Exploring the interplay between metabolic syndrome and brain volume in depression: Basis for Phenotype-Based Classification

Sujin Gan[1], Narae Kim[1,3], Bumhee Park[2,3], and Rae Woong Park[1, 3]
[1] Department of Biomedical Sciences, Ajou University Graduate School of Medicine, Suwon, Korea
[2] Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, Korea
[3] Office of Biostatistics, Medical Research Collaborating Center, Ajou Research Institute for Innovative Medicine, Ajou University Medical Center, Suwon, South Korea

The rising prevalence of major depressive disorder (MDD) is a significant global burden. MDD is closely linked to metabolic diseases like metabolic syndrome (MetS) through shared mechanisms such as inflammation, oxidative stress, and neuroendocrine dysregulation, suggesting a bidirectional relationship where MetS can worsen MDD and vice versa. While the potential pathophysiological connections between MetS and depression have not been extensively investigated, one promising candidate is the link between brain structure, depression, and metabolic disturbances.

In this study, we hypothesized that integrating brain volume and clinical features to examine the connections between MetS and MDD could reveal distinct clinical subgroups, enhancing prognosis and interventions. We used non-negative matrix factorization (NMF) to identify brain morphology subtypes of MDD using regional mean brain volume data. The condensed NMF components were evaluated for their correlation with metabolic features and used as factors for classifying MetS in the classification model.

Methods

This study utilized an electronic health record (EHRs) database from Ajou University School of Medicine (AUSOM) spanning the years 1994 to 2023. The AUSOM database was converted to the Observational Medical Outcomes Partnership – Common Data Model (OMOP – CDM) format, a standardized medical database that supports distributed research networks. The sample in this study was selected based on the following stringent inclusion criteria: 1) A first-time diagnosis of MDD, 2) brain MRI procedures within 365 days before and 30 days after the index date, and 3) had records for a year before the index date. Individuals with a history of psychiatric comorbidities (bipolar disorder, schizophrenia, or dementia), substance disorders, and those with brain injury, hydrocephalus, atrophy, or abscess were excluded.

The selected study population was classified into two groups: those with and without MetS. MetS was defined as meeting more than 2 out of the following 5 MetS criteria: 1) Use of antihypertensive drugs or systolic blood pressure (SBP) > 130, 2) Use of antidiabetic drugs or blood sugar level (BST) > 100, 3) Low HDL-cholesterol levels (<40 mg/dL for men and <50 mg/dL for women), 4) Hypertriglyceridemia (Triglyceride $\geq$ 150 mg/dL), 5) BMI $\geq$ 30.

Structural T1-weighted MRI data were collected from all participants. Voxel-based morphometry (VBM) was used to extract the regional mean gray matter volume for 200 brain regions from each patient.

The MetS variables (SBP, DBP, BMI, BST, Triglyceride, HDL) within 12 months before and one month after the index date were extracted, and individuals with at most one missing value were excluded. Additionally, we included 38 variables with less than 25% missing data. Missing data were addressed using Multiple Imputation by Chained Equations (MICE).

We applied NMF for identifying condensed features from brain volume of 200 regions. NMF method provides an optimized determination of network spatial locations for the sample, and allows identification

of discrete subregions/networks within networks identified by Yeo et al. The optimal number of NMF components was determined based on laplace PCA method. We then examined the similarity of extracted NMF components with canonical resting-state network with confirmation by visual inspection.

A regularized canonical correlation analysis (rCCA) was used to define a low-dimensional projection from these associations for use in clustering. Then, to identify potential MDD subtypes, the two-dimensional variates (representing N-NMF features and 44 clinical variables, respectively) of the first component derived from the significant rCCA models were subjected to K-means clustering. The first component represents canonical correlations that describes the maximum covariation between the two multidimensional variables.

Utilizing the extracted features including demographic (age, sex), MetS determining variables, 9 VBM-NMF, cluster features, we developed four classification models to assess the contribution of various feature sets in predicting MetS in patients with MDD:

1) Model 1: Using demographic variables and MetS features
2) Model 2: Incorporating clustering features alongside demographic and MetS features
3) Model 3: Utilizing N- NMF components, in addition to demographic and MetS features.
4) Model 4: Including demographic, MetS, N-NMF components, and clustering features.

The models were individually developed using XGB machine learning algorithm. The study population was randomly split into the training (75%) and test set (25%) in 5-fold cross-validation. The performance of the two models was compared using accuracy (ACC), sensitivity, specificity, and the area under the receiver operating characteristic curve (AUROC).

All analyses, except those involving brain MRI scans, were conducted using R software version 3.6, OHDSI's Health Analytics Data to Evidence Suite packages, and various open-source R statistical packages. Statistical analyses of brain MRI scans were performed using MATLAB.

Results

A total of 150 patients was selected based on the inclusion and exclusion criteria, 76 patients with MetS and 74 without MetS (with Mets: 52 females [68.4%]; age year, mean [SD] 61.5 ± 13.8; without Mets: 53 females [71.6%]; age year, mean [SD] 56.2 ± 1.66;).

Using NMF, 200 brain region volumes were reduced into 9 components that correspond to distinct structural brain networks. These networks, including the Default Mode Network, Visual Network, Dorsal Attention Network, Somato-motor Network, Auditory Network, Frontoparietal Control Network, Limbic Network, Ventral Attention Network, and Cerebellar Network, captured key patterns of brain structure relevant to the study population (Figure 1).

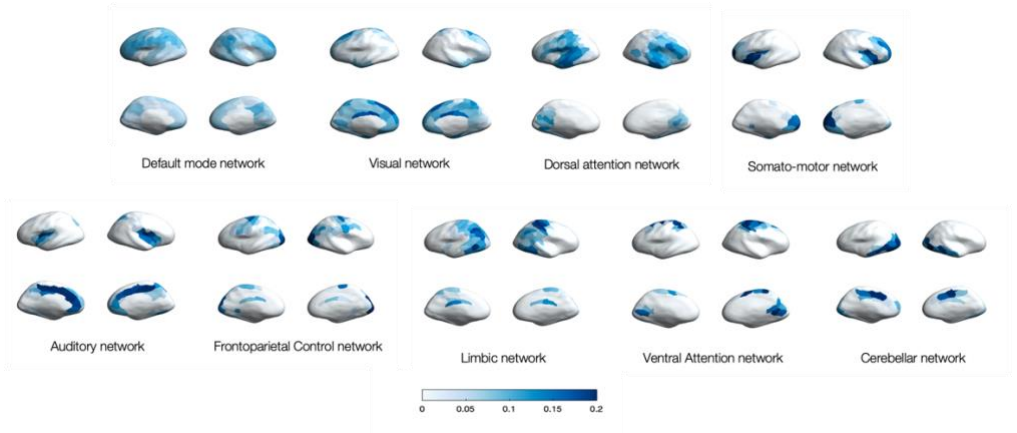

Figure 1. 9 Structural brain networks identified by non-negative matrix factorization analysis

The 9 NMF-derived brain networks were integrated with 44 clinical variables using CCA to investigate the relationship between neuroimaging data and clinical features. This integration revealed two distinct subtypes of patients. A K-means clustering algorithm was applied to these CCA components, and two clusters emerged with a mean silhouette score of 0.720.

Utilizing the NMF-derived brain features, we developed four models to classify MetS in patients with MDD. Model 4, which combined demographics, MetS, NMF features, and clustering, achieved the highest AUC of 0.79, indicating the strongest classification performance. Feature importance analysis using SHAP values revealed that BMI, triglyceride levels, and diastolic blood pressure were among the most influential features in predicting MetS. Additionally, brain regions such as VW5 and VW9 showed significant contributions. Higher feature values (e.g., elevated BMI and triglycerides) were associated with an increased likelihood of MetS (Table 1, Figure 2).

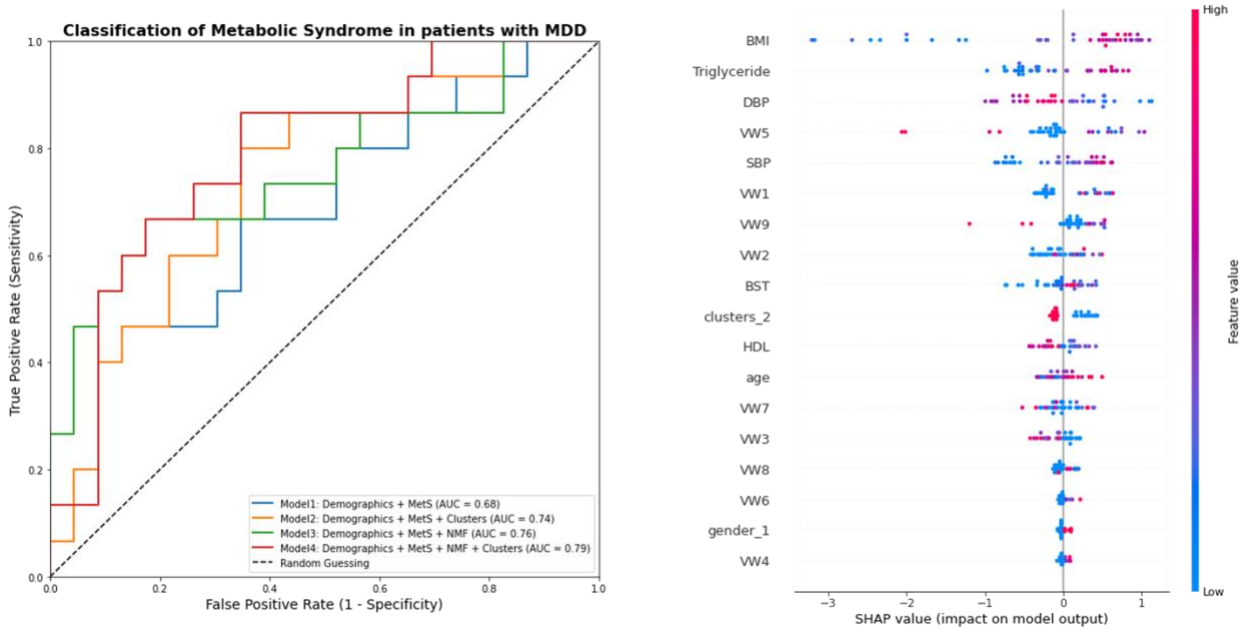

Figure 2. Model Comparison for Classifying Metabolic Syndrome in Patients with MDD and Feature Importance Using SHAP Values

Table 1. Classification model performance for metabolic syndrome in patients with depression

	Model 1	Model 2	Model 3	Model 4
AUROC	0.680	0.760	0.740	0.790
Accuracy	0.605	0.632	0.711	0.684
Precision	0.500	0.524	0.600	0.565
Recall	0.667	0.733	0.800	0.867
Specificity	0.565	0.565	0.652	0.565
F1-score	0.571	0.611	0.686	0.684

Conclusion

This study identified 9 brain components using NMF, revealing significant correlations with metabolic features. Integrating NMF-derived brain features with clinical variables improved the classification performance of MetS in MDD patients. These findings suggest that subgroups, defined by brain morphology and clinical features, may play a key role in understanding and managing metabolic conditions in this population.

Acknowledgement

This research was funded a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HR16C0001) and this research was supported by a Government-wide R&D Fund project for infectious disease research (GFID), Republic of Korea (grant number: HG22C0024, KH124685).

References

1. Lamers, F., Vogelzangs, N., Merikangas, K. et al. Evidence for a differential role of HPA-axis function, inflammation and metabolic syndrome in melancholic versus atypical depression. Mol Psychiatry 18, 692–699 (2013). https://doi.org/10.1038/mp.2012.144
2. An Pan, NaNa Keum, Olivia I. Okereke, Qi Sun, Mika Kivimaki, Richard R. Rubin, Frank B. Hu; Bidirectional Association Between Depression and Metabolic Syndrome: A systematic review and meta-analysis of epidemiological studies. Diabetes Care 1 May 2012; 35 (5): 1171–1180. https://doi.org/10.2337/dc11-2055