**Results**

We screened 106 subjects (5: not meeting inclusion criteria; 20: cannot comply with the study procedures) and eventually recruited 81 subjects (27 TRD, 27 nTRD, and 27 HCs) for the present study. Demographic data, including age, sex, and educational levels, did not differ between MDD and HC groups (Table 1). Regarding clinical variables and neurocognitive functioning, patients with MDD had significantly increased depressive symptoms (p < 0.001) on the basis of HDRS-17, decreased global cognitive function on the basis of MMSE (p = 0.039) and MoCA (p < 0.001), and decreased memory function on the basis of WLT 1st recall (p < 0.001) and 2nd recall (p = 0.006) (Table 1). If the TRD and nTRD groups were separately analyzed, results were similar, which revealed that patients with high depressive severity (i.e., TRD) had significantly high depressive symptoms (HDRS-17; TRD > nTRD > HC, p < 0.001), low global cognition (MoCA; TRD < nTRD = HC, p < 0.001), and low memory function (WLT 1st and 2nd recall; TRD < nTRD = HC, p < 0.001). No significant differences between TRD and nTRD were found for total illness duration, duration of the current episode, and ratios of late onset depression. In addition, no differences for the APOE gene were found across the 3 groups (Table 1). MDD patients with late-onset depression performed significantly worse than those with non-late-onset depression on the MMSE (T = -2.780, p = 0.008), MoCA (T = -2.028, p = 0.048), % errors of WCST (T = 2.047, p = 0.046), and % conceptual level responses of WCST (T = -2.190, p = 0.033). But, no differences of WLT 1st and 2nd recall were found between late-onset and non-late-onset depressives.

**3.1MDD versus HC: PIB SUVR and normalized brain glucose use**

Forty-two subjects (14 TRD, 14 nTRD, and 14 HCs) completed the PET scans. The results of neurocognitive functioning consistently showed that MDD patients had decreased global cognitive function (MMSE; MDD < HC, p < 0.001) and memory function (WLT 1st recall; MDD < HC, p = 0.001) (Table S1), and TRD patients had more impaired global function and memory function (Table S1). The voxel-wise comparisons of PIB scans showed that MDD had significantly higher PIB SUVR than HC in many brain regions, including the frontal cortex, parietal cortex, temporal cortex, occipital cortex, thalamus, brainstem, and even cerebellum, after adjusting for age, sex, education levels, PET machines, HDRS-17, and global counts (cluster-level FWE-corrected p < 0.001; Fig. 1). The most significant difference between MDD and HC was in the left middle frontal gyrus (peak MNI coordinate: −34, 10, 52; Fig. 1A: red arrows, Fig. 1B: a yellow circle; voxel-level FWE-corrected p < 0.05). No brain region showed significantly lower Aβ deposits for MDD than HC. If patients were divided into late-onset (onset ≥ 60 years old) and non–late-onset groups, no differences were observed between these groups in the frontal PIB SUVR [mean (standard deviation) for late-onset vs. non–late-onset depression = 1.15 (0.04) vs. 1.15 (0.09); T = −0.20, p = 0.984], temporal PIB SUVR [1.25 (0.06) vs. 1.26 (0.10); T = -0.325, p = 0.747], parietal PIB SUVR [1.14 (0.07) vs. 1.17 (0.14); T = -0.645, p = 0.525], or occipital PIB SUVR [1.41 (0.08) vs. 1.41 (0.12); T = 0.092, p = 0.928]. Furthermore, no differences were noted in the PIB SUVR among three different APOE genotypes [e.g., mean (standard deviation) frontal PIB SUVR for E2E3 vs. E3E3 vs. E3E4 = 1.12 (0.08) vs. 1.15 (0.08) vs. 1.09 (0.01); T = 0.798, p = 0.461]. For brain glucose use, no significant difference was found between the two groups under the threshold of either cluster-level or voxel-level FWE-corrected p < 0.05. However, if exploring at a cluster-level uncorrected p < 0.05, compared with HCs, patients with MDD had low glucose use in the bilateral frontal cortex, bilateral insula, and cerebellum and high use in the basal ganglia and parahippocampus/fusiform gyrus (Fig. S2).

**3.2TRD versus nTRD:** **PIB SUVR and normalized brain glucose use**

The voxel-wise PIB SUVR results showed that patients with TRD had more PIB deposits in the left-sided brain regions than did patients with nTRD after adjusting for age, sex, education levels, PET machines, HDRS-17, and global counts (cluster-level FWE-corrected p < 0.001; Fig. 2). The most prominent difference between the two groups existed in the left inferior parietal gyrus (peak MNI coordinate: −38, −36, 38; red arrows in Fig. 2A). Additionally, we compared 11C-PIB SUVRs between patients with nTRD and HCs (Fig. 2B) and between patients with TRD and HCs (Fig. 2C), revealing that both patients with TRD and nTRD have abnormally high Aβ deposits in many brain regions, including the frontal cortex, parietal cortex, temporal cortex, occipital cortex, thalamus, and even brainstem (cluster-level FWE-corrected p < 0.001). The identified brain regions were consistent with those found in the comparison between the MDD and HC groups (Fig. 1). If we compared 11C-PIB SUVR among the three groups, namely TRD, nTRD, and HC, only the left middle frontal gyrus (peak MNI coordinate: −34, 10, 52; Fig. 2D) survived after correcting for multiple comparisons (voxel-level FWE-corrected p < 0.05; post hoc analysis: TRD < HC, nTRD < HC). For brain glucose use, no significant difference was found between the TRD and nTRD groups or among the three groups under the threshold of either cluster-level or voxel-level FWE-corrected p < 0.05.

**3.3Differential correlations between PIB SUVR and neurocognition in HC and MDD**

The correlation analysis further showed that better neurocognitive functioning was significantly correlated with higher PIB SUVR in the prefrontal, parietal, and occipital cortex in the HC group (p < 0.05, Table 2), whereas opposite findings were found between neurocognitive functioning and PIB SUVR across all the cortical regions in the MDD group (p < 0.05, Table 2). For example, high MMSE scores, indicating high global cognitive functioning, were significantly correlated with high PIB SUVR in the prefrontal cortex in the HC group (r = 0.546, p < 0.05, dotted line in Fig. 3A), whereas low MMSE scores were correlated with high PIB SUVR in the prefrontal cortex in the MDD group (r = −0.395, p < 0.05, solid line in Fig. 3A). Similarly, high MoCA scores were significantly correlated with high PIB SUVR in the prefrontal cortex in the HC group (r = 0.766, p = 0.001, dotted line in Fig. 3B), whereas low MoCA scores were significantly correlated with high PIB SUVR in the prefrontal cortex in the MDD group (r = −0.381, p < 0.05, solid line in Fig. 3B). If the TRD and nTRD groups were separately analyzed, all the results were still negative correlations. Such an opposite finding of global cognition and PIB deposits between patients with MDD and HCs also existed for temporal cortex, parietal cortex, and occipital cortex (Fig. S3). Multiple linear regression analyses showed that PIB deposits in the prefrontal cortex predicted MoCA (β=20.055, T=2.977, p=0.018) in the HC group and MMSE (β=-12.686, T=-2.346, p=0.028) in the MDD group, but did not predicted MMSE (β=8.879, T=1.848, p=0.102) in the HC group and MoCA (β=-17.359, T=-1.546, p=0.136) in the MDD group. Furthermore, we investigated correlations between PIB SUVR and illness duration. None of the brain regions survived after correcting for multiple comparisons. However, exploratorily using a cluster-level uncorrected p < 0.001, a negative correlation between PIB SUVR and duration was observed in the bilateral cerebellum (Fig. S4) after adjusting for age, sex, education levels, PET machines, HDRS-17, and global counts. The result suggested that a late onset of LLD may lead to increased Aβ deposits in the cerebellum. In addition, if illness duration was entered as another covariate of no interest, the result of voxel-wise PIB SUVR analysis did not change, showing that patients with TRD had more PIB deposits in the left-sided brain regions than did patients with nTRD.

Table S1. Demographic data and clinical variables for PET scans

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | *HC* | *MDD* | *MDD* | | *HC vs. MDD*  *T (p-value)* | *HC vs. nTRD vs. TRD*  *F (p-value)* | *Post-hoc* |
| non-TRD  (n=14) | TRD  (n=14) |
| Age | 63.7 (6.0) | 64.53 (6.5) | 64.7 (5.7) | 65.1 (5.2) | 0.656 (0.516) | 0.353 (0.703) | - |
| Male/Female | 2/12 | 5/23 | 3/11 | 2/12 | 0.309 (0.967)b | 1.196 (0.694)b | - |
| Education | 13.3 (3.9) | 10.5 (4.1) | 10.4 (4.1) | 10.6 (4.3) | -1.638 (0.109) | 1.927 (0.160) | - |
| Duration (years) | - | 14.4 (12.3) | 11.8 (8.5) | 17.0 (15.1) | - | 1.287 (0.267)a | - |
| Late onset (>=60years) | - | 19 (35.2%) | 6 (42.9%) | 3 (21.4%) | - | 1.474 (0.420)b |  |
| Current episode  (<=12/12-24/>24months) | - | 12/2/14 | 7/1/6 | 5/1/8 | - | 0.876 (0.842)b |  |
| HDRS-17 | 1.8 (2.1) | 12.6 (10.2) | 6.7 (4.8) | 18.5 (10.9) | 5.275 (<0.001)\* | 18.9 (<0.001)\* | TRD>nTRD>HC |
| MMSE | 26.9 (1.3) | 26.4 (2.4) | 27.1 (1.7) | 25.6 (2.8) | -0.820 (0.417) | 2.181 (0.127) | - |
| MoCA | 24.9 (2.8) | 20.8 (5.0) | 22.5 (3.7) | 19.2 (5.7) | -3.618 (<0.001)\* | 5.782 (0.007)\* | TRD<nTRD=HC |
| WL 1st recall | 5.8 (1.1) | 4.1 (1.5) | 4.8 (1.4) | 3.5 (1.4) | 11.832 (0.001)\* | 10.045 (<0.001)\* | TRD<nTRD=HC |
| WL 2nd recall | 8.0 (2.9) | 5.9 (3.3) | 7.0 (3.8) | 4.9 (2.5) | 3.443 (0.071) | 3.449 (0.042)\* | TRD<HC |
| WCST % CLR | 35.4 (21.9) | 31.7 (25.8) | 35.1 (21.5) | 28.3 (30.0) | 0.192 (0.664) | 0.355 (0.704) |  |
| WCST CC | 2.0 (2.3) | 2.1 (2.1) | 2.4 (1.7) | 1.8 (2.5) | 0.021 (0.887) | 0.315 (0.732) |  |
| APOE  (E2E3/E3E3/E3E4) | 2/11/1 | 3/23/2 | 1/11/2 | 2/12/0 | 0.471 (1.000)b | 2.466 (0.841)b | - |

*Note.* MDD, major depressive disorder; HC, healthy control subjects; TRD, treatment-resistant depression; nTRD, non-TRD; HDRS-17, 17-item Hamilton depression rating scale; MMSE, mini-mental status examination; MoCA, Montreal cognitive assessment; APOE, apolipoprotein E gene

aP=0.446 as revealed by Mann-Whitney U analysis.

bFisher’s exact test was applied.

Fig S1. To improve the normalization accuracy, two group-specific MRI-aided 11C-PIB templates were created, one for PIB-negative participants and the other for PIB-positive participants, and used separately to normalize each participant’s PIB PET images. For PIB images, all voxel values were further normalized on the basis of the mean value of the cerebellar cortex (“Cerebellum\_Crus” in the AAL-ROI) to generate standardized uptake value ratio (SUVR)-scaled images (Akamatsu et al., 2016)

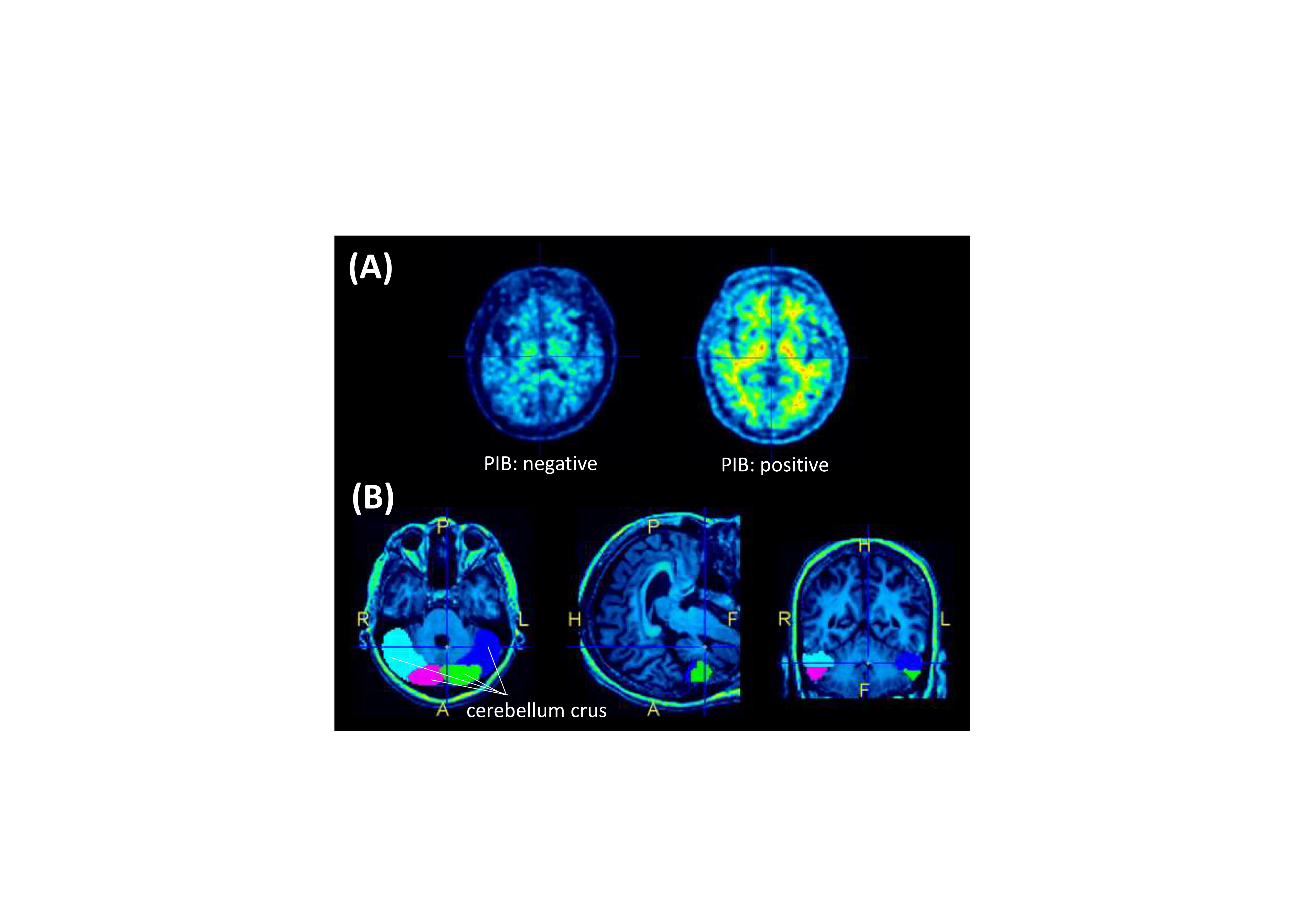


Fig S2. For brain glucose utilization, if exploring at a cluster-level uncorrected p < 0.05, compared with HCs, patients with MDD had low glucose use in the bilateral frontal cortex, bilateral insula, and cerebellum and high use in the basal ganglia and parahippocampus/fusiform gyrus

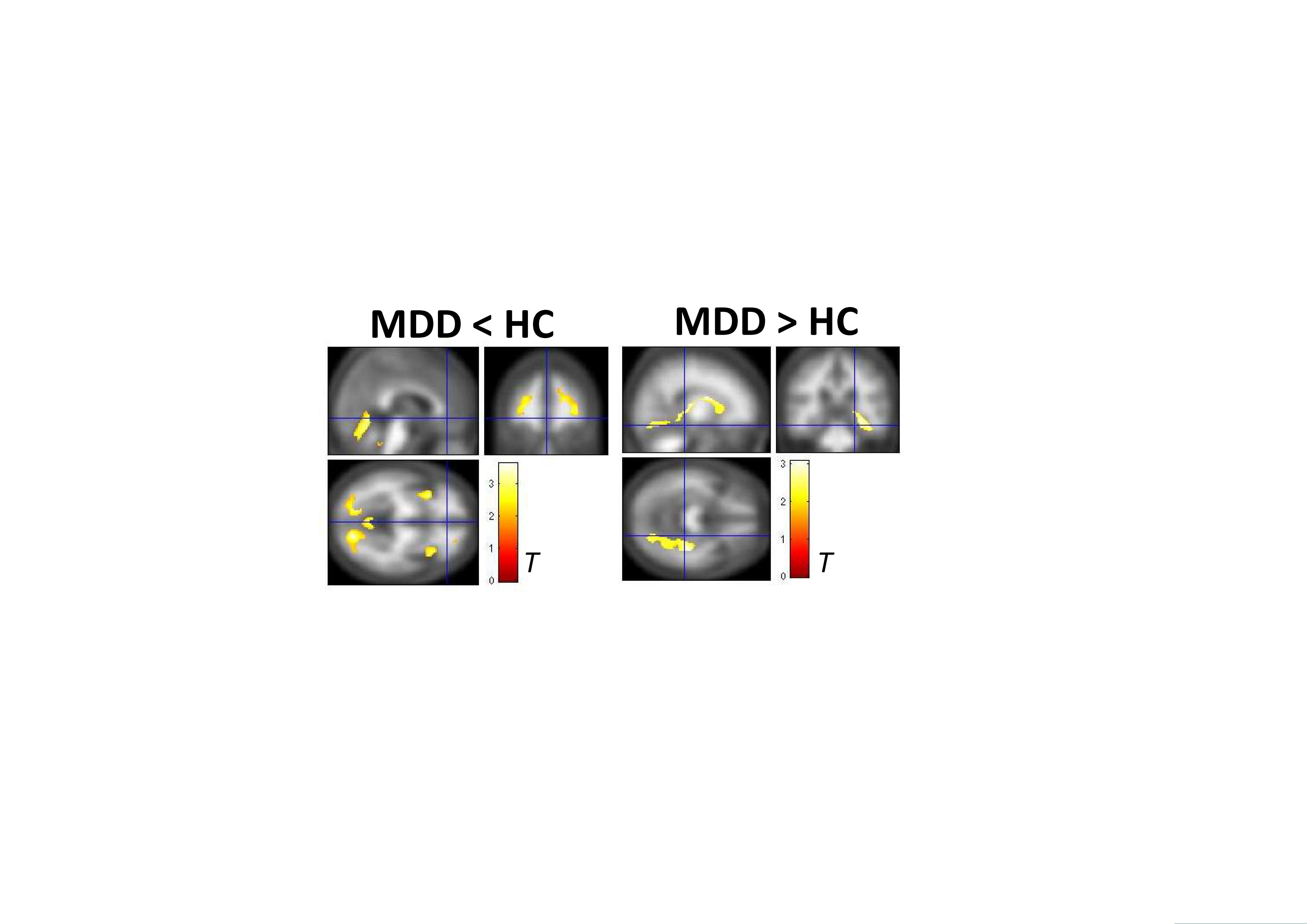


Fig S3. Correlations between MoCA and PIB SUVR in the prefrontal cortex (PFC), temporal cortex (TEMP), parietal cortex (PAR), and occipital cortex (OCC), consistently showing positive correlations in the HC group, but negative correlations in both of the TRD and nTRD group. SUVR, standardized uptake value ratio.

C:\Users\on508\Dropbox\980309 TEMP(self)\106年度_科技部MOST計畫_PIB FDG PET\manuscript_MDD_TRD_1090505\figures\Fig 4_MoCA_PIB_all cortex_01.tif

Fig S4. Regarding correlations between PIB SUVR and illness duration, if exploratorily using a cluster-level uncorrected p < 0.001, a negative correlation between PIB SUVR and duration was observed in the bilateral cerebellum.

