

A Comparative Study of White Matter, Gray Matter, and Cerebrospinal Fluid Volumes in Schizophrenia and Healthy Controls Using MRI

M. R. Meshram^{1*}, A. D. Sakhare²

¹ R.T.M. Nagpur University, Nagpur, India

² Department of Electronics and Computer Science, R.T.M. Nagpur University, Nagpur, India

*Corresponding author: mihir25_2010@rediffmail.com

Abstract

Background: Schizophrenia has been associated with structural abnormalities of the brain including CSF, GM, and WM. The quantification of brain volume based on T1-weighted MRI together with machine learning will provide more information regarding the abnormalities in order to facilitate classification.

Objectives: To investigate the differences in whole-brain CSF, GM, and WM volumes in schizophrenia patients and healthy controls, and to assess the discriminating capabilities of eight different supervised machine learning classifiers based on the volumetric features.

Methods: T1-weighted MRI scans from 97 individuals with schizophrenia and 93 healthy controls (total N = 190) were obtained from the OpenNeuro repository. Tissue volumes were extracted using intensity-based K-means segmentation and expressed in milliliters. Group difference were assessed using the Welch's t-test and Mann-Whitney U test. Classification was done on eight algorithms(SVM,RF,XGB,KNN,LR,DT,NB,GB) based on the following performance metrics:accuracy,sensitivity,specificity,F1 score and ROC-AUC.

Results: CSF volume was significantly higher in the schizophrenia group (334.61 ± 76.25 mL) than in healthy controls (306.87 ± 87.35 mL; $t = -2.329$, $p = 0.021$; Mann-Whitney $p = 0.0065$), corresponding to a 9.04% increase. Differences in GM and WM volumes were small and not consistently significant. Random Forest obtained the highest classification performance (Accuracy = 0.753, Precision = 0.784, Sensitivity = 0.711, Specificity = 0.796, F1-score = 0.746, ROC-AUC = 0.780).

Conclusion: Elevated CSF volume represents the most robust volumetric finding in this cohort. Global tissue volumes provide moderate discriminative information, with Random Forest emerging as the most effective classifier. These results support the utility of structural MRI volumetry as a complementary approach in schizophrenia research.

Keywords: Schizophrenia(SZ),Gray matter(GM), White matter(WM),Cerebrospinal fluid(CSF), Machine learning(ML), Random Forest(RF)

I. INTRODUCTION

Schizophrenia is a serious and heterogeneous neuropsychiatric disorder that affects about 1% of the world population [5], usually appears in late adolescence or early adulthood. It is characterized by positive symptoms (hallucinations and delusions), negative symptoms (anhedonia and avolition) and cognitive deficits that seriously impair social and occupational functioning[9].

The MRI technique has played a great role in the identification of the neuroanatomy related to schizophrenia.The results that have been consistently repeated are the enlargement of the lateral and third ventricles through an increase in the volume of CSF, along with a decrease in the volume of GM in frontal, temporal, and limbic regions [6][10]. White matter

(WM) changes have been reported but are not consistent and are thought to indicate disrupted connections [3][7].

Global tissue volumetry offers a straightforward, reproducible, and computationally efficient approach to quantifying these changes. When combined with supervised machine learning, volumetric features can be evaluated for their ability to discriminate patients from healthy individuals. Ensemble methods such as Random Forest and gradient boosting have shown particular promise in neuroimaging classification tasks because of their capacity to model non-linear relationships and interactions among features [1][12].

The present study had three primary objectives:

- (1) to quantify and compare whole-brain CSF, GM, and WM volumes between

individuals with schizophrenia and healthy controls;

- (2) to evaluate the statistical significance of these volumetric differences using both parametric and non-parametric tests; and
- (3) to systematically compare the classification performance of eight supervised machine learning algorithms using CSF, GM, and WM volumes as input features.

II. DATA AND EXPERIMENTAL PROCEDURE

1. Participants and Dataset

A total of 190 high-resolution T1-weighted MRI scans were included in the analysis. The schizophrenia group consisted of 97 individuals, and the healthy control group consisted of 93 age-comparable participants. All images were obtained from the publicly available OpenNeuro neuroimaging repository. Only scans acquired in the standard anatomical orientation and of sufficient quality for volumetric analysis were retained. No additional demographic or clinical metadata (for example, age, sex, illness duration, or medication status) were available in the dataset.

2. Image Preprocessing and Tissue Segmentation

All T1-weighted MRI volumes were analyzed through a pipeline in Python. The following packages were utilized in the process: nibabel library to handle NIfTI format, NumPy and SciPy libraries for numerical calculations, scikit-learn library for machine learning and clustering and pandas library for data manipulation.

The non-zero voxels were selected from each T1-weighted volume to build an approximate mask of the brain area thus removing background air. Segmentation of tissue based on intensity was done by applying the K-means clustering algorithm with $k = 3$. Three clusters that emerged from this procedure were sorted by their means: the cluster with the smallest mean was assigned as CSF, the second smallest cluster was GM and the third one with the largest mean value was assigned as WM. The voxel counts for each tissue type were then transformed into the absolute volume measured in milliliters (mL). Since the volume of the intracranium (ICV) was not available for all patients, all volumes were analyzed in absolute terms.

3. Statistical evaluation

Descriptive statistics (mean, standard deviation, and median) were calculated for CSF, GM, and WM volumes within each diagnostic group. Group differences were evaluated using Welch's t-test, which does not assume equal variances, and the Mann-Whitney U test as a non-parametric alternative.

Percentage differences were computed relative to the healthy control mean. Statistical significance was defined as $p < 0.05$. All statistical analyses were performed using SciPy.

4. Machine Learning Classification

Three volumetric features (CSF, GM, and WM volumes) served as input variables. The target variable was diagnostic group (Schizophrenia versus Healthy). The dataset was partitioned into stratified training (70%) and test (30%) sets to preserve class distribution. Features were standardized using z-score normalization based on training-set statistics.

Eight supervised classifiers were trained and evaluated: Support Vector Machine (SVM) with a radial basis function kernel, Random Forest (150 trees), XGBoost, K-Nearest Neighbors ($k = 5$), Logistic Regression, Decision Tree, Gaussian Naive Bayes, and Gradient Boosting. Because the two classes were nearly balanced (97 schizophrenia versus 93 healthy), no additional class-weighting or resampling techniques were applied.

Model performance was assessed on the held-out test set using accuracy, precision, sensitivity (recall), specificity, F1-score, and area under the receiver operating characteristic curve (ROC-AUC). All machine learning experiments were conducted using scikit-learn and XGBoost in Python 3.10.

III. RESULTS

1. Dataset Characteristics

The final dataset encompassed 190 subjects (schizophrenia: $n = 97$; healthy controls: $n = 93$).

2. Group Differences in Tissue Volumes

Table 1 summarizes the volumetric measurements for both groups. Patients with schizophrenia exhibited a significantly higher mean CSF volume (334.61 ± 76.25 mL) compared with healthy controls (306.87 ± 87.35 mL), corresponding to an absolute difference of 27.74 mL (9.04% increase). Differences in GM (+8.67 mL, +1.37%) and WM (+14.81 mL, +2.89%) volumes were modest.

TABLE 1. DESCRIPTIVE STATISTICS OF TISSUE VOLUMES (ML) BY GROUP

Tissue	Group	n	Mean	SD	Median
CSF	Healthy	93	306.87	87.35	296.86
CSF	Schizophrenia	97	334.61	76.25	332.64
GM	Healthy	93	633.62	107.69	619.77
GM	Schizophrenia	97	642.29	115.61	636.69
WM	Healthy	93	513.32	48.79	509.30
WM	Schizophrenia	97	528.13	53.78	520.28

Note. Volumes are expressed in milliliters (mL). SD = standard deviation.

Table 2 presents the results of formal statistical comparisons. CSF volume differed significantly between groups on both Welch's t-test ($t = -2.329$, $p = 0.021$) and the Mann-Whitney U test ($p = 0.0065$). WM volume showed a borderline difference on the parametric test ($p = 0.048$) but did not reach significance on the non-parametric test ($p = 0.086$). GM volume did not differ significantly between groups.

TABLE 2. STATISTICAL COMPARISON OF TISSUE VOLUMES BETWEEN GROUPS

Note. $*p < 0.05$. Welch's t-test assumes unequal variances.

3. Correlation Among Tissue Volumes

The Pearson correlation test showed that there was a very high correlation between GM and CSF volumes ($r = 0.86$), a moderate correlation between GM and WM ($r = 0.51$), and a moderate correlation between WM and CSF ($r = 0.45$). Figure 1 shows the correlation matrix.

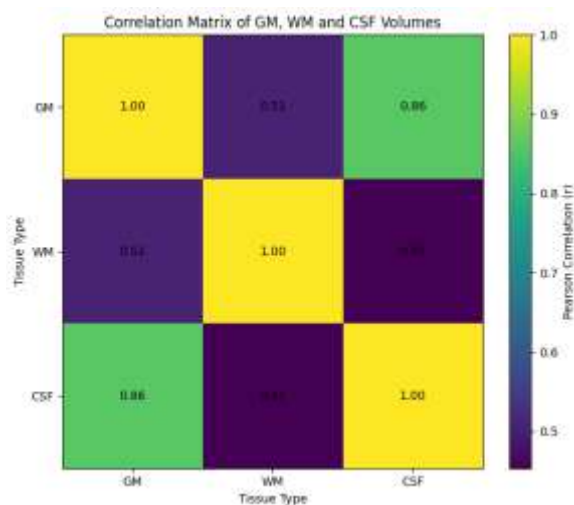


FIG. 1. CORRELATION MATRIX OF GM, WM, AND CSF VOLUMES

Figure 1. Heatmap displaying Pearson correlation coefficients among gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) volumes. A strong positive correlation is observed between GM and CSF ($r = 0.86$).

4. Machine Learning Classification performance

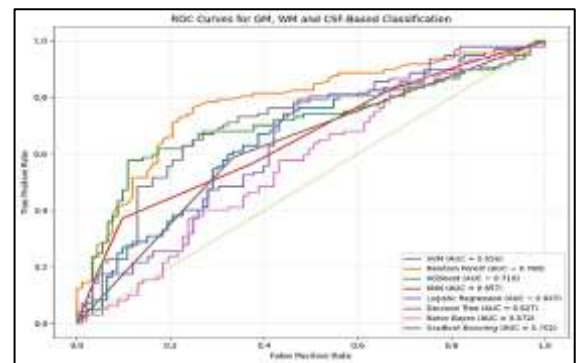
Classification results for the eight algorithms are summarized in Table 3. Random Forest achieved the highest overall performance across nearly all metrics (Accuracy = 0.753, Precision = 0.784, Sensitivity = 0.711, Specificity = 0.796, F1-score = 0.746, ROC-AUC = 0.780). Gradient Boosting and XGBoost ranked second and third, respectively. Naive Bayes yielded the lowest performance.

TABLE 3. CLASSIFICATION PERFORMANCE OF EIGHT SUPERVISED ALGORITHMS

Classifier	Accuracy	Precision	Sensitivity	Specificity	F1-score	ROC-AUC
Random Forest	0.753	0.784	0.711	0.796	0.746	0.780
Gradient Boosting	0.695	0.701	0.701	0.688	0.701	0.701
XGBoost	0.658	0.660	0.680	0.634	0.670	0.710
SVM	0.637	0.632	0.691	0.581	0.660	0.656
Decision Tree	0.626	0.648	0.588	0.667	0.616	0.627
Logistic Regression	0.611	0.614	0.639	0.581	0.626	0.647
KNN	0.595	0.611	0.567	0.624	0.588	0.657
Naive Bayes	0.553	0.567	0.526	0.581	0.546	0.572

Note. Random Forest achieved the best performance across most metrics.

Receiver Operating Characteristic(ROC) curves for all classifiers are presented in Figure 2. The Random Forest classifier produced the highest area under the curve (AUC = 0.780), indicating good discriminative ability.



Tissue	t-statistic	Welch p	Mann-Whitney U	Mann-Whitney p	Interpretation
CSF	-2.329	0.021*	3479.0	0.0065*	Higher in Schizophrenia
GM	-0.535	0.593	4387.0	0.745	No significant difference
WM	-1.989	0.048*	3860.0	0.086	Borderline / trend

FIG. 2. ROC CURVES FOR GM, WM, AND CSF-BASED CLASSIFICATION

Figure 2. Receiver operating characteristic (ROC) curves for eight classifiers trained on CSF, GM, and WM volumes. Random Forest achieved the highest AUC (0.780), followed by XGBoost (0.710) and Gradient Boosting (0.701).

IV. DISCUSSION

A statistically significant rise in cerebrospinal fluid volumes in schizophrenics relative to healthy subjects is considered to be the main result of the presented research. The established deviation between groups is 9.04%, which is in line with multiple papers devoted to ventricular dilatation as one of the most common abnormalities in schizophrenia patients revealed via imaging studies (Shenton et al., 2001; Haijma et al., 2013; van Erp et al., 2016). Such increase is typically associated with the surrounding tissues atrophy or abnormal neurodevelopment that leads to increased CSF volumes. Contrary to many other findings in the existing literature, no significant differences in global gray and white matter volumes were reported. There are several potential reasons for this. First, whole-brain volumetric analysis is less effective when measuring regionally specific atrophy (in prefrontal cortex, superior temporal gyrus, or hippocampus), which is usually found in the schizophrenia patients (Cropley et al., 2017; Lee et al., 2020). Second, intensity-based K-means approach provides faster processing but is less accurate in terms of anatomy than such approaches as FreeSurfer, FSL, and deep-learning approaches like SynthSeg. Third, clinical heterogeneity, medications, and disease duration (none of which are included in the current database) can affect the results of the volumetric analysis.

Of the eight machine learning algorithms that were assessed, Random Forest demonstrated superior performance. The ability of the algorithm to model non-linear relationships between the three tissue volumes, along with its resistance to feature scaling, are some possible factors responsible for the superiority. With ROC-AUC of 0.780, the model has been shown to have moderate discriminative power with a relatively modest number of features. Non-linear decision boundaries, as applied in ensemble tree-based models (Random Forest, Gradient Boosting, and XGBoost) performed better than linear and instance-based models in this feature space (Arbabshirani et al., 2017; Vieira et al., 2020). The highly significant positive correlation between the volumes of GM and CSF tissue ($r = 0.86$) is important and may be explained by the common variance attributable to the total size of the brain. Further research should take into account the adjustment of volumes with respect to intracranial volume.

V. LIMITATIONS

The method used for segmentation is fairly simple and could lead to a systemic bias. The fact that this research employs a cross-sectional design limits the inference that can be drawn on progression or longitudinal changes. There is no clinical metadata included such as age, gender, period of illness, medicated versus unmedicated, and symptoms.

VI. CONCLUSION

This study reports significant changes in cerebrospinal fluid (CSF) volumes in people with schizophrenia and shows that global tissue volumes have sufficient discriminative power. Out of the classifiers that were tested, the one that performed best was the Random Forest classifier, with accuracy of 75.3% and ROC-AUC of 0.780. The study highlights the importance of structural MRI volumetry and also sheds light on the usefulness of ensembles in neuroimaging classification tasks. Future research should focus on employing more sophisticated techniques in segmentation, regional morphometry analysis, and using multiple modality features.

VII. FUTURE RESEARCH DIRECTIONS

1. Implementation of sophisticated segmentation pipelines (such as FreeSurfer, FastSurfer, SynthSeg, or SPM) to increase the accuracy of tissue segmentation.
2. Introduction of regional volumes, cortical thickness measures, and surface-based morphometry.
3. Multimodal analysis combining structural, diffusion, and functional MRI measurements.
4. Consideration of additional demographic and clinical covariates for increased generalizability of the model.
5. Assessment of reproducibility in large-scale studies involving multiple sites and scanner types.

VIII. REFERENCES

- [1] M. R. Arbabshirani, S. Plis, J. Sui, and V. D. Calhoun, "Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls," *NeuroImage*, vol. 145, pp. 137–165, 2017, doi: 10.1016/j.neuroimage.2016.02.079.
- [2] M. S. Bagary, M. R. Symms, G. J. Barker, S. H. Mutsatsa, E. M. Joyce, and M. A. Ron, "Gray and white matter brain abnormalities in first-episode schizophrenia inferred from magnetization transfer imaging," *Archives of General Psychiatry*, vol. 60, no. 8, pp. 779–788, 2003, doi: 10.1001/archpsyc.60.8.779.
- [3] V. L. Cropley, P. Klauser, R. K. Lenroot, J. Bruggemann, S. Sundram, C. Bousman, A. Pereira, M. A. Di Biase, T. W. Weickert, C. S. Weickert, C. Pantelis, and A. Zalesky, "Accelerated gray and white matter deterioration with age in schizophrenia," *American Journal of Psychiatry*, vol. 174, no. 3, pp. 286–295, 2017, doi: 10.1176/appi.ajp.2016.16050610.
- [4] R. De Filippis, E. Carbone, R. Gaetano, A. Bruni, V. Pugliese, C. Segura-Garcia, and P. De Fazio, "Machine learning techniques in a structural and functional MRI diagnostic approach in schizophrenia: A systematic review," *Neuropsychiatric Disease and Treatment*, vol. 15, pp. 1605–1627, 2019, doi: 10.2147/NDT.S202418.
- [5] H. Göker, "1D-convolutional neural network approach and feature extraction methods for automatic detection of schizophrenia," *Signal*,

Image and Video Processing, vol. 17, no. 5, pp. 2627–2636, 2023, doi: 10.1007/s11760-022-02479-7.

[6] S. V. Haijma, N. van Haren, W. Cahn, P. C. Koolschijn, H. E. Hulshoff Pol, and R. S. Kahn, “Brain volumes in schizophrenia: A meta-analysis in over 18,000 subjects,” *Schizophrenia Bulletin*, vol. 39, no. 5, pp. 1129–1138, 2013, doi: 10.1093/schbul/sbs118.

[7] D. K. Lee, H. Lee, K. Park, E. Joh, C. E. Kim, and S. Ryu, “Common gray and white matter abnormalities in schizophrenia and bipolar disorder,” *PLoS ONE*, vol. 15, no. 5, Art. no. e0232826, 2020, doi: 10.1371/journal.pone.0232826.

[8] J. Oh, B. L. Oh, K. U. Lee, J. H. Chae, and K. Yun, “Identifying schizophrenia using structural MRI with a deep learning algorithm,” *Frontiers in Psychiatry*, vol. 11, Art. no. 16, 2020, doi: 10.3389/fpsyt.2020.00016.

[9] N. N. Raju, “Psychiatry training in India,” *Indian Journal of Psychiatry*, vol. 64, no. 5, pp. 433–439, 2022, doi: 10.4103/indianjpsychiatry.indianjpsychiatry_649_22.

[10] M. E. Shenton, C. C. Dickey, M. Frumin, and R. W. McCarley, “A review of MRI findings in schizophrenia,” *Schizophrenia Research*, vol. 49, nos. 1–2, pp. 1–52, 2001, doi: 10.1016/S0920-9964(01)00163-3.

[11] T. G. M. van Erp, D. P. Hibar, J. M. Rasmussen, D. C. Glahn, G. D. Pearlson, O. A. Andreassen, *et al.*, “Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium,” *Molecular Psychiatry*, vol. 21, no. 4, pp. 547–553, 2016, doi: 10.1038/mp.2015.63.

[12] S. Vieira, Q. Y. Gong, W. H. L. Pinaya, C. Scarpazza, S. Tognin, B. Crespo-Facorro, *et al.*, “Using machine learning and structural neuroimaging to detect first-episode psychosis: Reconsidering the evidence,” *Schizophrenia Bulletin*, vol. 46, no. 1, pp. 17–26, 2020, doi: 10.1093/schbul/sby189.