# Supplementary Methods

## PET acquisition

*Sample 1 – BACS/UCSF*

PIB and FTP PET scans were acquired at Lawrence Berkeley National Laboratory on an ECAT EXACT HR scanner (for 4 PIB scans) or a Biograph Truepoint 6 PET/CT tomograph (Siemens Medical Systems, Erlangen, Germany) for the rest of PIB scans and all FTP scans. We analyzed data that was acquired in list mode from 80 to 100 minutes after intravenous injection of approximately 370 MBq of FTP and reconstructed as 4x5min frames using an ordered subset expectation maximization algorithm with weighted attenuation and smoothed with a 4 mm Gaussian kernel with scatter correction (calculated image resolution 6.5 x 6.5 x 7.25 mm$^3$ using Hoffman). Frames were realigned and averaged; the resulting image was used for the visual read in order to make the visual interpretation entirely qualitative. The lack of reference region normalization was also intended to resemble a clinical scenario, where SUVR might not be available. Details on PIB-PET acquisition and processing are described in details elsewhere (see Villeneuve et al, Brain 2015). Aβ positivity was based on a PIB distribution volume ratio > 1.065, a cutoff adapted from previous studies conducted in our lab.

*Sample 2 - ADNI*

All FTP images used for the visual assessment were processed up to stage 2 of the ADNI PET protocol using the 80-100 min PET data (i.e. co-registered, averaged frames) at each scanner's specific resolution (http://adni.loni.usc.edu/methods/pet-analysis-method/pet-analysis) to mimic real life heterogeneity in image quality. Resulting images were smoothed with a 4 mm Gaussian kernel in order to make them more readable; this step did no bring scans into a common resolution. Similar to sample 1, readers did not look at SUVR images to resemble a clinical scenario. Aβ positivity was based on the Aβ-scan closest to the tau-PET, using either ^{18}F-Florbetapir (n=128) or ^{18}F-Florbetaben (n=9) and corresponding thresholds to determine positivity based on global SUVR (1.11 for ^{18}F-Florbetapir and 1.08 for ^{18}F-Florbetaben).
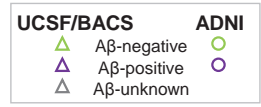
# Supplementary Figures

**Supplementary Figure 1. Reliability analyses for the global visual score.**

A-B. Bubble plots showing inter (A) and intra (B) rater reliability of the 0-14 global visual score for FTP-PET. Bubble area is proportional to the number of cases, and the dotted diagonal line represents the identity line (i.e. similar scores assigned by the two raters (A) or by the same rater twice (B) for a given scan).

C. Inter- and intra-rater reliability for the FTP distribution patterns.

95%CI: 95% confidence interval.

# A. Inter-rater reliability

All cases (137 BACS/UCSF + 137 ADNI)

Spearman's rho = .805
Linear weighted κ = .73, $^{95\%}$CI [.68-.78]



# B. Intra-rater reliability

15 BACS/UCSF + 15 ADNI cases read twice by each rater



Reader 1

Linear weighted
κ = .79, $^{95\%}$CI [.68-.91]

Reader 2

Linear weighted
κ = .84, $^{95\%}$CI [.76-.93]

**Supplementary Figure 2. Associations between FTP cortical SUVR and global visual score from each reader and in each cohort.**
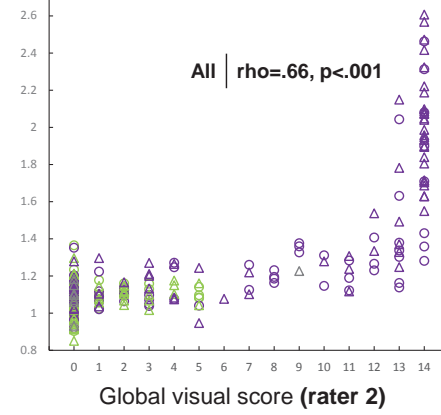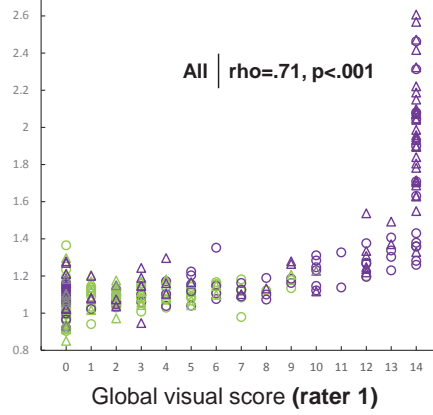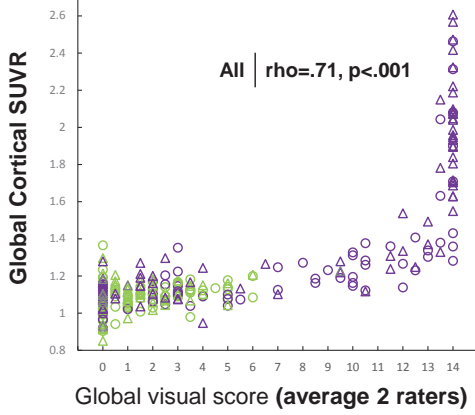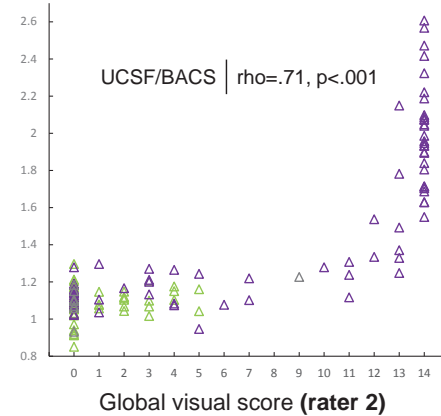
Figure legend (top left):

**UCSF/BACS** — Aβ-negative (green triangle), Aβ-positive (purple triangle), Aβ-unknown (gray triangle)
**ADNI** — Aβ-negative (green circle), Aβ-positive (purple circle), Aβ-unknown (gray circle)

Column headers: Average 2 raters | Rater 1 | Rater 2

Row labels: UCSF/BACS + ADNI | UCSF/BACS | ADNI

Panel annotations:
- All | rho=.71, p<.001
- All | rho=.71, p<.001
- All | rho=.66, p<.001
- UCSF/BACS | rho=.74, p<.001
- UCSF/BACS | rho=.74, p<.001
- UCSF/BACS | rho=.71, p<.001
- ADNI | rho=.69, p<.001
- ADNI | rho=.71, p<.001
- ADNI | rho=.59, p<.001

Y-axis: Global Cortical SUVR
X-axis: Global visual score (average 2 raters) | Global visual score (rater 1) | Global visual score (rater 2)
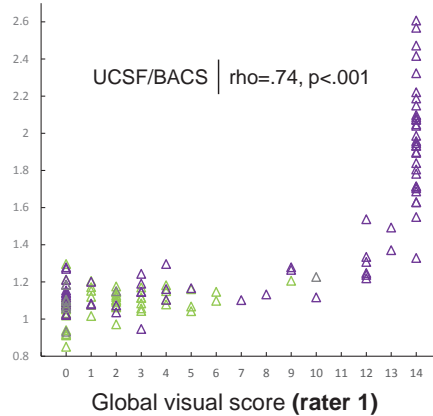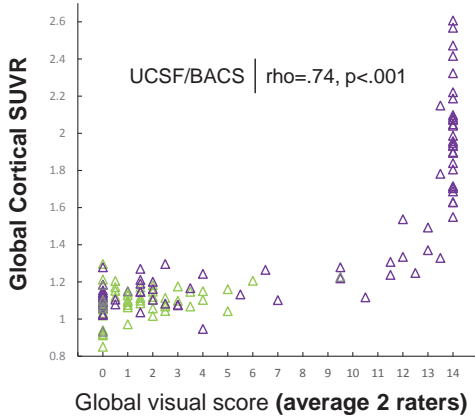
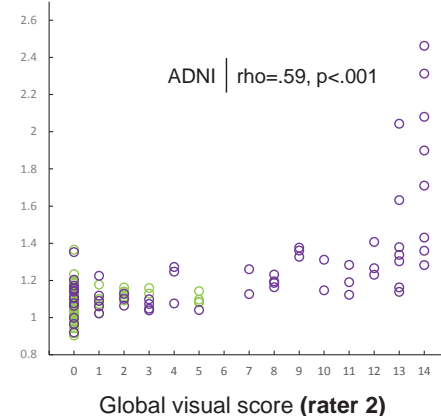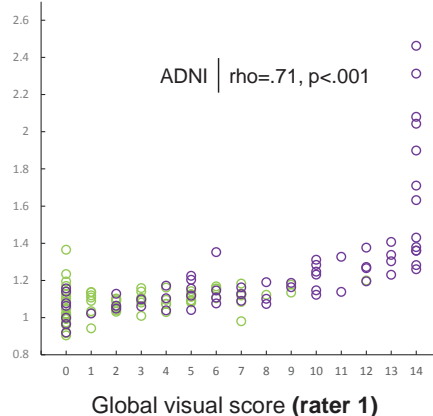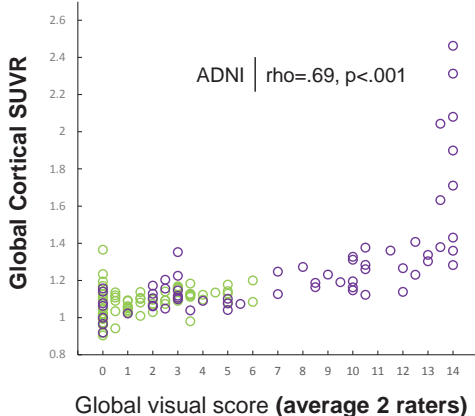**Supplementary Figure 3. Association of clinical diagnosis and Aβ status on the global visual score from each reader or global cortical SUVR.**

AUROC: Area Under the Receiver Operating Curve (comparing Aβ-positive and Aβ-negative cases within a given clinical group. *: p<.05, **:p<.01, ***:p<.001.
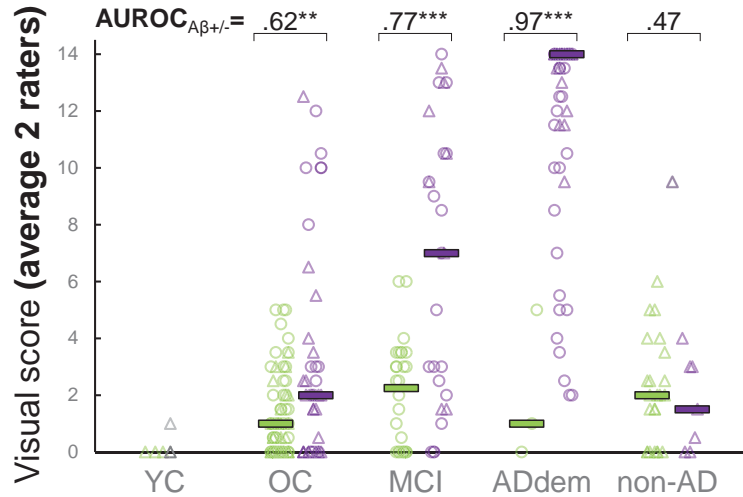
YC: Young Controls, OC: Older Controls, MCI: patients with a clinical diagnosis of Mild Cognitive Impairment, ADdem: patients with a clinical diagnosis of Alzheimer's disease dementia, non-AD: patients with a clinical diagnosis of non-Alzheimer's disease neurodegenerative syndrome. Clinical diagnoses were independent from Aβ status.

**Legend:**

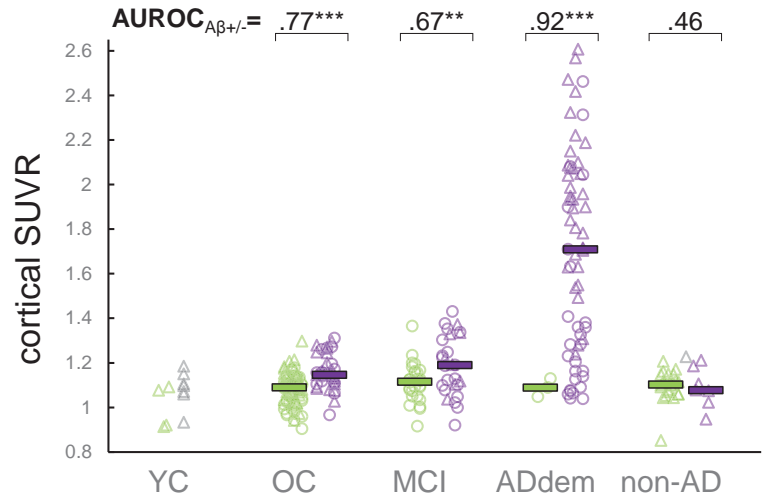UCSF/BACS    ADNI
- △ Aβ-negative ◯ (green)
- △ Aβ-positive ◯ (purple)
- △ Aβ-unknown (grey)
- ▬ Median Aβ-negative (green)
- ▬ Median Aβ-positive (purple)

General Linear Models exclude all Young Controls and the single non-AD patient without available amyloid status

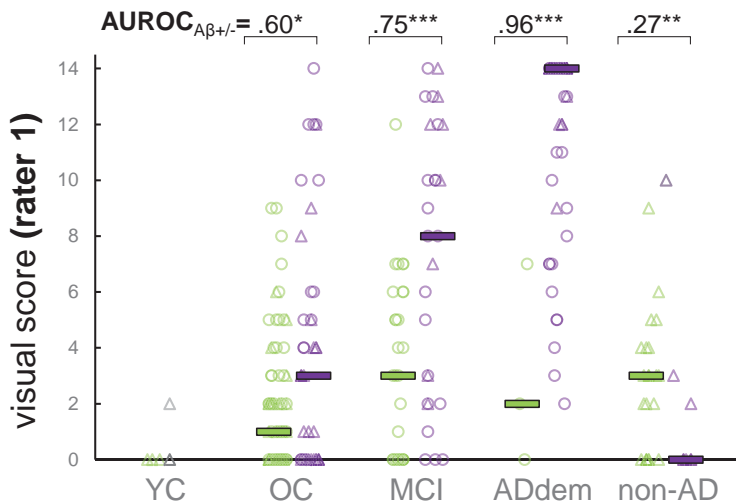**Top-left panel (Visual score, average 2 raters):**

Group   $\eta^2$=.05, p<.001
Aβ   $\eta^2$=.06, p<.001    $R^2$=.68
Group x Aβ   $\eta^2$=.04, p<.001

$AUROC_{A\beta+/-}$=   .62**   .77***   .97***   .47

Groups: YC, OC, MCI, ADdem, non-AD

**Top-right panel (cortical SUVR):**

Group   $\eta^2$=.03, p=.002
Aβ   $\eta^2$=.03, p<.001    $R^2$=.54
Group x Aβ   $\eta^2$=.03, p=.002

$AUROC_{A\beta+/-}$=   .77***   .67**   .92***   .46

Groups: YC, OC, MCI, ADdem, non-AD

**Bottom-left panel (visual score, rater 1):**

Group   $\eta^2$=.06, p<.001
Aβ   $\eta^2$=.04, p<.001    $R^2$=.61
Group x Aβ   $\eta^2$=.04, p<.001

$AUROC_{A\beta+/-}$=   .60*   .75***   .96***   .27**

Groups: YC, OC, MCI, ADdem, non-AD

**Bottom-right panel (visual score, rater 2):**

Group   $\eta^2$=.04, p<.001
Aβ   $\eta^2$=.07, p<.001    $R^2$=.69
Group x Aβ   $\eta^2$=.04, p<.001

$AUROC_{A\beta+/-}$=   .62**   .85***   .97***   .64
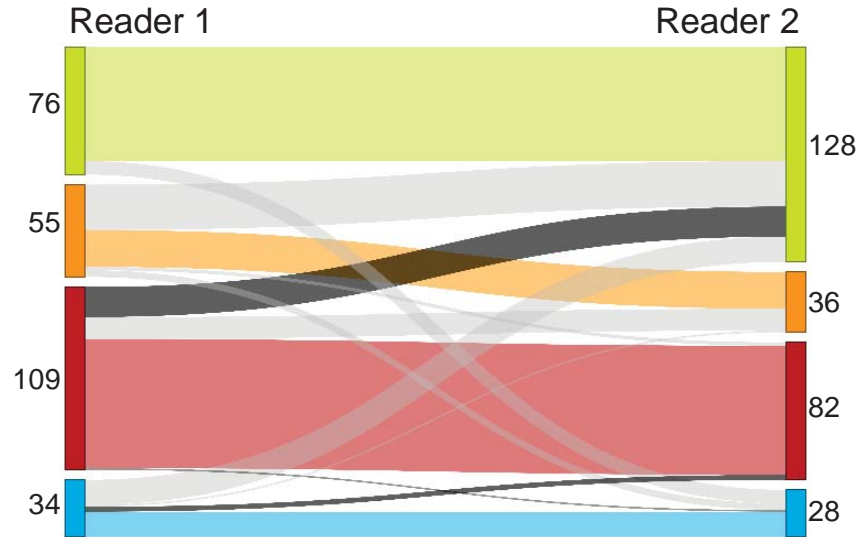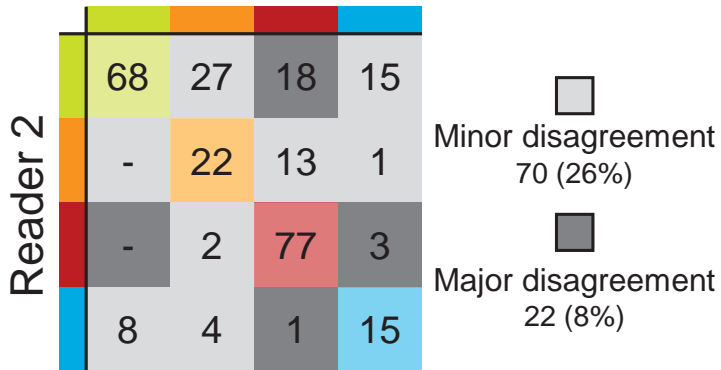
Groups: YC, OC, MCI, ADdem, non-AD
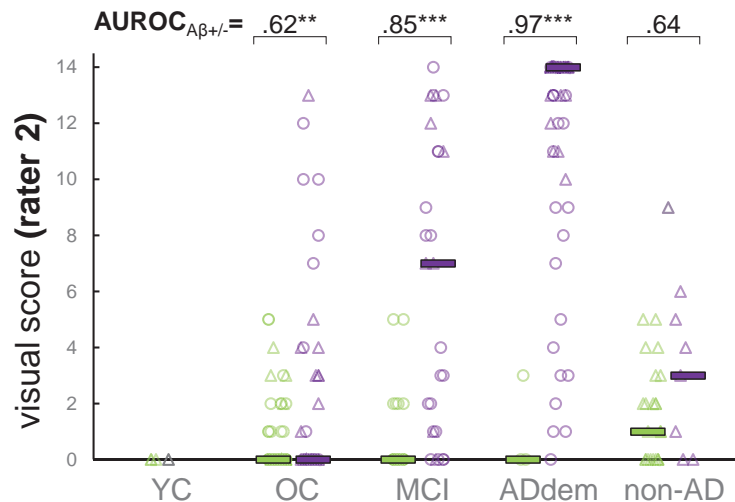
**Supplementary Figure 4. Reliability analyses for the FTP visual pattern.**

**Four visual FTP-PET patterns**

Negative | Mild Temporal only | AD-like | Non-AD-like

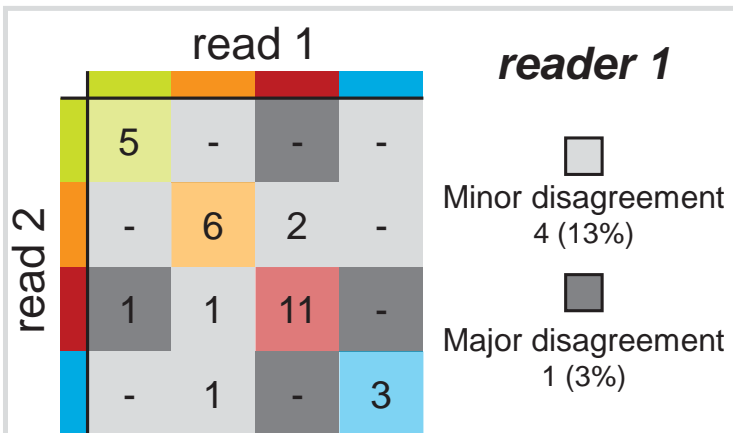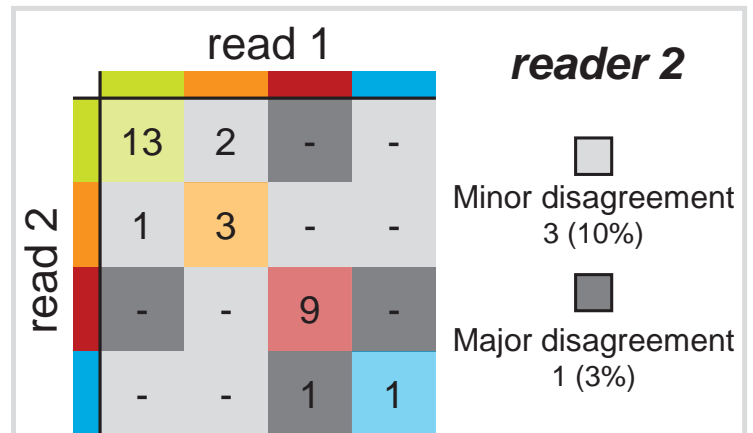## A. Inter-rater agreement — All 274 cases (137 from each cohort)

Reader 1

| | | | |
|---|---|---|---|
| 68 | 27 | 18 | 15 |
| - | 22 | 13 | 1 |
| - | 2 | 77 | 3 |
| 8 | 4 | 1 | 15 |

Reader 2

Minor disagreement 70 (26%)

Major disagreement 22 (8%)

| Contrast | K [95%CI] |
|---|---|
| 4 pattern classification | .53 [.45-.61] |
| AD-like vs others | .71 [.62-.79] |
| Negative vs others | .49 [.38-.59] |
| Temporal vs others | .39 [.23-.55] |
| Non-AD-like vs others | .41 [.23-.61] |

Reader 1 — Reader 2

76 → 128
55 → 36
109 → 82
34 → 28

## B. Intra-rater reliability — 15 cases from each cohort read twice by each reader

**reader 1**

read 1

| | | | |
|---|---|---|---|
| 5 | - | - | - |
| - | 6 | 2 | - |
| 1 | 1 | 11 | - |
| - | 1 | - | 3 |

read 2

Minor disagreement 4 (13%)

Major disagreement 1 (3%)

4 pattern classification: κ = .76, 95%CI [.57-.95]
AD-like vs others: κ = .73, 95%CI [.48-.98]

**reader 2**

read 1

| | | | |
|---|---|---|---|
| 13 | 2 | - | - |
| 1 | 3 | - | - |
| - | - | 9 | - |
| - | - | 1 | 1 |

read 2

Minor disagreement 3 (10%)

Major disagreement 1 (3%)

4 pattern classification: κ = .79, 95%CI [.60-.98]
AD-like vs others: κ = .92, 95%CI [.77-1]

**Supplementary Figure 5. Illustration of major disagreement between the visual patterns assigned by the two readers.**

For both cases, reader 1 assigned pattern III (AD-like) while reader 2 assigned pattern I (negative). Each image is presented twice, with slightly different color scale adjustments to highlight the potentially ambiguous low-level signal in some cortical areas that was differentially interpreted by the two readers.

58 yo, Aβ-PET-negative control, global cortical SUVR= 1.10; Meta ROI SUVR= 1.13
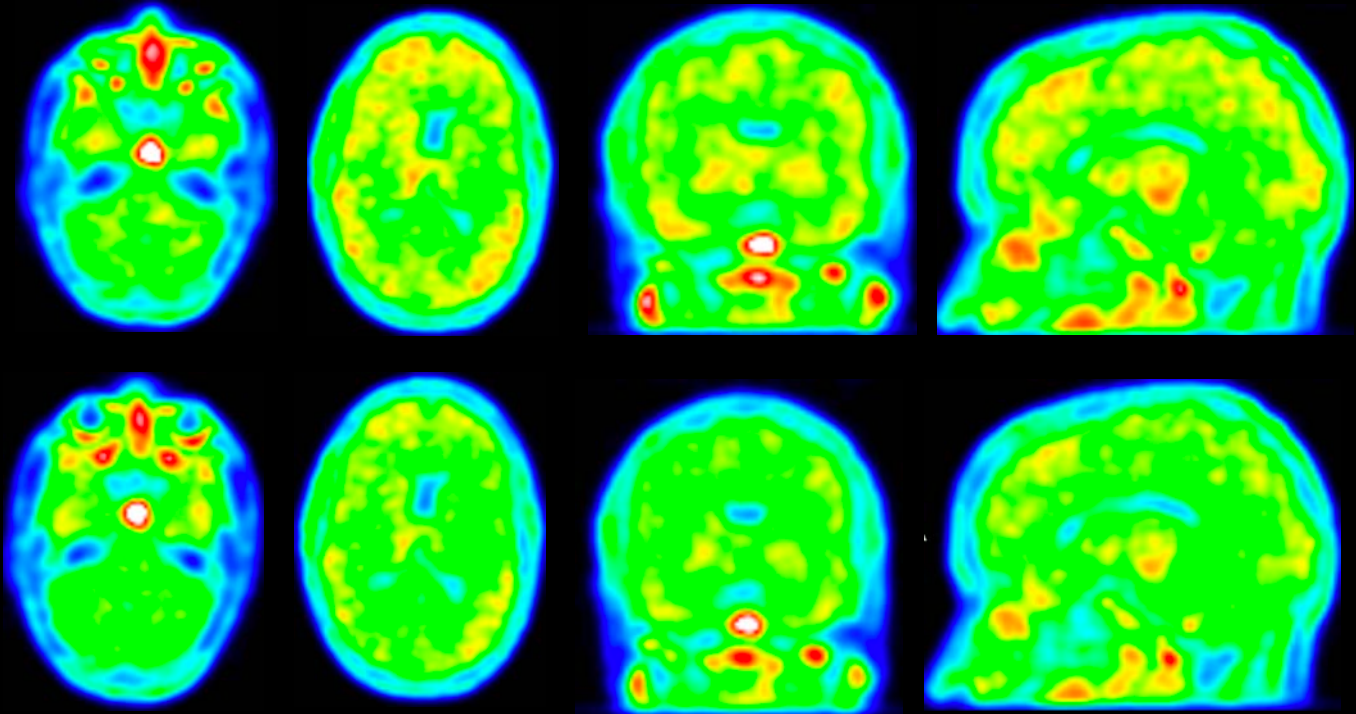


73 yo, Aβ-PET-negative patient with bvFTD, global cortical SUVR= 1.17; Meta ROI SUVR= 1.23
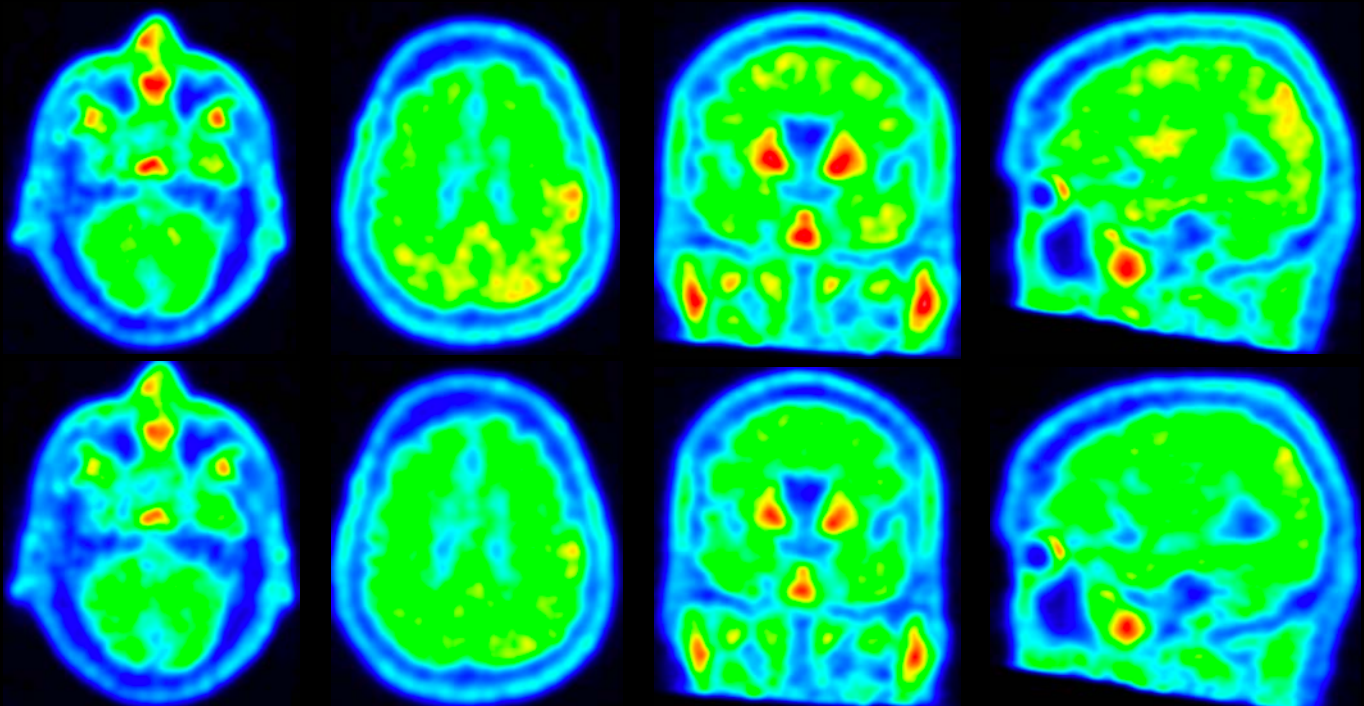
**Supplementary Figure 6. Relationships between the two indices derived from the FTP-PET visual reading scheme.**

Associations were tested for reader 1 (A), reader 2 (B), and for both readers combined (C). For each panel, plots illustrate the distribution of global scores within each visual pattern (x-axis: visual score, y-axis: density of scores within each visual pattern).
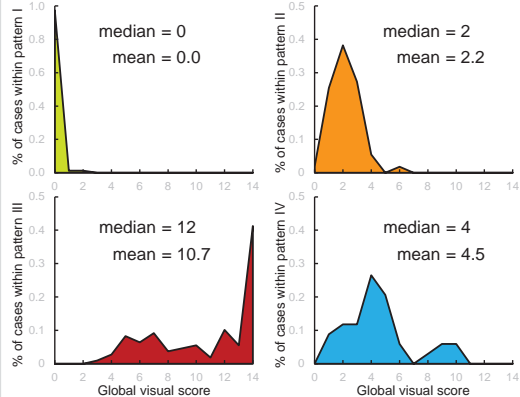
**Four visual FTP-PET patterns**

I Negative  II Mild Temporal only  III AD-like  IV Non-AD-like

**A**  **Reader 1**: pattern versus global score

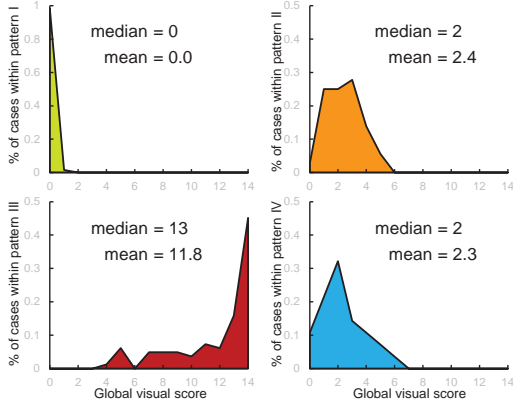**Kruskal-Wallis:** $\chi^2$=239.6 (df=3), p<$10^{-3}$, $\epsilon^2$=.88
All pairwise comparisons: p<.001

median = 0
mean = 0.0

median = 2
mean = 2.2

median = 12
mean = 10.7

median = 4
mean = 4.5

**B**  **Reader 2**: pattern versus global score

**Kruskal-Walli:** $\chi^2$=253.7 (df=3), p<$10^{-3}$, $\epsilon^2$=.93
All pairwise comparisons p<.001, except  II vs IV p=.94

median = 0
mean = 0.0

median = 2
mean = 2.4

median = 13
mean = 11.8

median = 2
mean = 2.3

**C**  **Both readers**: patterns versus average score

**Kruskal-Wallis:** $\chi^2$=240.7 (df=5), p<$10^{-3}$, $\epsilon^2$=.88

I vs all p<.001
III vs all p<.001

II vs IV p=.22
II vs grey p=.17
II vs black p=.06

IV vs grey p=.008
IV vs black p=1
grey vs black p<.001

*Consensus on negative pattern*
median = 0
mean = 0.0

*Consensus on temporal pattern*
median = 2.5
mean = 2.5

*Consensus on AD-like pattern*
median = 13.5
mean = 12.3

*Consensus on non-AD-like pattern*
median = 3.5
mean = 3.4

*Minor disagreement*
median = 1.5
mean = 2.0

*Major disagreement*
median = 3
mean = 3.9
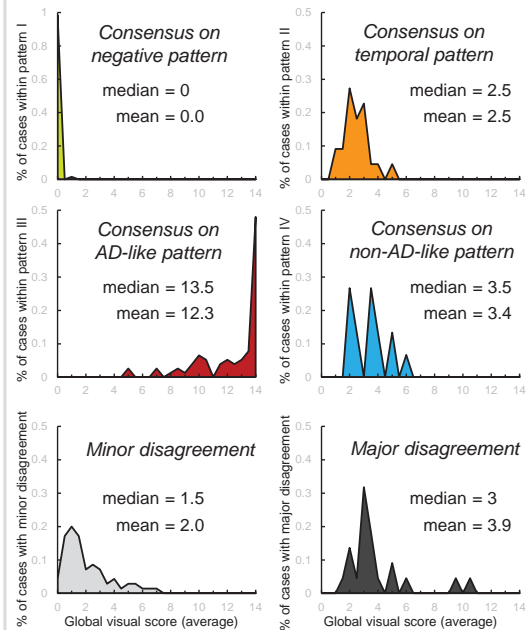
**Supplementary Figure 7. Distribution of visual FTP patterns in clinical and Aβ subgroups - combining patterns from the two readers.**

Left panel shows the distribution of visual patterns in each clinical/Aβ subgroup. Middle panel includes Sankey diagram illustrating the combination of visual patterns assigned by both readers (specific n indicated on top), and in Aβ-positive MCI/ADdem (circled numbers for each pattern, n=90 total) versus all other participants (n=184). Right panel shows the diagnostic properties of the AD-like pattern to distinguish Aβ-positive MCI/ADdem versus all other participants.

YC: Young Controls, OC: Older Controls, MCI: patients with a clinical diagnosis of Mild Cognitive Impairment, ADdem: patients with a clinical diagnosis of Alzheimer's disease dementia, non-AD: patients with a clinical diagnosis of non-Alzheimer's disease neurodegenerative syndrome. Clinical diagnoses were independent from Aβ status.
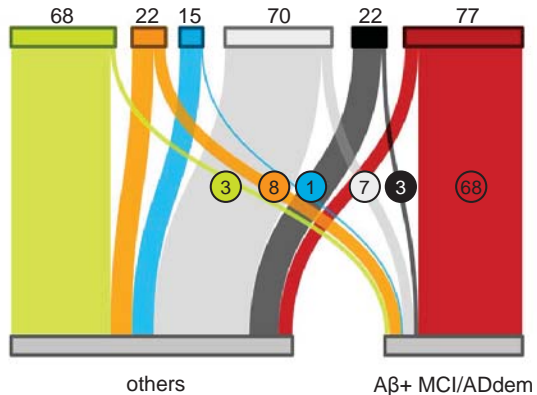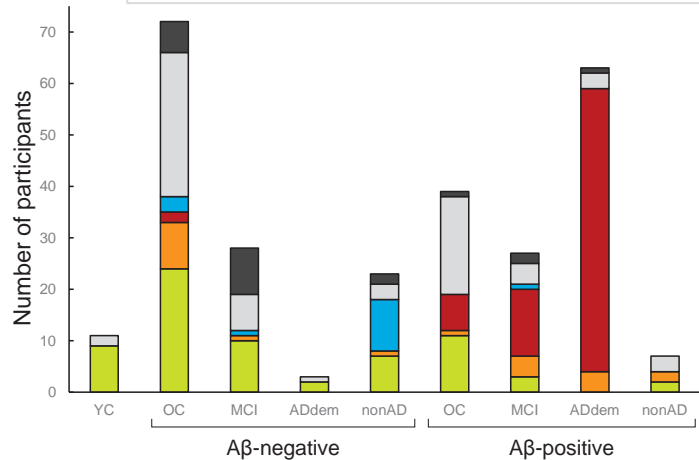
Se: Sensitivity, Sp: Specificity, PLR: positive likelihood ratio, NLR: Negative Likelihood ratio.