to Quantitative Measures

For the total cohort, mean (± SD) global SUVr and BP_{ND} values were 1.33 (± 0.21) and 0.16 (± 0.12), respectively. There was a good agreement between both measures (ICC = 0.89, $p < 0.01$). Inter-reader concordant positive cases had significantly higher SUVr and BP_{ND} values than concordant negative

cases ($p < 0.01$) (**Table 1**). Based on the visual read concordant cohort alone ($N = 149$), the cut-off value for positivity using SUVr was calculated to be 1.52 (AUC = 0.98, sensitivity = 95%, specificity = 98%) and for BP_{ND} 0.26 (AUC = 1.00, sensitivity = 100%, specificity = 98%) using a ROC analysis (Supplemental Fig. 1). After applying both cut-offs to the dataset, the agreement between the SUVr majority visual read and semi-quantitative negative/positive classification was good ($\kappa = 0.78$), with 16 cases (9%) discordant between the two classification methods. The agreement analysis was also done with a literature based cut-off value (1.56)(*8,29*) resulting in a kappa increase of 0.01. For BP_{ND}, the agreement between majority visual read and quantitative negative/positive classification was excellent ($\kappa = 0.93$), with 3 cases (2%) discordant between the two classification methods. The majority of the 35 cases (91%) where SUVr image assessment was discordant between readers, were classified as negative using either cut-off value (**Fig. 2B**). In addition, in the 8 cases with discordant inter-method visual read, there was full agreement between visual and quantitative measurements when using BP_{ND}, which was not the case with SUVr (**Fig. 2A**).

SUVr ≠ BP_{ND} Quantification

We investigated the relationship between the two quantitative measures with regard to majority visual read to assess any violations of the equilibrium assumptions (i.e. SUVr – 1 = BP_{ND}) in this population. For all cases except one, global SUVr – 1 values overestimated the corresponding global BP_{ND} values. Participants with a positive read ($M = 0.37 \pm 0.11$) had a significantly higher overestimation compared to participants with a negative read ($M = 0.14 \pm 0.07$; $p < 0.01$). This relationship was also observed on a regional level, with the frontal lobe displaying the highest mean binding and the largest mean SUVr overestimation, compared to the parietal ($p < 0.01$) and temporal ($p < 0.01$) lobes. In turn, the parietal lobe did not show a significantly higher mean binding ($p = 0.1$), but did show a significantly larger overestimation ($p < 0.01$) compared to the temporal lobe (supplemental Fig. 2 and table 1). The SUVr overestimation seems to have a limited influence on the visual read of the high binding group (i.e. $BP_{ND} > 0.26$), considering no cases were visually assessed as positive on SUVr and negative on BP_{ND} and only 2 SUVr images (7%) had a discordanct read. For the low binding group (i.e. $BP_{ND} \leq 0.26$), the SUVr overestimation might have influenced the visual read, considering that 26 cases (16%) had a SUVr discordant visual read. However, no obvious pattern was discernible (**Fig. 3**).

Discussion

In a cognitively normal elderly population with low amyloid burden we show a considerable improvement in inter-reader agreement of 18F-Flutemetamol visual assessment when using BP_{ND} rather than standard SUVr images. Misclassifications can be reduced using semi-quantitative SUVr measures and avoided using fully quantitative BP_{ND} measures.

Our results are in line with 11C-PiB findings of Zwan et al., where a comparable improvement in inter-reader agreement using BP_{ND} images was observed.(*11*) This suggests that the underlying reason for discrepant inter-reader agreements tracer-independent and likely related to the distinctive metrics being used (SUVr and BP_{ND}). SUVr is commonly used as a proxy for BP_{ND}, under the assumption that a secular equilibrium is reached during scanning. However, these equilibrium conditions are rarely met in practice. As such, while parametric BP_{ND} images reflect density of available receptors (amyloid plaques), SUVr images are affected by a non-displaceable (free and non-specific) signal and may be affected by changes in regional flow and wash-out effects.(*28,30*) As a result, SUVr can overestimate specific binding,(*10*) and influence visual assessments (**Fig. 3**). Furthermore, our and existing data show that this overestimation is not constant, but instead increases with higher tracer binding.(*10,28*)

The inter-reader agreement for the SUVr images and the concordance between semi-quantitative and corresponding visual read classifications in our study, is lower than previously reported.(*7-9*) However, previous results were based on a clinical population and/or end-of-life subjects with higher incidence of moderate to severe amyloid burden, which highlights the challenge of assessing amyloid pathology in a population with low amyloid burden. The challenge could be due to the non-specific white matter uptake seen with 18F-Flutemetamol, which together with the overestimation resulting from static scanning may translate into a tendency to visually assign regions as positive.(*31*) In our study, the frontal regions were most often perceived as difficult to assess, leading to most doubt for final classification. Although the 18F-Flutemetamol reader training focuses on disentangling the white matter pattern from the cortical signal, assessment in this population seems additionally challenging, especially for less experienced readers. Indeed, the positive assigning tendency was the strongest for the reader with the least experience, who also showed the lowest *intra*-reader agreement between methods. This stresses the need for experienced readers to make early assessments and/or updating the reading guidelines, with

focus on a cognitively normal elderly population. Of note, while the reference region used for visual assessment (i.e. pons) is different from that used for quantitative assessment (i.e. grey matter cerebellum), a separate agreement analysis using pons for quantification did not affect the agreement between classification methods (data not shown).

Our results may have consequences for drug-intervention studies focused on early populations, since using the visual assessment of SUVr images either as an inclusion criteria could result in false-positive inclusion due to the observed overestimation of cortical amyloid burden.(*32,33*) Also, studies indicate that cerebral blood flow can change with age and disease progression.(*34,35*) Therefore, using BP_{ND} images in clinical trials could avoid false-positive classification in visual assessment(*28*) and ensure that measured changes are due to the treatment instead of a measurement error or blood flow confounders.

An important factor in considering dynamic PET acquisition is participant burden. In this cohort, 95% of participants indicated they had no objections to undergoing a second dynamic PET scan. The coffee-break protocol used in this study may have facilitated this response and suggests the feasibility of longitudinal dynamic acquisition in cognitively normal elderly.

In a clinical setting, however, amyloid burden will more likely be moderate to severe and dynamic acquisition more challenging. In addition, the clinical utility of SUV or SUVr visual reads for the diagnosis of AD-type dementia in a clinical setting has been extensively shown.(*36*) Thus, in this context, visual assessment of SUVr images may indeed be sufficient. Nevertheless, the present results illustrate that semi-quantification using SUVr can help reduce false-positive classification, especially in a challenging population. Thus, the clinical preference for visual assessment could be revised in light of more available automatic (semi-)quantification methods, such as the one already provided for 18FFlutemetamol PET scans.(*8*)

In this study, the standard manufacturer guidelines were used for reading both SUVr and BP_{ND} images. Nonetheless, an interesting finding was the improvement in inter-reader agreement when using BP_{ND} images despite the lack of official guidelines and the limited experience of readers in assessing such images. However, it might still be of interest to formally assess whether the current guidelines are optimal for assessing BP_{ND} images. In addition, optimizing visual assessment of SUVr images by updating

the current guidelines and providing training specifically focused on early accumulation, may also result in improved classification certainty, comparable to the observed using dynamically derived measures. Studies have suggested that specifically medial frontal, anterior/posterior/isthmus cingulate cortex, and the precuneus are early accumulating regions.(*37,38*) These regions can be visually assessed using the sagittal view of the PET image. Thus, the importance of this plane may be of most interest for updating guidelines.

A limitation of this study is the lack of a gold standard, as no post-mortem data were available, hampering the understanding of the findings in relation to underlying neuropathology. Furthermore, although the frequency of amyloid positivity in this cohort is comparable with previous reports,(*39*) the low incidence may have induced reader bias with regards to searching for amyloid positivity. Lastly, both quantification and visual assessment of the PET images in this study were accompanied by a structural MRI, which might not always be available.

Conclusion

The use of parametric BP_{ND} images for visual assessment of 18F-Flutemetamol in a population with low amyloid burden improves inter-reader agreement. Implementing semi-quantification in addition to visual assessment of SUVr images can reduce false-positive classification in this population.

Disclosures

L.E. Collij; Elles Konijnenberg; Juhan Reimand; Mara ten Kate; Anouk den Braber; Isadora Lopes Alves; Marissa Zwan; Maqsood Yaqub; Daniëlle M.E. van Assema; Alle Meije Wink; Adriaan A. Lammertsma; & Bart N.M. van Berckel all report no existing potential conflicts of interest relevant to this article .

Dr. Philip Scheltens received grants from GE Healthcare, Piramal, and Merck, paid to his institution; he has received speaker's fees paid to the institution Alzheimer Center, VU University Medical Center, Lilly, GE Healthcare, and Roche

Dr. Pieter Jelle Visser received research support from Biogen, grants from EU/EFPIA Innovative Medicines Initiative Joint Undertaking, EU Joint Programme–Neurodegenerative Disease Research (JPND), ZonMw, and Bristol-Myers Squibb; having served as member of the advisory board of Roche Diagnostics; and having received nonfinancial support from GE Healthcare.

Dr. Frederik Barkhof received payment and honoraria from Bayer-Schering Pharma, Sanofi-Aventis, Genzyme, Biogen-Idec, TEVA, Merck-Serono, Novartis, Roche, Jansen Research, IXICO Ltd, GeNeuro, and Apitope Ltd for consulting; payment from the Serono Symposia Foundation, IXICOLtd, and MedScape for educational presentations; research support via grants from EU/EFPIA Innovative Medicines Initiative Joint Undertaking (AMYPAD consortium), EuroPOND (H2020), UK MS Society, Dutch MS Society, PICTURE (IMDI-NWO), NIHR UCLH Biomedical Research Centre (BRC), ECTRIMS-MAGNIMS.

No other potential conflicts of interest relevant to this article exist.

Acknowledgements

Data collection and sharing for this project was funded by the EU/EFPIA Innovative Medicines Initiative Joint Undertaking EMIF grant agreement (grant No115372). This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 115952. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA. Supported by the NIHR UCLH biomedical research centre. This work also received in kind sponsoring of the PET-tracer from GE Healthcare.

REFERENCES

1. Alzheimer's A. 2016 Alzheimer's disease facts and figures. *Alzheimers Dement.* 2016;12:459-509.

2. Jansen WJ, Ossenkoppele R, Knol DL, et al. Prevalence of cerebral amyloid pathology in persons without dementia: a meta-analysis. *JAMA.* 2015;313:1924-1938.

3. Jansen WJ, Ossenkoppele R, Knol DL, et al. Prevalence of cerebral amyloid pathology in persons without dementia: a meta-analysis. *JAMA : the journal of the American Medical Association.* 2015;313:1924-1938.

4. Dubois B, Feldman HH, Jacova C, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet neurology.* 2007;6:734-746.

5. Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association.* 2011;7:280-292.

6. Sperling RA, Rentz DM, Johnson KA, et al. The A4 study: stopping AD before symptoms begin? *Sci Transl Med.* 2014;6:228fs213.

7. Buckley CJ, Sherwin PF, Smith AP, Wolber J, Weick SM, Brooks DJ. Validation of an electronic image reader training programme for interpretation of [18F]flutemetamol beta-amyloid PET brain images. *Nucl Med Commun.* 2017;38:234-241.

8. 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;55:1623-1628.

9. Salloway S, Gamez JE, Singh U, et al. Performance of [(18)F]flutemetamol amyloid imaging against the neuritic plaque component of CERAD and the current (2012) NIA-AA recommendations for the neuropathologic diagnosis of Alzheimer's disease. *Alzheimers Dement (Amst).* 2017;9:25-34.

10. van Berckel BN, Ossenkoppele R, Tolboom N, et al. Longitudinal amyloid imaging using 11C-PiB: methodologic considerations. *J Nucl Med.* 2013;54:1570-1576.

11. Zwan MD, Ossenkoppele R, Tolboom N, et al. Comparison of simplified parametric methods for visual interpretation of 11C-Pittsburgh compound-B PET images. *J Nucl Med.* 2014;55:1305-1307.

12. Morris JC, Heyman A, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology.* Vol 39; 1989:1159-1165.

13. de Jager CA, Budge MM, Clarke R. Utility of TICS-M for the assessment of cognitive function in older adults. *International journal of geriatric psychiatry.* Vol 18; 2003:318-324.

14. Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *Journal of psychiatric research.* Vol 17; 1982:37-49.

15. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology.* Vol 43; 1993:2412-2414.

16. Nelissen N, Van Laere K, Thurfjell L, et al. Phase 1 study of the Pittsburgh compound B derivative 18F-flutemetamol in healthy volunteers and patients with probable Alzheimer disease. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine.* Vol 50: Society of Nuclear Medicine; 2009:1251-1259.

17. Heeman F, Yaqub M, Heurling K, Lopes Alves I, Gispert JD, Foley C, Lammertsma AA. P20: Optimized coffee-break protocol for quantitative

[18F]flutemetamol studies. *Human Amyloid Imaging Conference.* 2018:83.

18. Gunn RN, Lammertsma AA, Hume SP, Cunningham VJ. Parametric imaging of ligand-receptor binding in PET using a simplified reference region model. *NeuroImage.* Vol 6. 1998/02/07 ed; 1997:279-287.

19. Wu Y, Carson RE. Noise reduction in the simplified reference tissue model for neuroreceptor functional imaging. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism.* Vol 22. 2002/12/07 ed; 2002:1440-1452.

20. Boellaard R, Yaqub M, Lubberink M, Lammertsma A. PPET: a software tool for kinetic and parametric analyses of dynamic PET studies. *NeuroImage.* 2006;31:T62.

21. Svarer C, Madsen K, Hasselbalch SG, et al. MR-based automatic delineation of volumes of interest in human brain PET images using probability maps. *Neuroimage.* 2005;24:969-979.

22. Hammers A, Allom R, Koepp MJ, et al. Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Hum Brain Mapp.* 2003;19:224-247.

23. Tolboom N, Yaqub M, van der Flier WM, et al. Detection of Alzheimer pathology in vivo using both 11C-PIB and 18F-FDDNP PET. *J Nucl Med.* 2009;50:191-197.

24. Guo T, Dukart J, Brendel M, Rominger A, Grimmer T, Yakushev I. Rate of β-amyloid accumulation varies with baseline amyloid burden: Implications for anti-amyloid drug trials. *Alzheimer's & Dementia.* 2018.

25. Koedam EL, Lehmann M, van der Flier WM, et al. Visual assessment of posterior atrophy development of a MRI rating scale. *Eur Radiol.* 2011;21:2618-2625.

26. Scheltens P, Leys D, Barkhof F, et al. Atrophy of Medial Temporal Lobes on Mri in Probable Alzheimers-Disease and Normal Aging - Diagnostic-Value and Neuropsychological Correlates. *Journal of Neurology Neurosurgery and Psychiatry.* 1992;55:967-972.

27. Fazekas F, Chawluk JB, Alavi A, Hurtig HI, Zimmerman RA. Mr Signal Abnormalities at 1.5-T in Alzheimers Dementia and Normal Aging. *American Journal of Neuroradiology.* 1987;8:421-426.

28. Lammertsma AA. Forward to the Past: The Case for Quantitative PET Imaging. *J Nucl Med.* 2017;58:1019-1024.

29. Vandenberghe R, Van Laere K, Ivanoiu A, et al. 18F-flutemetamol amyloid imaging in Alzheimer disease and mild cognitive impairment: a phase 2 trial. *Ann Neurol.* 2010;68:319-329.

30. Carson RE, Channing MA, Blasberg RG, et al. Comparison of bolus and infusion methods for receptor quantitation: application to [18F]cyclofoxy and positron emission tomography. *J Cereb Blood Flow Metab.* 1993;13:24-42.

31. Mountz JM, Laymon CM, Cohen AD, et al. Comparison of qualitative and quantitative imaging characteristics of [11C]PiB and [18F]flutemetamol in normal control and Alzheimer's subjects. *Neuroimage Clin.* 2015;9:592-598.

32. Salloway S, Sperling R, Fox NC, et al. Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. *N Engl J Med.* 2014;370:322-333.

33. Doody RS, Thomas RG, Farlow M, et al. Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease. *N Engl J Med.* 2014;370:311-321.

34. Zhang N, Gordon ML, Goldberg TE. Cerebral blood flow measured by arterial spin labeling MRI at resting state in normal aging and Alzheimer's disease. *Neurosci Biobehav Rev.* 2017;72:168-175.

35. Sierra-Marcos A. Regional Cerebral Blood Flow in Mild Cognitive Impairment and Alzheimer's Disease Measured with Arterial Spin Labeling Magnetic Resonance Imaging. *Int J Alzheimers Dis.* 2017;2017:5479597.

36. Morris E, Chalkidou A, Hammers A, Peacock J, Summers J, Keevil S. Diagnostic accuracy of (18)F amyloid PET tracers for the diagnosis of Alzheimer's disease: a systematic review and meta-analysis. *Eur J Nucl Med Mol Imaging.* 2016;43:374-385.

37. Grothe MJ, Barthel H, Sepulcre J, et al. In vivo staging of regional amyloid deposition. *Neurology.* 2017;89:2031-2038.

38. Palmqvist S, Scholl M, Strandberg O, et al. Earliest accumulation of beta-amyloid occurs within the default-mode network and concurrently affects brain connectivity. *Nat Commun.* 2017;8:1214.

39. Jansen W, Ossenkoppele R, Verhey FR, Visser PJ. The prevalence of amyloid pathology in healthy individuals and individuals with mild cognitive impairment: A meta-analysis of amyloid-PET studies. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association.* 2013;9:P533.

Table 1. Demographics, MRI Measurements and (Semi-)Quantitative PET Values		
Total Cohort N = 190		
Demographics		
Gender	113 women (59.5%)	
Age	70.4 ± 7.56 y	
MMSE	29 ± 1.13	
Years of education	15.15 ± 4.42 y	
Brain measures		
GCA*	0.79 ± 0.72	
Average MTA†	0.65 ± 0.72	
Fazekas‡	1.18 ± 0.82	
	0 = 35; 1=101; 2 = 40; 3 = 14	
Quantitative Cohort *N* = 185		
Quantitative Measures		
SUVr	1.33 ± 0.21, range = 0.79 – 2.13	
BP_{ND}	0.16 ± 0.12, range = 0.20 - 0.66	
Concordant Cohort *N* = 149		
Brain measures	**PET- (*N* = 130)**	**PET+ (*N* = 19)**
GCA*	0.74 ± 0.67	0.89 ± 0.81
Average MTA†	0.57 ± 0.64	0.82 ± 0.75
Fazekas‡	1.18 ± 0.83	1.26 ± 0.87
Quantitative Measures		
SUVr	1.25 ± 0.09	1.83 ± 0.16§
	range = 1.08 – 1.63	range = 1.54 – 2.13
BP_{ND}	0.12 ± 0 .05	0.43 ± 0.12§
	range = 0.02 - 0.27	range = 0.27 - 0.66

Results are displayed as mean ± SD

*Global Cortical Atrophy score (0-3)

†Medial Temporal Atrophy score (0-4)

‡White matter hyperintensity score (0-3)

§ $p < 0.01$

Fig. 1. Examples of SUVr (top) and BP_{ND} (bottom) [^{18}F]flutemetamol images of three different patients.

From left to right are shown axial, coronal, and sagittal views. The three boxes on the right represent the amyloid classification by the three readers (red = negative, green = positive).

Subject 1: Example of a difficult case, represented by discordant visual reads on both SUVr and BP_{ND} image. **Subject 2:** Example of a possible overestimation of amyloid pathology when only assessing the SUVr image. **Subject 3:** Example of a clear positive case.

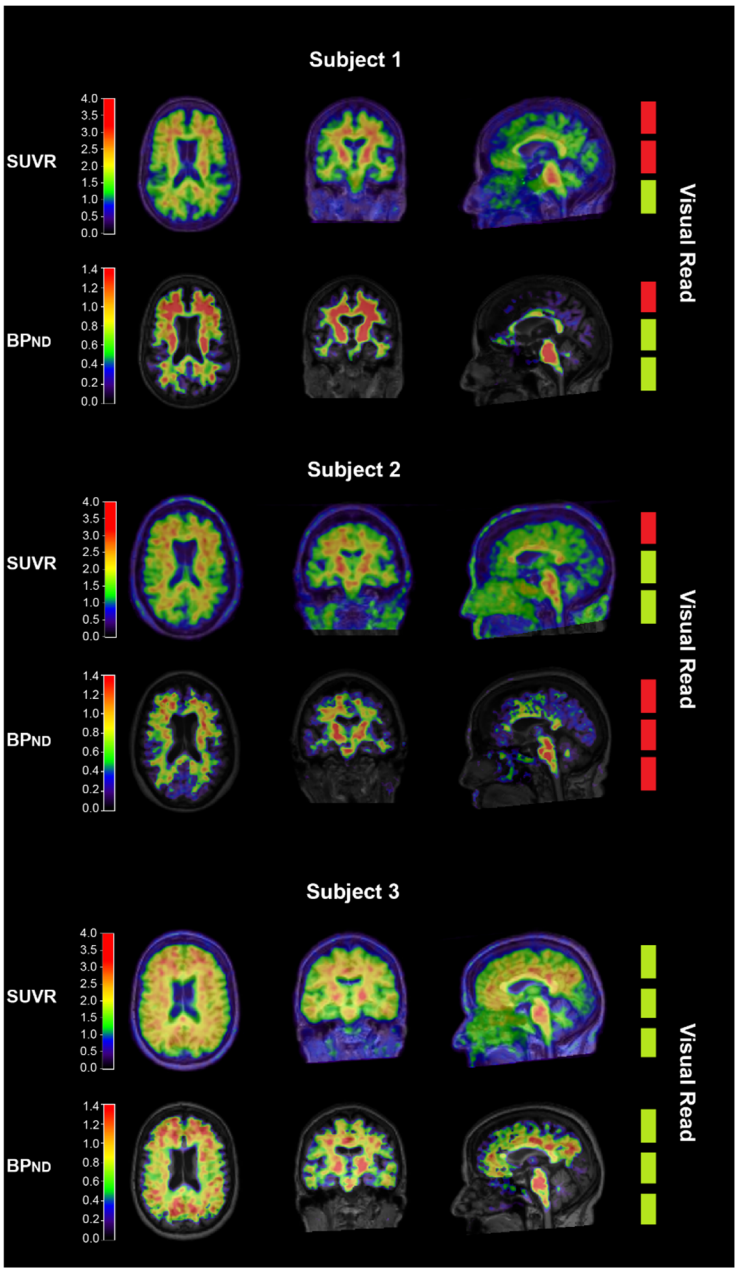


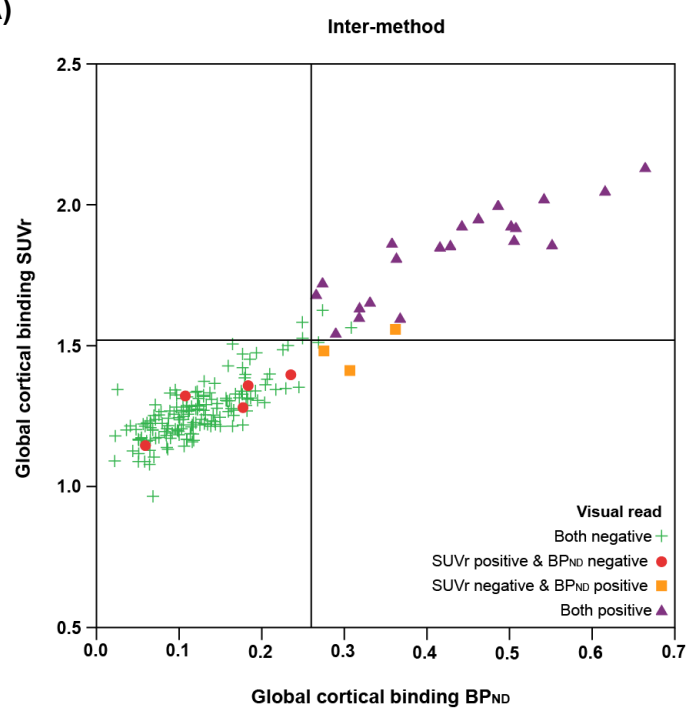
Fig. 2. Scatterplot of quantitative measures compared to visual read.

On X-axis the global cortical binding derived from BP_{ND}. On Y-axis the global cortical binding derived from SUVr. Reference lines denote the cut-off (1.52 for SUVr and .26 for BP_{ND}). Different colours demonstrate discordance/concordance between SUVr and BP_{ND} visual read.

A) Based on majority rules visual read. For all inter-method discordant cases (red circles and orange squares) the BP_{ND} visual read was in accordance with the quantitative value, while SUVr was not.

B) Most SUVr inter-reader discordant cases (red circles) are below the cut-off for both SUVr and BP_{ND}.

A)



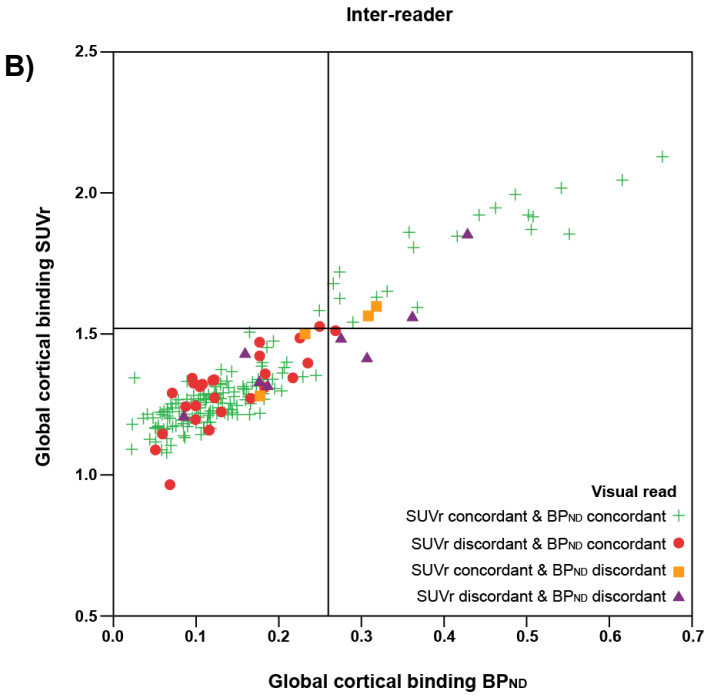
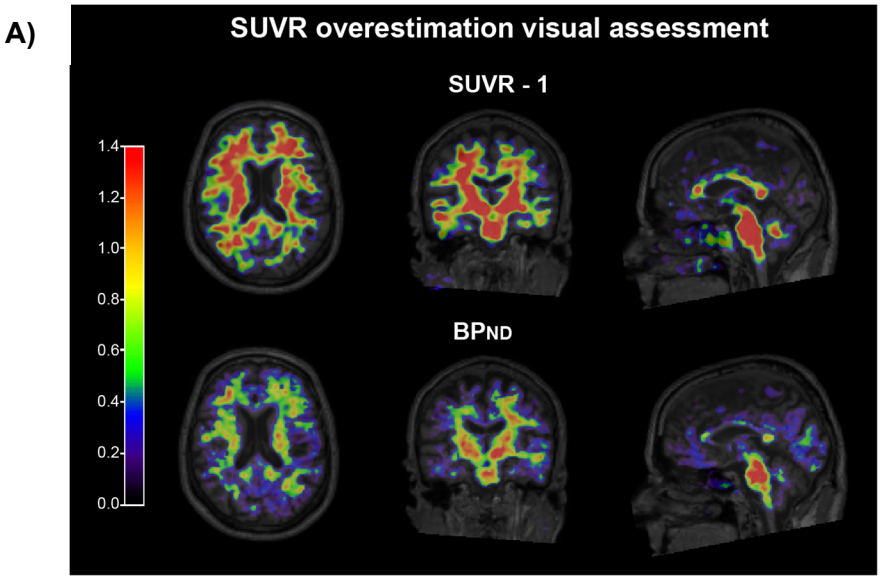
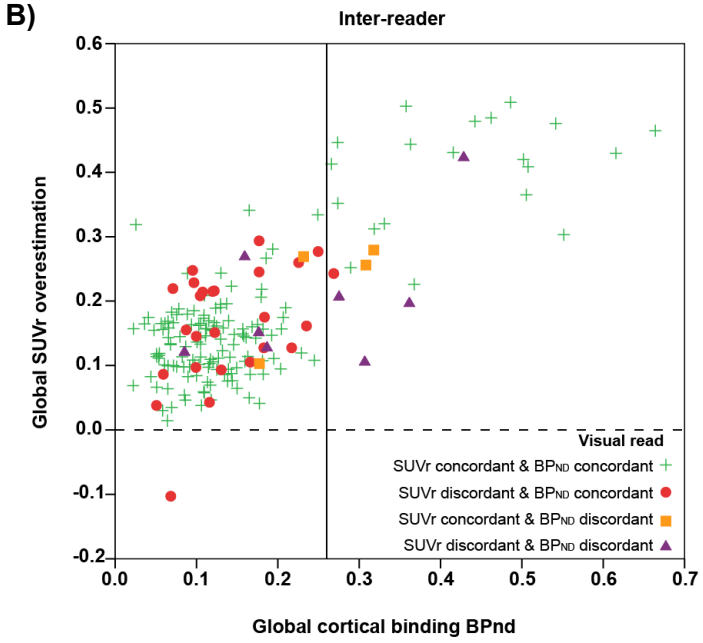

B)
Inter-reader
Global cortical binding SUVr
Global cortical binding BPND
2.5
2.0
1.5
1.0
0.5
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Visual read
SUVr concordant & BPND concordant
SUVr discordant & BPND concordant
SUVr concordant & BPND discordant
SUVr discordant & BPND discordant

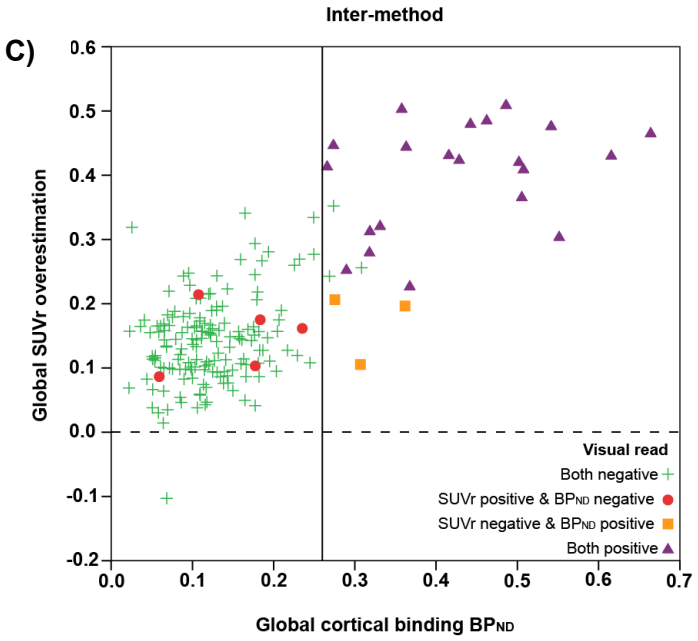
Fig. 3. Illustration of binding overestimation when using semi-quantitative PET acquisition.

(**A**) From left to right are shown axial, coronal, and sagittal views. (**Top**) A SUVr image with a subtraction of 1, showing clearly higher binding values than the BP_{ND} image (**bottom**), while in theory this should be the same image.

(**C/D**) Diagrams showing the difference between SUVr-1 and BP_{ND} for each subject with regard to visual read. The overestimation of SUVr is higher with increasing cortical binding.






C)
Inter-method
Global SUVr overestimation
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Global cortical binding BPND
Visual read
Both negative
SUVr positive & BPND negative
SUVr negative & BPND positive
Both positive

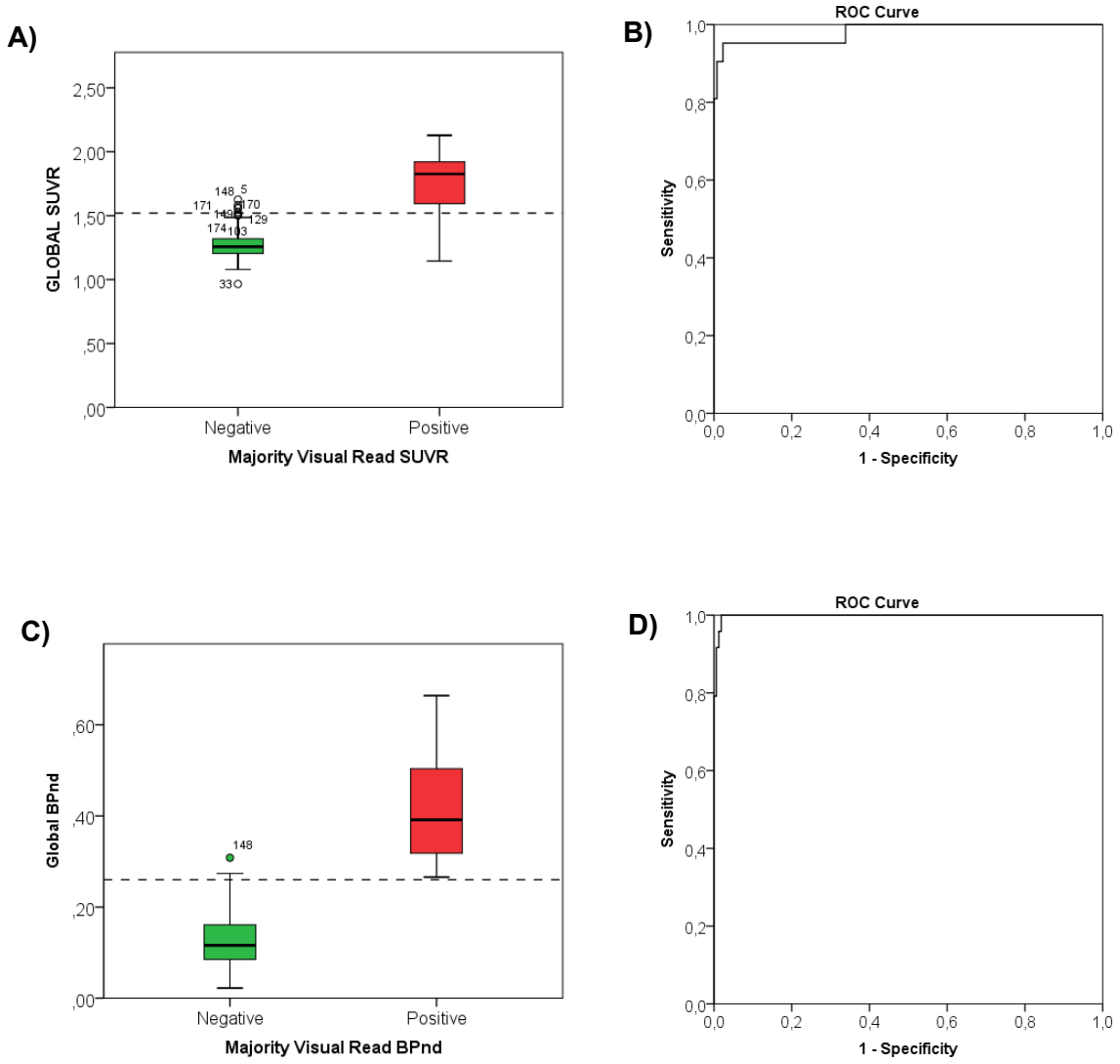
Supplementary Table 1

Supplementary Table 1. Regional Binding and SUVr Overestimation		
Regional ROI	**Binding (BP_{ND})**	**SUVr overestimation**
Frontal	.20 ± .14*	.18 ± .12*
Parietal	.14 ± .12	.17 ± .11†
Temporal	.14 ± .09	.15 ± .10

*Significantly higher than parietal and temporal ROI ($p < .01$).

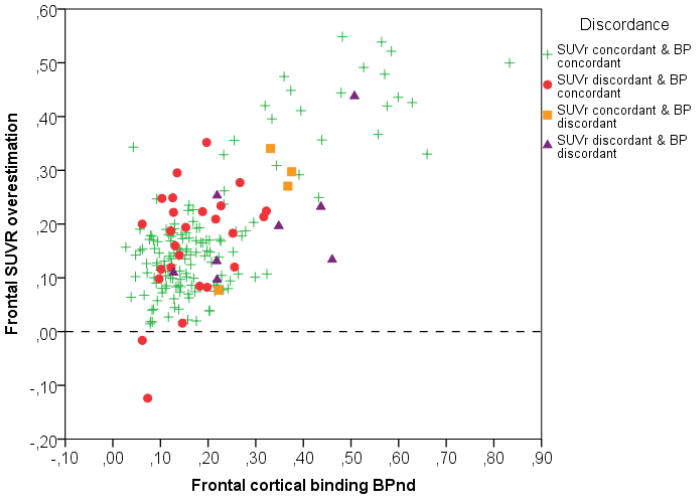
†Significantly higher than temporal ROI ($p < .01$).

Supplementary Figure 1. Negative/positive classification based on (semi-)quantitative measures. A total of 149 cases had a concordant visual read on both the SUVr and BP_{ND} image and were used in a ROC analysis in order to determine the dichotomous negative/positive cut-off value of the (semi-)quantitative measurements.
A) Boxplot displaying the separation of groups based on the majority visual read of SUVr images and the spread of semi-quantitative values within groups. **B**) ROC curve representing the sensitivity/specificity corresponding to the decision threshold (1.52) based on semi-quantitative values. **C**) Boxplot displaying the separation of groups based on the majority visual read of BP_{ND} images and the spread of quantitative values within groups. **D**) ROC curve representing the sensitivity/specificity corresponding to the decision threshold (0.26) based on quantitative values.

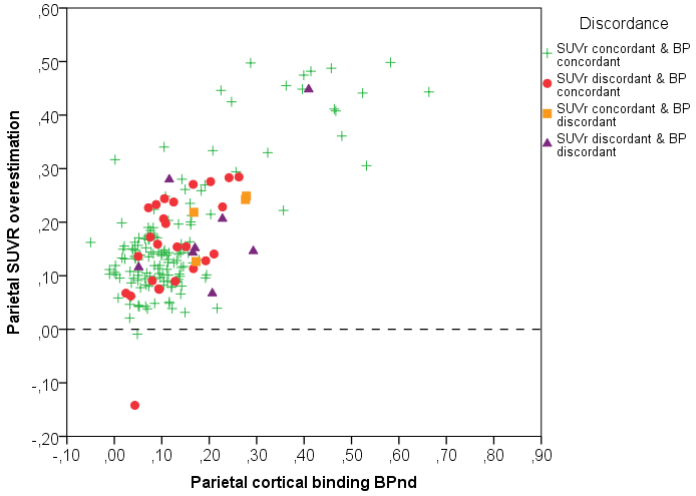


Supplementary Figure 2. Regional SUVr overestimations.
Diagrams showing the difference between **A**) frontal, **B**) parietal, and **C**) temporal SUVr overestimation and BP_{ND} for each subject with regard to visual read. The overestimation of SUVr is higher with increasing cortical binding in all regions.

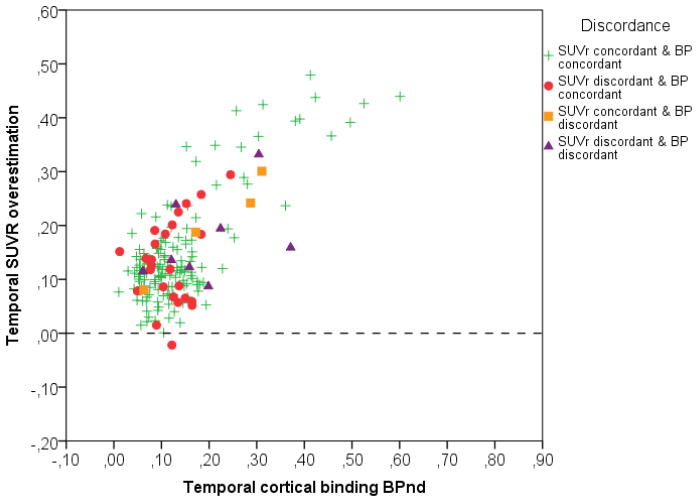
A)



B)



C)


Discordance
SUVr concordant & BP concordant
SUVr discordant & BP concordant
SUVr concordant & BP discordant
SUVr discordant & BP discordant
Temporal SUVR overestimation
,60
,50
,40
,30
,20
,10
,00
-,10
-,20
-,10
,00
,10
,20
,30
,40
,50
,60
,70
,80
,90
Temporal cortical binding BPnd