**Integrating brain imaging features and genomic profiles for the subtyping of major depression**

Liangying Yin1, Yuping Lin1, Jinghong Qiu1, Yong Xiang1, Ming Li2,10-12, Xiao Xiao2,10-12, Simon Sai-Yu Lui2-3, Hon-Cheong So1-2,5-9\*

1School of Biomedical Sciences, The Chinese University of Hong Kong, Shatin, Hong Kong

2KIZ-CUHK Joint Laboratory of Bioresources and Molecular Research of Common Diseases, Kunming Institute of Zoology and The Chinese University of Hong Kong, China

3Department of Psychiatry, The University of Hong Kong, Hong Kong, China

4Castle Peak Hospital, Hong Kong, China

5Department of Psychiatry, The Chinese University of Hong Kong, Hong Kong

6CUHK Shenzhen Research Institute, Shenzhen, China

7Margaret K.L. Cheung Research Centre for Management of Parkinsonism, The Chinese University of Hong Kong, Shatin, Hong Kong

8Brain and Mind Institute, The Chinese University of Hong Kong, Hong Kong SAR, China

9Hong Kong Branch of the Chinese Academy of Sciences Center for Excellence in Animal Evolution and Genetics, The Chinese University of Hong Kong, Hong Kong SAR, China

10 State Key Laboratory of Genetic Evolution & Animal Models, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, China

11 Yunnan Key Laboratory of Animal Models and Human Disease Mechanisms, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, China

12 Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan, China

\*Corresponding author Hon-Cheong SO, email: hcso@cuhk.edu.hk

**Supplementary information**

**Data**

For our subtyping analysis, we selected 69 depression-related clinical features with missing data rates below 10% across all MDD patients. These features encompassed demographics (e.g., sex), socioeconomic factors (e.g., Townsend deprivation index), and self-reported symptoms (e.g., nervous feelings, frequency of depressed mood in the last 2 weeks). Please refer to Table S1 for a detailed list.

**Sex-adjusted clustering analysis**

 We also conducted sex-adjusted clustering analysis for MDD patients. We identified 331, 293, 346 and 271 genes for cortex, frontal cortex, nucleus accumbens basal ganglia and putamen basal ganglia, respectively. Table S1.b lists the number of selected features in the corresponding data view. The best performance was achieved when the depression subjects were stratified into 2 different subgroups, with 20 subjects in the first subgroup and 332 (cluster 2) subjects in the second.

Table S3e-i illustrate the differences in neuroimaging and clinical features, as well as gene expression levels, across the identified subgroups from our sex-adjusted analysis. Additionally, based on results from the sex-adjusted analysis, we compared the frequencies of psychiatric admissions and treatment-resistant depression (TRD) between the two derived subgroups. We found significant differences in TRD (p=0.034) and hospitalization frequencies (p=0.046), as shown in Fig. S2 and S3.

**Comparison of the original and sex-adjusted clustering solutions**

Moreover, we compared the clustering results from the sex-adjusted analysis with our previous analysis that did not adjust for sex. The concordance of cluster labels identified from two solutions for the smaller and larger subgroups were respectively 50% and 97.0%.

Regarding the overlap of subtype-defining features, most clinical and brain imaging features were shared between the two clustering solutions. However, more genes were identified as subgroup-defining in the sex-adjusted analysis. This may be due to potentially better statistical power in covariate-adjusted models (Jiang et al., 2017). Table S3.e and Table S3.f-i present a comparison of the selected features across the two analyses. No significant differences were observed in treatment-resistant depression (TRD) or hospitalization frequencies between subgroups derived from the original and sex-adjusted clustering solutions. Overall, these findings suggest that adjusting for sex prior to clustering does not significantly affect the identification of MDD subgroups.

**Clustering analysis using the PCs with >90% variance explained in each view as input**

We also conducted clustering analysis using principal components (PCs) accounting for >90% variance in each view as input. Consistent with our earlier findings, we set the minimum subgroup size to 20. The best performance was achieved with 2 MDD subgroups, consisting of 20 and 332 subjects.

External validation showed no significant differences in outcome variables across the derived subgroups (Table S6). Given these results, the original clustering strategy, i.e., using PCA to select features, remains more effective. While the PC-as-input approach can better capture information from all variables, it also has the potential to include a larger number of irrelevant (clinical and genetic) features, which may affect the identification of distinct MDD subgroups. Nevertheless, dimension reduction (e.g. by PCA) prior to clustering analysis is an important direction for future studies, especially with very high-dimensional features.

**Supplementary Tables**

Table S1 Features (brain image features included) used for the subtyping analysis

Table S2Comparison results for subtype-defining clinical features

|  |  |
| --- | --- |
| Subtype 1 | Subtype 2 |
| Features | P-value | Features | P-value |
| Volume of grey matter in Frontal Pole (left) | 5.89E-08 | Volume of grey matter in Frontal Pole (left) | 5.89E-08 |
| Volume of grey matter in Frontal Pole (right) | 1.30E-05 | Volume of grey matter in Frontal Pole (right) | 1.30E-05 |
| Volume of grey matter in Insular Cortex (left) | 1.11E-04 | Volume of grey matter in Insular Cortex (left) | 1.11E-04 |
| Volume of grey matter in Insular Cortex (right) | 5.02E-04 | Volume of grey matter in Insular Cortex (right) | 5.02E-04 |
| Volume of grey matter in Superior Frontal Gyrus (left) | 3.11E-04 | Volume of grey matter in Superior Frontal Gyrus (left) | 3.11E-04 |
| Volume of grey matter in Superior Frontal Gyrus (right) | 4.66E-03 | Volume of grey matter in Superior Frontal Gyrus (right) | 4.66E-03 |
| Volume of grey matter in Middle Frontal Gyrus (left) | 2.35E-04 | Volume of grey matter in Middle Frontal Gyrus (left) | 2.35E-04 |
| Volume of grey matter in Middle Frontal Gyrus (right) | 3.56E-08 | Volume of grey matter in Middle Frontal Gyrus (right) | 3.56E-08 |
| Volume of grey matter in Precentral Gyrus (left) | 2.85E-03 | Volume of grey matter in Inferior Frontal Gyrus, pars opercularis (right) | 1.25E-04 |
| Volume of grey matter in Precentral Gyrus (right) | 1.40E-04 | Volume of grey matter in Precentral Gyrus (left) | 2.85E-03 |
| Volume of grey matter in Temporal Pole (left) | 2.39E-04 | Volume of grey matter in Precentral Gyrus (right) | 1.40E-04 |
| Volume of grey matter in Temporal Pole (right) | 1.61E-04 | Volume of grey matter in Temporal Pole (left) | 2.39E-04 |
| Volume of grey matter in Superior Temporal Gyrus, anterior division (right) | 2.77E-03 | Volume of grey matter in Temporal Pole (right) | 1.61E-04 |
| Volume of grey matter in Superior Temporal Gyrus, posterior division (right) | 2.26E-05 | Volume of grey matter in Superior Temporal Gyrus, posterior division (right) | 2.26E-05 |
| Volume of grey matter in Middle Temporal Gyrus, anterior division (left) | 2.62E-04 | Volume of grey matter in Middle Temporal Gyrus, anterior division (left) | 2.62E-04 |
| Volume of grey matter in Middle Temporal Gyrus, anterior division (right) | 7.79E-03 | Volume of grey matter in Middle Temporal Gyrus, posterior division (left) | 1.24E-04 |
| Volume of grey matter in Middle Temporal Gyrus, posterior division (left) | 1.24E-04 | Volume of grey matter in Middle Temporal Gyrus, posterior division (right) | 2.27E-04 |
| Volume of grey matter in Middle Temporal Gyrus, posterior division (right) | 2.27E-04 | Volume of grey matter in Postcentral Gyrus (left) | 2.91E-03 |
| Volume of grey matter in Postcentral Gyrus (left) | 2.91E-03 | Volume of grey matter in Postcentral Gyrus (right) | 4.03E-05 |
| Volume of grey matter in Postcentral Gyrus (right) | 4.03E-05 | Volume of grey matter in Lateral Occipital Cortex, superior division (left) | 5.54E-05 |
| Volume of grey matter in Supramarginal Gyrus, posterior division (right) | 1.37E-04 | Volume of grey matter in Lateral Occipital Cortex, superior division (right) | 1.23E-03 |
| Volume of grey matter in Lateral Occipital Cortex, superior division (left) | 5.54E-05 | Volume of grey matter in Lateral Occipital Cortex, inferior division (left) | 2.95E-06 |
| Volume of grey matter in Lateral Occipital Cortex, inferior division (right) | 2.70E-05 | Volume of grey matter in Lateral Occipital Cortex, inferior division (right) | 2.70E-05 |
| Volume of grey matter in Frontal Medial Cortex (left) | 9.62E-04 | Volume of grey matter in Subcallosal Cortex (left) | 7.64E-04 |
| Volume of grey matter in Frontal Medial Cortex (right) | 2.61E-03 | Volume of grey matter in Subcallosal Cortex (right) | 1.63E-03 |
| Volume of grey matter in Subcallosal Cortex (left) | 7.64E-04 | Volume of grey matter in Paracingulate Gyrus (left) | 2.68E-05 |
| Volume of grey matter in Subcallosal Cortex (right) | 1.63E-03 | Volume of grey matter in Paracingulate Gyrus (right) | 3.51E-04 |
| Volume of grey matter in Paracingulate Gyrus (left) | 2.68E-05 | Volume of grey matter in Cingulate Gyrus, posterior division (left) | 3.41E-05 |
| Volume of grey matter in Paracingulate Gyrus (right) | 3.51E-04 | Volume of grey matter in Cingulate Gyrus, posterior division (right) | 6.63E-05 |
| Volume of grey matter in Cingulate Gyrus, anterior division (right) | 1.98E-03 | Volume of grey matter in Precuneous Cortex (left) | 2.29E-05 |
| Volume of grey matter in Cingulate Gyrus, posterior division (left) | 3.41E-05 | Volume of grey matter in Precuneous Cortex (right) | 3.65E-05 |
| Volume of grey matter in Cingulate Gyrus, posterior division (right) | 6.63E-05 | Volume of grey matter in Cuneal Cortex (right) | 7.68E-04 |
| Volume of grey matter in Precuneous Cortex (left) | 2.29E-05 | Volume of grey matter in Frontal Orbital Cortex (left) | 1.81E-07 |
| Volume of grey matter in Precuneous Cortex (right) | 3.65E-05 | Volume of grey matter in Frontal Orbital Cortex (right) | 1.30E-05 |
| Volume of grey matter in Cuneal Cortex (left) | 8.96E-03 | Volume of grey matter in Parahippocampal Gyrus, anterior division (left) | 1.94E-03 |
| Volume of grey matter in Cuneal Cortex (right) | 7.68E-04 | Volume of grey matter in Parahippocampal Gyrus, anterior division (right) | 9.30E-03 |
| Volume of grey matter in Frontal Orbital Cortex (left) | 1.81E-07 | Volume of grey matter in Lingual Gyrus (left) | 3.75E-04 |
| Volume of grey matter in Frontal Orbital Cortex (right) | 1.30E-05 | Volume of grey matter in Lingual Gyrus (right) | 1.98E-04 |
| Volume of grey matter in Lingual Gyrus (left) | 3.75E-04 | Volume of grey matter in Temporal Fusiform Cortex, anterior division (left) | 1.78E-03 |
| Volume of grey matter in Lingual Gyrus (right) | 1.98E-04 | Volume of grey matter in Temporal Fusiform Cortex, anterior division (right) | 7.86E-04 |
| Volume of grey matter in Temporal Fusiform Cortex, posterior division (left) | 6.63E-03 | Volume of grey matter in Temporal Fusiform Cortex, posterior division (left) | 6.63E-03 |
| Volume of grey matter in Temporal Fusiform Cortex, posterior division (right) | 1.29E-02 | Volume of grey matter in Temporal Fusiform Cortex, posterior division (right) | 1.29E-02 |
| Volume of grey matter in Occipital Fusiform Gyrus (right) | 3.42E-03 | Volume of grey matter in Occipital Fusiform Gyrus (right) | 3.42E-03 |
| Volume of grey matter in Central Opercular Cortex (left) | 1.14E-05 | Volume of grey matter in Frontal Operculum Cortex (left) | 1.17E-03 |
| Volume of grey matter in Central Opercular Cortex (right) | 4.64E-05 | Volume of grey matter in Frontal Operculum Cortex (right) | 7.39E-04 |
| Volume of grey matter in Parietal Operculum Cortex (right) | 5.08E-05 | Volume of grey matter in Central Opercular Cortex (left) | 1.14E-05 |
| Volume of grey matter in Planum Polare (right) | 1.05E-04 | Volume of grey matter in Central Opercular Cortex (right) | 4.64E-05 |
| Volume of grey matter in Heschl's Gyrus (includes H1 and H2) (left) | 1.63E-03 | Volume of grey matter in Parietal Operculum Cortex (left) | 6.00E-04 |
| Volume of grey matter in Heschl's Gyrus (includes H1 and H2) (right) | 6.25E-04 | Volume of grey matter in Parietal Operculum Cortex (right) | 5.08E-05 |
| Volume of grey matter in Planum Temporale (left) | 1.29E-02 | Volume of grey matter in Planum Polare (left) | 5.24E-05 |
| Volume of grey matter in Planum Temporale (right) | 5.08E-05 | Volume of grey matter in Planum Polare (right) | 1.05E-04 |
| Volume of grey matter in Supracalcarine Cortex (left) | 1.54E-03 | Volume of grey matter in Heschl's Gyrus (includes H1 and H2) (left) | 1.63E-03 |
| Volume of grey matter in Supracalcarine Cortex (right) | 5.60E-06 | Volume of grey matter in Heschl's Gyrus (includes H1 and H2) (right) | 6.25E-04 |
| Volume of grey matter in Occipital Pole (left) | 3.96E-05 | Volume of grey matter in Planum Temporale (left) | 1.29E-02 |
| Volume of grey matter in Occipital Pole (right) | 3.53E-04 | Volume of grey matter in Planum Temporale (right) | 5.08E-05 |
| Volume of grey matter in Thalamus (left) | 5.66E-02 | Volume of grey matter in Supracalcarine Cortex (left) | 1.54E-03 |
| Volume of grey matter in Thalamus (right) | 3.27E-02 | Volume of grey matter in Supracalcarine Cortex (right) | 5.60E-06 |
| Volume of grey matter in Hippocampus (left) | 5.17E-04 | Volume of grey matter in Occipital Pole (left) | 3.96E-05 |
| Volume of grey matter in Hippocampus (right) | 5.08E-03 | Volume of grey matter in Occipital Pole (right) | 3.53E-04 |
| Volume of grey matter in Amygdala (right) | 6.36E-06 | Volume of grey matter in Hippocampus (left) | 5.17E-04 |
| Volume of grey matter in V Cerebellum (left) | 3.67E-02 | Volume of grey matter in Hippocampus (right) | 5.08E-03 |
| Volume of grey matter in V Cerebellum (right) | 3.03E-02 | Volume of grey matter in Amygdala (left) | 1.79E-04 |
| Volume of grey matter in VI Cerebellum (right) | 3.19E-02 | Volume of grey matter in Amygdala (right) | 6.36E-06 |

Table S3 Identified subtype-defining gene set in each tissue

Table S4 Enriched pathways for each depression subtype

Table S5 Enriched drugs for each depression subtype

Table S6 External validation result for clustering analysis the PCs with >90% variance explained in each view as the input to the bi-clustering model

|  |  |
| --- | --- |
| Clinical outcome | p-value |
| TRD | 0.261 |
| Hospitalization frequencies | 0.178 |

**Supplementary Figures**



Fig. S1 Model performance with varied subgroup numbers. BBD/WBD indicates the ratio of between bicluster and within bicluster distance. A larger BBD/WBD indicates better separated subgroups.



Fig. S2 Comparison of TRD status by subgroups for depression patients, based on the sex-adjusted analysis



Fig. S3 Comparison of admission frequencies by subgroups for depression patients, , based on the sex-adjusted analysis

**References**

Hastie, T. (2009). The elements of statistical learning: Data mining, inference, and prediction.

Jiang, H., Kulkarni, P. M., Mallinckrodt, C. H., Shurzinske, L., Molenberghs, G., & Lipkovich, I. (2017). Covariate adjustment for logistic regression analysis of binary clinical trial data. *Statistics in Biopharmaceutical Research, 9*(1), 126–134.