Namgung *et al.*
The Journal of Headache and Pain (2025) 26:5
https://doi.org/10.1186/s10194-024-01946-5

The Journal of Headache and Pain

RESEARCH

Open Access

A robust multimodal brain MRI-based diagnostic model for migraine: validation across different migraine phases and longitudinal follow-up data

Jong Young Namgung[1], Eunchan Noh[2], Yurim Jang[3], Mi Ji Lee[4,5*] and Bo-yong Park[6,7*]

Abstract

Inter-individual variability in symptoms and the dynamic nature of brain pathophysiology present significant challenges in constructing a robust diagnostic model for migraine. In this study, we aimed to integrate different types of magnetic resonance imaging (MRI), providing structural and functional information, and develop a robust machine learning model that classifies migraine patients from healthy controls by testing multiple combinations of hyperparameters to ensure stability across different migraine phases and longitudinally repeated data. Specifically, we constructed a diagnostic model to classify patients with episodic migraine from healthy controls, and validated its performance across ictal and interictal phases, as well as in a longitudinal setting. We obtained T1-weighted and resting-state functional MRI data from 50 patients with episodic migraine and 50 age- and sex-matched healthy controls, with follow-up data collected after one year. Morphological features, including cortical thickness, curvature, and sulcal depth, and functional connectivity features, such as low-dimensional representation of functional connectivity (gradient), degree centrality, and betweenness centrality, were utilized. We employed a regularization-based feature selection method combined with a random forest classifier to construct a diagnostic model. By testing the models with varying feature combinations, penalty terms, and spatial granularities within a strict cross-validation framework, we found that the combination of curvature, sulcal depth, cortical thickness, and functional gradient achieved a robust classification performance. The model performance was assessed using the test dataset and achieved 87% accuracy and 0.94 area under the curve (AUC) at distinguishing migraine patients from healthy controls, with 85%, 0.97 and 84%, 0.93 during the interictal and ictal/peri-ictal phases, respectively. When validated using follow-up data, which was not included during model training, the model achieved 91%, 94%, 89% accuracies and 0.96, 0.94, 0.98 AUC for the total, interictal, and ictal/peri-ictal phases, respectively, confirming its robustness. Feature importance and clinical association analyses exhibited that the somatomotor, limbic, and default mode regions could be reliable markers of migraine. Our findings, which demonstrate a robust diagnostic performance using multimodal MRI features and a machine-learning framework, may offer a valuable approach for clinical diagnosis across diverse cohorts and help alleviate the decision-making burden for clinicians.

Keywords Migraine, Diagnosis model, Multimodal MRI, Machine learning, Longitudinal validation

*Correspondence:
Mi Ji Lee
mijilee.md@snu.ac.kr
Bo-yong Park
boyongpark@korea.ac.kr
Full list of author information is available at the end of the article

Introduction

Migraine is a neurological disease with a significant global impact, affecting approximately 15% of the world's population. In patients with migraine, symptoms such as nausea, vomiting, and hypersensitivity to light and sound commonly accompany headaches during the ictal phase [1]. Patients with migraine commonly experience interictal symptoms suggestive of altered sensory and autonomic processing, including increased sensitivity to light, sound, smell, thermal stimuli, and motion, and delayed gastric emptying [2–8]. Currently, the diagnosis of migraine relies solely on clinical symptoms. Although the International Classification of Headache Disorders (ICHD) provides a standardized framework for classifying migraine among primary headache disorders, it does not offer a pathophysiology-based diagnostic approach [9]. Given the intra- and inter-individual variability and complexity of clinical manifestations, the diagnosis of migraine could be challenging for physicians without specialized expertise in headache medicine, thus underscoring the need for accessible biomarkers [10]. Although blood biomarkers such as calcitonin gene-related peptide (CGRP) have been investigated for the diagnosis of migraine [11, 12], they have not been widely implemented in clinical practice owing to limitations in feasibility and reliability [10, 13–15].

Neuroimaging studies have highlighted the potential role of imaging markers in the diagnosis of migraine. Structural and functional magnetic resonance imaging (MRI) studies have revealed topological alterations in the migraine brain [16–18]. Structural MRI studies have further identified cortical thinning in the temporal and occipital areas, as well as volume reductions in the hypothalamus [19–21]. Although these changes showed high accuracy in distinguishing chronic migraine from episodic migraine and healthy controls in one study [22], they have not been widely replicated or validated in classifying episodic migraine from healthy controls. Functional MRI studies have further demonstrated significant alterations in the functional connectivity of the visual, limbic, and default mode systems, particularly in those involving the insula, thalamus, and brainstem [23–26]. A recent study on migraine classification using multimodal imaging achieved 89% accuracy in distinguishing between patients with migraine and healthy controls [27]. However, recent large-scale studies have raised questions regarding the reliability of these findings and whether the proposed models could be generalized across various data configurations because most classification performances in prior studies were based on optimal settings [28, 29]. Given these findings, it is evident that more robust diagnostic models are required before neuroimaging can serve an independent role in migraine diagnosis. A promising avenue for future research is the development of models that integrate multimodal structural and functional MRI data with the validation of diagnostic performance across different migraine phases and times.

In this context, the present study aimed to develop a robust multimodal diagnostic model for migraine that demonstrated reliability and generalizability across different technical settings (multimodal feature combinations, feature selection hyperparameters, and spatial granularities). In addition, we validated the stability of the model across different migraine phases and longitudinally repeated data.

Methods

Study participants

We recruited 50 patients with episodic migraine and 50 age- and sex-matched healthy controls from an academic headache clinic between August 2017 and July 2018 [30]. The inclusion criteria for patients were as follows: 1) age 18–50 years, 2) not taking preventive medications, and 3) premenopausal status in female patients. The exclusion criteria were as follows: 1) chronic migraine, medication-overuse headaches, chronic pain disorders other than migraine, and psychiatric disorders such as bipolar affective disorder or schizophrenia; 2) contraindications for 3 T MRI, including use of a tissue expander, pacemaker, non-detachable metal objects, orthodontic devices, electrical leads, or implants in the body; 3) pregnancy; 4) claustrophobia requiring sedation during MR scanning; 5) inability to report headache or complete the headache diary due to cognitive decline; and 6) disagreement with the study procedures. Controls were recruited using the same inclusion and exclusion criteria after confirming that they had no history of memorable headache episodes in their lifetime or during the last year and no family history of migraine. Migraine was diagnosed by a headache specialist (MJL) according to the ICHD 3rd version [9]. The study was approved by the Samsung Medical Center Institutional Review Board, and all participants provided written informed consent.

Study scheme

At baseline, the patients completed headache diaries and clinical evaluations and underwent structural and functional MRIs following the protocol described below. Headache status within 2 days of MRI scanning was obtained. Patients were considered ictal if they had headaches of any intensity on the day of scanning, interictal if they were headache-free for 2 days before and after scanning, and peri-ictal if they were headache-free on the day of scanning, but developed headaches within two days of scanning. The same participants underwent multimodal

MRI scanning and clinical evaluation using the same protocol one year after the baseline scan. This study was registered at ClinicalTrials.gov (NCT03487978).

MRI acquisition

Participants underwent T1-weighted and resting-state functional MRI (rs-fMRI) using a Philips Ingenia 3 T scanner. The acquisition parameters for T1-weighted imaging were as follows: repetition time (TR) = 9.9 ms; echo time (TE) = 4.6 ms; flip angle = 8°; field of view (FOV) = 240 mm × 240 mm; the number of slices = 180 (reconstructed to 360); slice thickness = 1 mm (reconstructed to 0.5 mm); and in-plane resolution = 1 mm × 1 mm (reconstructed to 0.5 mm × 0.5 mm). The rs-fMRI data were acquired as follows: TR = 3,000 ms; TE = 30 ms; flip angle = 90°; FOV = 220 mm × 220 mm; number of slices = 33; slice thickness = 4 mm; in-plane resolution = 3 mm × 3 mm (reconstructed to 1.25 mm × 1.25 mm); and number of volumes = 200. During the scan, the participants were instructed to keep their eyes open and focus on a fixation cross to prevent falling asleep.

Data preprocessing

The imaging data were preprocessed using the fusion of a neuroimaging preprocessing (FuNP) surface-based pipeline incorporating FSL, AFNI, FreeSurfer, ANT, and Workbench (https://github.com/CAMIN-neuro/FuNP/) [31–35]. The gradient nonlinearity and B0 distortions were corrected from the T1-weighted images, non-brain tissues were removed, and intensity was normalized. Pial and white surfaces were generated, and the two surfaces were averaged to obtain a mid-thickness surface [33, 36–38]. For the rs-fMRI data, preprocessing involved the following steps: the first four volumes were discarded to allow for magnetic field saturation, and slice timing was corrected. Volumes with large head motion (i.e., framewise displacement > 0.5 mm) were discarded [39], and motion correction was performed. Skull removal and intensity normalization were performed, and nuisance variables, including head motion, white matter, cerebrospinal fluid signals, cardiac and respiration signals, and large vein-related contributions, were removed using the FMRIB's independent component analysis-based X-noiseifier (ICA-FIX) [40]. The rs-fMRI data were registered onto preprocessed T1-weighted structural data and subsequently to the Montreal Neurological Institute standard space. The band-pass filter with a frequency range between 0.009 and 0.08 Hz was applied. Finally, the preprocessed volumetric rs-fMRI data were mapped to the cortical surface using a cortical ribbon-constrained volume-to-surface mapping algorithm, and spatial smoothing with a full width at half maximum of 5 mm was applied.

Multimodal imaging features

We calculated the morphological features from preprocessed T1-weighted MRI and the functional features from rs-fMRI (Fig. 1).

Morphological features

The morphological features included cortical thickness, curvature, and sulcal depth. Cortical thickness assesses the thickness of gray matter by measuring the distance between the pial and white surfaces. Curvature was defined as the degree of folding of the cortical surface, while the sulcal depth was the vertical distance between the sulci and pial surfaces. Vertex-level data were mapped to brain regions defined using the Shaefer atlas with 200, 300, and 400 parcels [41] to assess the robustness of the model across spatial granularities.

Functional features

From the rs-fMRI data, a functional connectivity matrix was constructed by calculating the linear correlations of the time series between different brain regions defined by the Schaefer atlas containing 200, 300, and 400 parcels [41]. The correlation coefficients were calculated using Fisher's r-to-z transformations. The nodal degree and betweenness centrality were computed after leaving the top 10% of the elements per node. Degree centrality was defined as the sum of all edge weights connected to a given node, while betweenness centrality was defined as the number of shortest paths between any two nodes running through that node. Additionally, we generated a low-dimensional representation of functional connectivity (gradient) by applying a nonlinear dimensionality-reduction technique (diffusion map embedding) [42, 43]. The functional gradient represents the hierarchy of the brain because it shows an axis that distinguishes between the low-level sensory and higher-order default mode regions [42]. First, we generated a group-level template gradient from the averaged functional connectivity matrix and aligned the individual gradients to the template gradients using Procrustes rotation [43, 44].

Feature selection

Because the size of each imaging feature was equal to the number of brain regions (200, 300, or 400), the combination of N imaging features resulted in $N \times$ *number of brain regions*. When constructing machine learning models, too many features commonly result in overfitting problems. Thus, a feature selection process is required to identify a subset of features to build a classification model to distinguish patients with migraine from healthy controls. We utilized the elastic net regularization method, which combines the properties of the least absolute shrinkage and selection

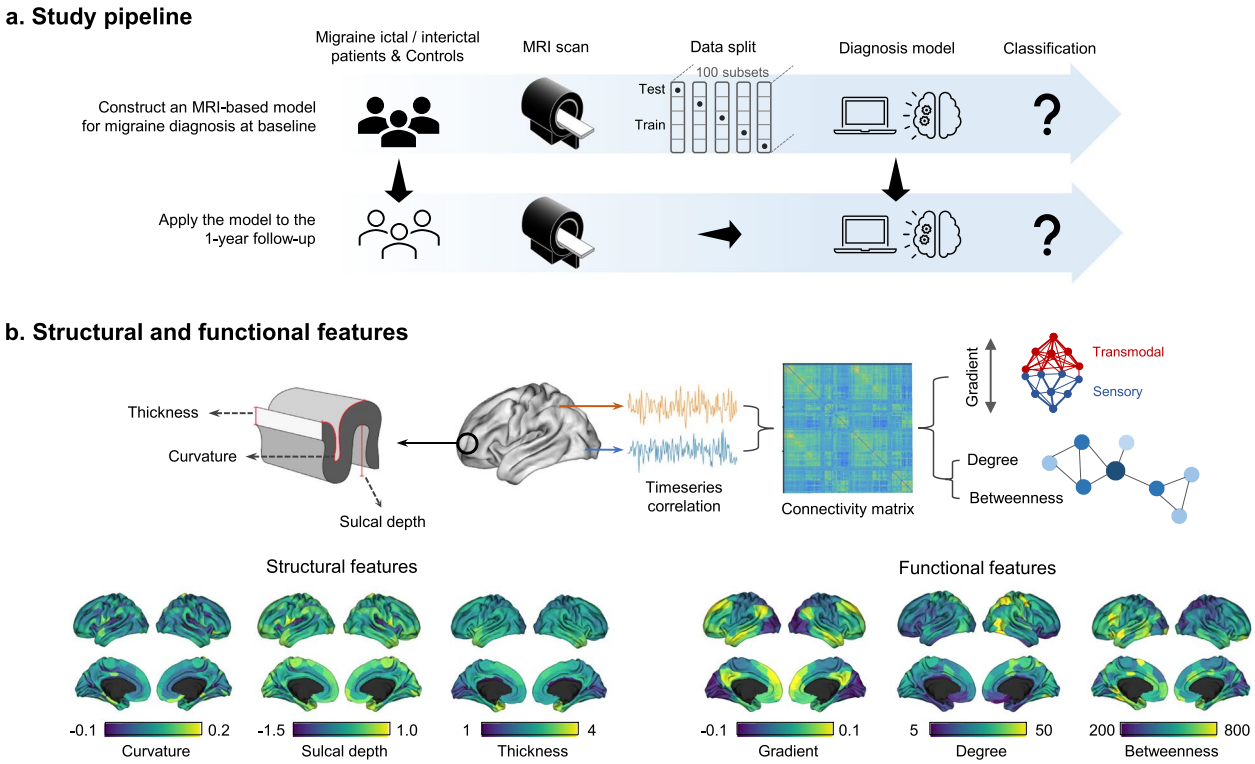

Fig. 1 The migraine diagnostic model constructed using multimodal MRI features. **a** Schema of the research pipeline to construct a multimodal MRI-based migraine diagnosis model using data at baseline and validating its performance using data at follow-up. **b** The concepts of the features, including curvature, sulcal depth, cortical thickness, functional gradient, degree centrality, and betweenness centrality, are shown. Mean values of the features at baseline are shown on brain surfaces

operator (LASSO) and ridge regression models [45]. It adjusts the weights of each feature by penalizing the coefficients of the model with the L1 (LASSO) and L2 (ridge) norms, as follows:

$$\left(\widehat{\beta_0}, \widehat{\boldsymbol{\beta}}\right) = \underset{\beta_0, \boldsymbol{\beta}}{argmin}\left(\sum_{i=1}^{n}(y_i - \beta_0 - \boldsymbol{\beta}^T \boldsymbol{x_i})^2 + \lambda \sum_{j=1}^{p}\left(\frac{1-\alpha}{2}\beta_j^2 + \alpha|\beta_j|\right)\right)$$

where $\boldsymbol{x_i}$ is the feature set of i^{th} subject and y_i is the dependent variable indicating the disease state. β_j denotes the regression coefficient of j^{th} feature, λ determines the overall strength of the regularization, while α determines the relative weight of the ridge and LASSO penalties (i.e., L1 ratio). We determined the optimal λ using greedy search algorithms by varying the L1 ratio as {0.5, 0.7, 0.9, 0.95, 0.99, 1}. A hundred λ values were generated at each L1 ratio, and the optimal λ showing the best performance was selected. The feature selection procedure was performed using the baseline data, and the selected features were applied to the follow-up data.

Diagnostic model

We constructed a multimodal MRI-based diagnostic model to distinguish patients with migraine from healthy controls using a random forest classifier with selected features. Given the constraints of small sample size and the high dimensionality of imaging features, the random forest classifier, an ensemble-based machine-learning approach that aggregates decisions from multiple decision trees, is an appropriate option as it effectively captures feature interactions in high-dimensional data [46]. It is more suitable than deep learning or boosting methods, which are prone to overfitting and unstable training with limited data. In this study, a hundred trees and the maximum depth at which all leaves become pure nodes were considered. The model was trained using all the subjects at baseline. We conducted a classification task with

a five-fold cross-validation. Specifically, 100 participants were randomly divided into five subsets of equal size (i.e., 20 participants per subset), ensuring that the proportions of patients and controls were balanced. Four of the five folds were used as training data, while the remaining fold served as the test set. We repeated this process 100 times with different training and test datasets to mitigate potential subject selection bias. To assess the stability of the trained model across different migraine phases, we applied the model constructed using all data to patients in the ictal/peri-ictal phase and matched controls, as well as patients in the interictal phase and matched controls. In addition, the model built using all participants at baseline was applied to the follow-up data. The classification performance was assessed based on accuracy, sensitivity, specificity, F1 score, and area under the curve (AUC). The model accuracy was assessed for all designed L1 ratios and parcellation scales. We further evaluated the classification performance by configuring the features in all possible combinations of morphological (i.e., cortical thickness, curvature, and sulcal depth) and functional measures (i.e., degree centrality, betweenness centrality, and gradient).

Between-group differences in the features

To evaluate how each feature differed between the groups, we quantified the selected probability of the feature set that showed the highest average classification accuracy across baseline and follow-up, interictal and ictal/peri-ictal phases, L1 ratios, and parcellation scales. We subsequently compared the features of the brain regions selected during the feature selection process between patients with migraine and healthy controls using two-sample t-tests. We visualized the selected probabilities and t-statistics according to seven intrinsic functional networks: visual, somatomotor, dorsal attention, ventral attention, limbic, frontoparietal, and default mode [47]. Additionally, we calculated the Shapley values from the random forest classifier across various parameter settings to identify the features contributing to the classification between patients with migraine and healthy controls.

Clinical associations

The clinical applicability of our findings was demonstrated by examining the relationship between the features with the highest classification performance (i.e., curvature, sulcal depth, thickness, and gradient) and clinical measures, including headache frequency and disease duration. Specifically, we predicted the clinical scores by applying a linear regression model to the top 20 features that showed high correlations with each clinical score.

Statistical analysis

Data are presented as the mean ± standard deviation (SD), unless otherwise specified. All analyses and visualizations were performed using Python v3.10.9. Elastic net regularization and random forest classifiers were conducted using scikit-learn v1.3.2, and a two-sample t-test was conducted using SciPy v1.10.0.

Results

The characteristics of the patients and controls are summarized in Tables 1 and 2. At baseline, we enrolled 50 patients (age 33.46 ± 9.01 years; 70% female) and 50 matched healthy controls (age 33.5 ± 8.95 years; 70% female) to undergo the clinical assessment and MRI scanning. After one year, 46 (92.0%) patients (age 35.63 ± 8.98 years; 72% female) and 43 (86.0%) healthy controls (age 35.02 ± 8.76 years; 72% female) completed the follow-up MRI scan and clinical evaluation. The headache phase was interictal in 16 and 21 patients at baseline and at follow-up, respectively, and ictal/peri-ictal in 34 and 25 patients, respectively.

Classification using unimodal data

Single feature

We evaluated classification performance using unimodal features from T1-weighted MRI or rs-fMRI. When a single feature was considered, cortical thickness and degree centrality were not selected during the feature selection process. The curvature showed the highest classification performance (accuracy [%] = 78.0 and 94.0 at baseline and follow-up, respectively). Although the order of classification accuracy changed at follow-up for other features, the baseline accuracy was modest for the functional gradient (73.4 and 56.8), followed by betweenness centrality (72.6 and 54.4) and sulcal depth (68.1 and 75.7; Fig. 2a). These results remained consistent when applied

Table 1 Demographic information of the study participants at baseline

	Patients ($N = 50$)	Controls ($N = 50$)	*P*-value
Age (years)	33.5 (9.01)	33.5 (8.95)	> 0.999
Sex (female)	35 (70.0%)	35 (70.0%)	> 0.999
Migraine with aura	9 (18.0%)	-	
Disease duration (years)	12.0 (7.45)		
Headache days per month	7.4 (5.22)	-	
Headache frequency at the month of fMRI acquisition	5.2 (3.55)		

Abbreviations: fMRI functional magnetic resonance imaging, *BMI* Body mass index

Table 2 Study participants at interictal and ictal/peri-ictal phases

Information		Interictal patients	Matched controls	Ictal/peri-ictal patients	Matched controls
Number	Baseline	16	16	34	34
	Follow-up	21	21	25	22
Sex (female)	Baseline	13 (81.3%)	13 (81.3%)	22 (64.7%)	22 (64.7%)
	Follow-up	11 (52.4%)	11 (52.4%)	22 (88.0%)	20 (90.9%)
Age (years)	Baseline	31.93 (9.10)	32.37 (8.65)	34.17 (8.88)	34.02 (9.03)
	Follow-up	35.52 (9.58)	36.19 (9.56)	35.72 (8.45)	33.9 (7.76)

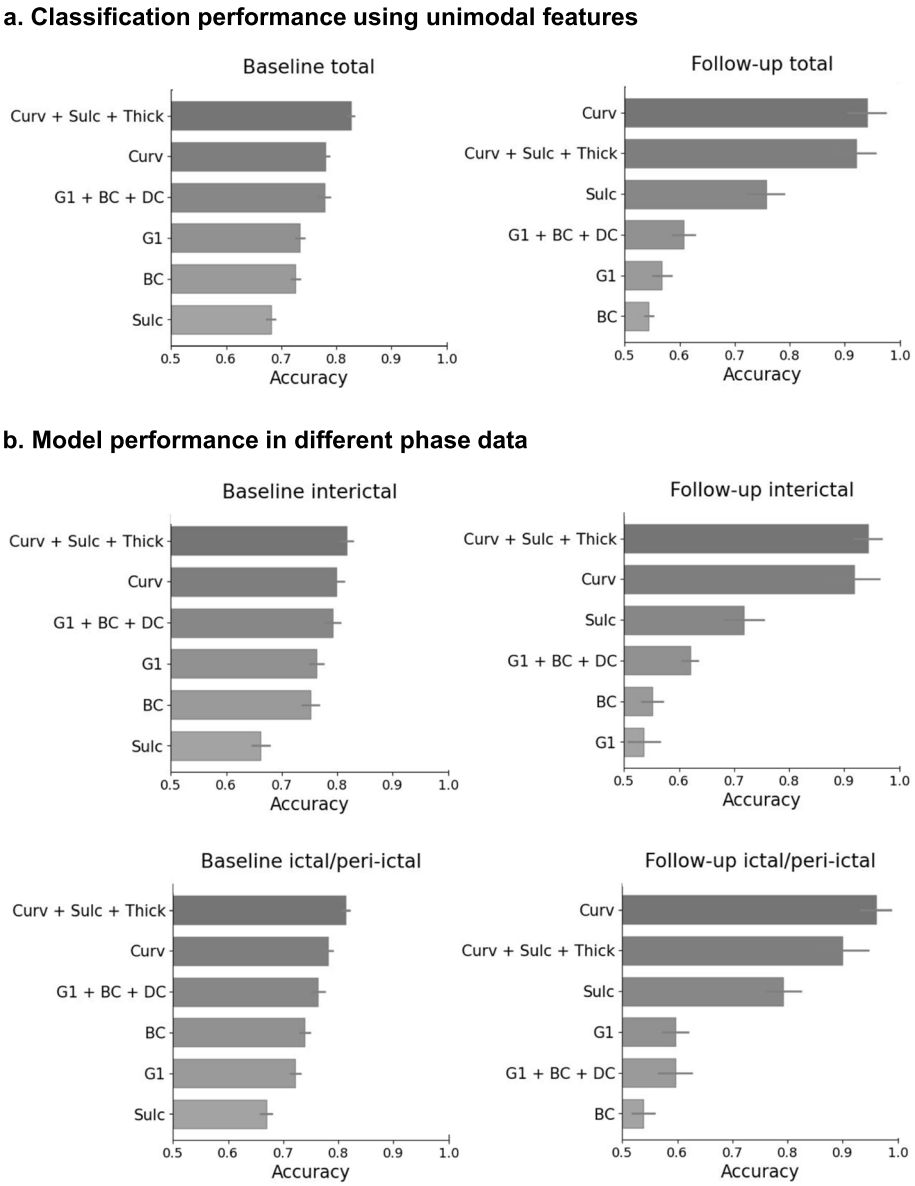

Fig. 2 Classification performance using single modal features. **a** We constructed classification models using a single or multiple features within unimodal imaging data. The bar plots present the classification accuracy, and the error bars indicate 99% confidence interval across the L1 ratios and parcellations. **b** The classification model was evaluated using the interictal and ictal/peri-ictal datasets. *Abbreviations:* Curv, curvature; Sulc, sulcal depth; Thick, cortical thickness; G1, first gradient; BC, betweenness centrality; DC, degree centrality

separately to the interictal and ictal/peri-ictal datasets. For the interictal patient-control dataset, the classification accuracies were 78.2 and 94.7 for curvature, 76.3 and 53.6 for functional gradient, 75.2 and 55.2 for betweenness centrality, and 66.2 and 71.7 for sulcal depth at baseline and at follow-up, respectively. For the ictal and peri-ictal patient-control datasets, we observed 77.9 and 93.2 for curvature, 72.2 and 59.7 for functional gradient, 73.9 and 53.8 for betweenness centrality, and 66.9 and 79.2 for sulcal depth, respectively (Fig. 2b).

Multiple features

As no single feature showed excellent performance, we combined the morphological features (i.e., curvature, sulcal depth, and cortical thickness) from T1-weighted MRI. This yielded performances of 82.6 and 90.0 in the total dataset, 81.7 and 94.3 in the interictal patient-control dataset, and 81.3 and 90.0 in the ictal/peri-ictal patient-control dataset, respectively, at baseline and follow-up (Fig. 2), outperforming the combined functional features (gradient, betweenness centrality, and degree centrality) from the rs-fMRI (77.8 and 60.7 in the total dataset, 79.2 and 62.0 in the interictal dataset, and 76.3 and 59.6 in the ictal/peri-ictal dataset; Fig. 2). The accuracy, sensitivity, specificity, F1 score, and AUC for all cases are presented in Supplementary Tables 1–3.

Classification using multimodal data

We constructed classifiers using all possible combinations of multimodal features. The combination of curvature, sulcal depth, cortical thickness, functional gradient, betweenness centrality, and degree centrality showed the highest classification performance (accuracy [%] = 90.4 and 86.0 at baseline and follow-up, respectively; Fig. 3a), which was retained in different phases (90.4 and 87.2 in the interictal patient-control dataset; and 89.6 and 84.9 in the ictal/peri-ictal patient-control dataset; Fig. 3b). Although the order of classification accuracy changed slightly at follow-up, the curvature, sulcal depth, cortical thickness, and functional gradient consistently contributed to distinguishing the groups. The performance of all feature combinations is presented in Supplementary Table 4.

Between-group differences using the most differentiable feature set

Considering the classification accuracy of the baseline and follow-up data, we identified the most robust feature sets for migraine diagnosis. The feature combination of curvature, sulcal depth, cortical thickness, and functional gradient achieved the best accuracy of 86.7% with robustness (99% confidence interval: 0.865–0.869; Fig. 4a). Confusion matrices and receiver operating characteristic (ROC) curves are shown in Supplementary Fig. 1. We subsequently focused on these four features for further analysis. The highest selection probability was observed for curvature (mean ± standard deviation across L1 ratios and parcellations = 5.9 ± 0.98%), followed by functional gradient (5.6 ± 0.82%), sulcal depth (4.3 ± 0.67%), and cortical thickness (3.5 ± 0.68%; Fig. 4b). Among the different brain networks, dorsal attention, limbic, and frontoparietal networks were relatively frequently selected (6.3 ± 0.86, 6.0 ± 0.66, 6.0 ± 0.80%, respectively), and the somatomotor and visual networks showed moderate selection probability (4.8 ± 0.63, 4.3 ± 0.47%, respectively). Finally, we assessed the differences in features between patients with migraine and healthy controls and observed decreased statistics in patients with migraine in the somatomotor and frontoparietal networks (Fig. 4c). Between-group differences for each feature are presented in Supplementary Fig. 2. When we calculated the Shapley values, both structural and functional features indicated that the dorsal attention and limbic networks are important for the classification task (Fig. 4c). These areas are involved in sensory integration, attention regulation, and emotional processing, which are often disrupted in patients with migraine. When calculating the Shapley value for each feature, we found that the curvature of the somatomotor network, the sulcal depth of the dorsal attention network, and the functional gradient of the limbic and dorsal attention networks were particularly notable (Supplementary Fig. 3).

Clinical applicability

When we associated the features with clinical scores, the dorsolateral and medial prefrontal cortices, superior and inferior temporal lobes, precuneus, and visual areas showed significant associations for headache frequency (adjusted $R^2 = 0.34 \pm 0.10$ and $p = 0.03$ across parcellations and L1 ratios) and disease duration (adjusted $R^2 = 0.41 \pm 0.05$ and $p = 0.01$; Fig. 5).

Discussion

The development of reliable and robust diagnostic models for migraine is essential to overcome the complexities arising from intra- and inter-individual variability in patient pathophysiology. In the present study, we integrated multimodal MRI features to develop a diagnostic model using regularization-based feature selection and random forest classification techniques. Curvature, sulcal depth, cortical thickness, and functional gradient delivered the best performance and stability across various settings, including different time points, migraine phases, feature selection parameters, and parcellation scales. Notably, patients with migraine exhibit reduced morphological properties and functional connectivity within the somatomotor and frontoparietal regions, and these areas showed significant associations with clinical variables.

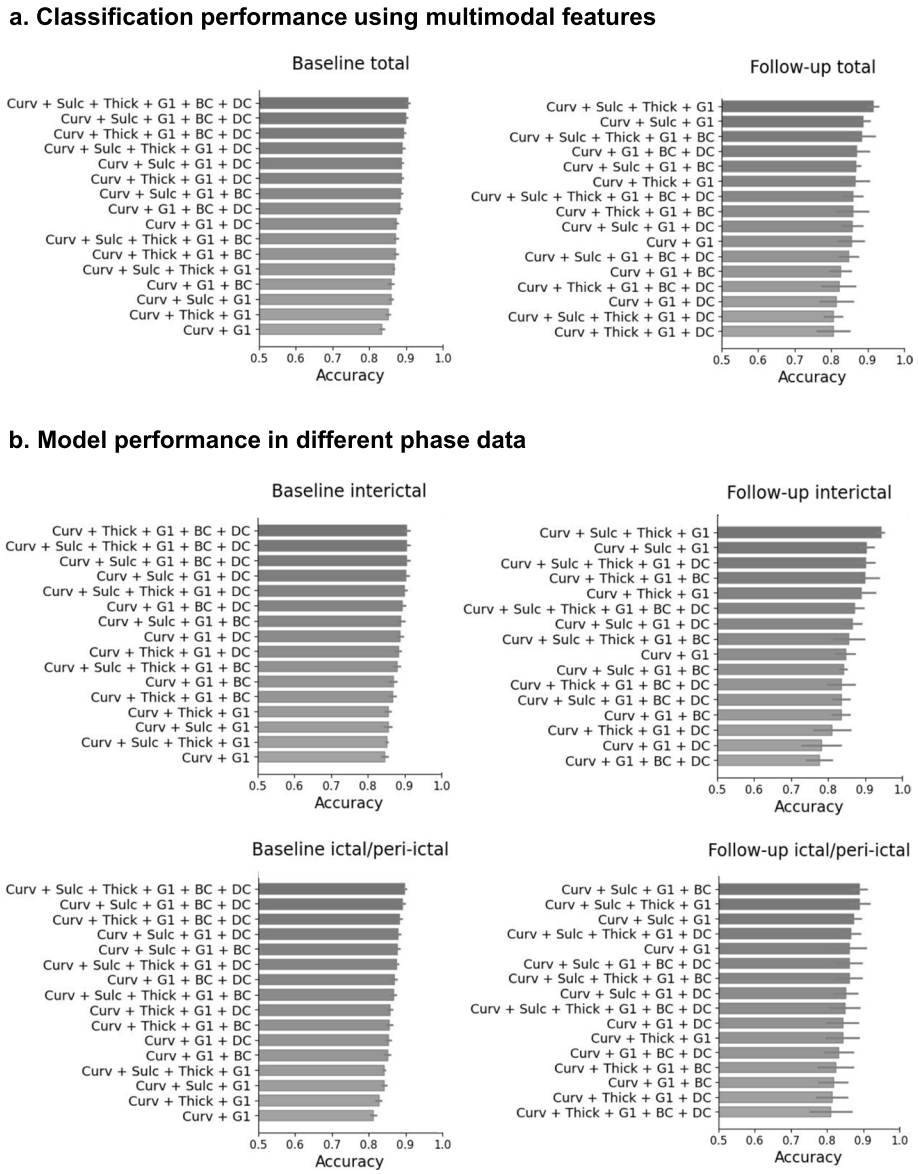

Fig. 3 Classification performance using multimodal features. **a** We constructed classification models using all feature combinations from T1-weighted and rs-fMRI. The bar plots present the classification accuracy, and the error bars indicate 99% confidence interval across L1 ratios and parcellations. **b** The classification model was evaluated using the interictal and ictal/peri-ictal datasets. *Abbreviations:* Curv, curvature; Sulc, sulcal depth; Thick, cortical thickness; G1, first gradient; BC, betweenness centrality; DC, degree centrality

Several classification models for migraine have been proposed, with accuracy rates ranging from 65 to 85% [22, 48–52]. Although several prior studies have validated model accuracies using external datasets [53], most have been conducted using cross-sectional data, while the reliability of these models has not been tested in a longitudinal setting with the same participants. Furthermore, the robustness and consistency of these models across different phases of headache have not been thoroughly addressed. A recent study introduced a machine learning-based approach to classifying individuals with migraine with aura from healthy controls using morphometric measures, employing extremely randomized trees for feature selection and linear discriminant analysis for classification [54]. The study identified key features such as the cortical thickness of the temporal pole, lingual gyrus, and pars opercularis, which are related to migraine pathophysiology, including sensory hypersensitivity, aura phenomena, and visual processing disruptions. Our findings were consistent with this study, but also identified

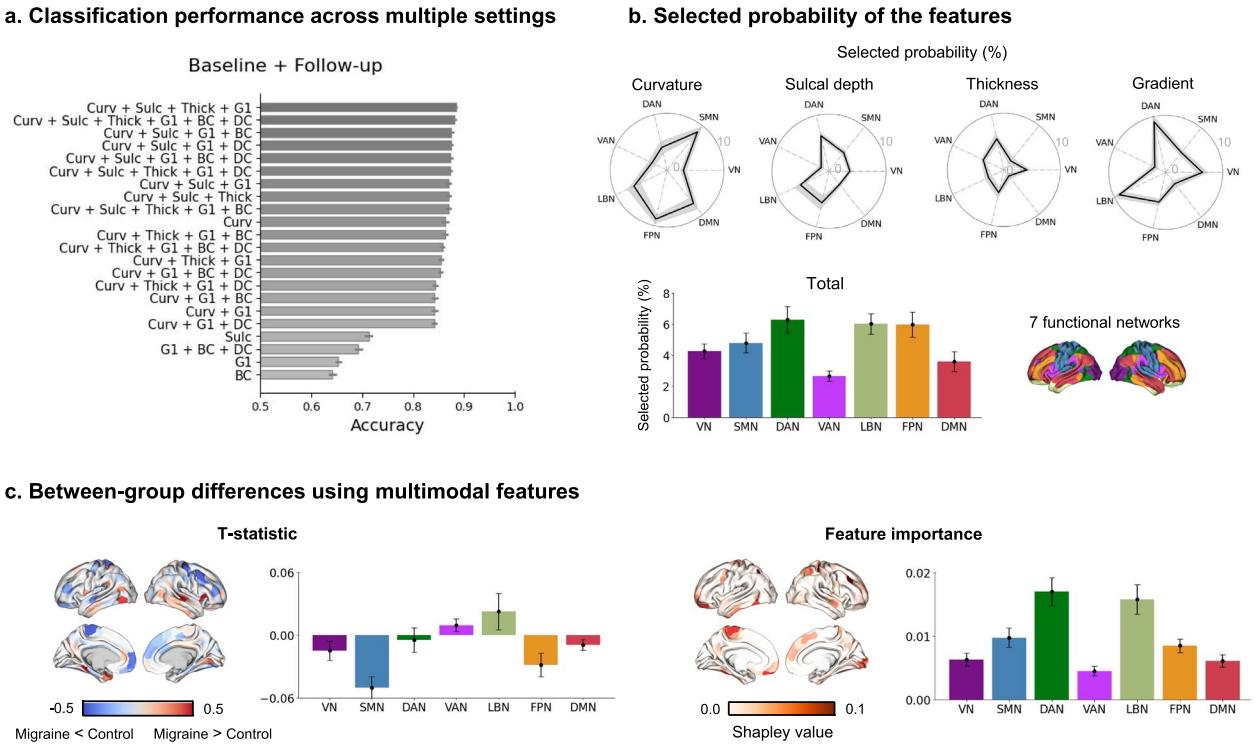

Fig. 4 Differences in multimodal features between patients with migraine and healthy controls. **a** The bar plots present the mean classification accuracy across baseline and follow-up, interictal and ictal/peri-ictal phases, feature selection parameters, and parcellations. The error bars indicate the 99% confidence interval. **b** The spider plots present the selection probability of brain regions per network for each feature. The bar plot represents the mean selection probability across all features. Gray areas in the spider plots and error bars in the bar plot indicate the standard error. **c** Brain surfaces showing the t-statistics of the between-group differences (left). The bar plot indicates the network-level t-statistics, and the error bars indicate the standard error. Negative values represent decreases in features in patients with migraine. The classification contributions between patients with migraine and healthy controls are assessed using the Shapley values and are visualized using the brain networks (right). Abbreviations: Curv, curvature; Sulc, sulcal depth; Thick, cortical thickness; G1, first gradient; BC, betweenness centrality; DC, degree centrality; VN, visual network; SMN, somatomotor network; DAN, dorsal attention network; VAN, ventral attention network; LBN, limbic network; FPN, frontoparietal network; DMN, default mode network

regions associated with pain modulation and sensory-related cognitive control processing, including the somatomotor, dorsolateral prefrontal, and medial prefrontal cortices. Compared to this prior work, our study offered a more comprehensive representation of migraine pathophysiology by integrating morphological and functional MRI features, enabling the identification of additional key regions for migraine classification. Critically, both studies support the notion that migraine is not localized to a single brain region but involves widespread brain network alterations, and sensory processing areas particularly reinforce their role as robust markers for migraine diagnosis. However, differences in the primary focus—subtype classification versus phase-general classification—highlight the complementary strengths of each approach. Furthermore, our study used more expansive and adaptive feature subsets through elastic net regularization and utilized an ensemble-based classifier, which is robust in handling high-dimensional data and capturing nonlinear relationships. These distinctions underline the broader applicability and robustness of our approach for diagnostic modeling in diverse migraine presentations.

Unlike single-modal designs, multimodal neuroimaging offers complementary insights into brain organization by capturing diverse data types [55], resulting in improved classification performance. In our study, the multimodal feature-based diagnostic model achieved 86.7% classification accuracy and demonstrated robustness across various spatial granularities, feature subsets, and migraine phases. Unlike prior studies, we used regularization techniques to avoid overfitting in the classification models with large feature sets. The key advantages of this approach are that it automatically selects features without manual intervention and is computationally efficient due to the low variability in the model parameters, contributing to the development of more reliable disease diagnosis models.

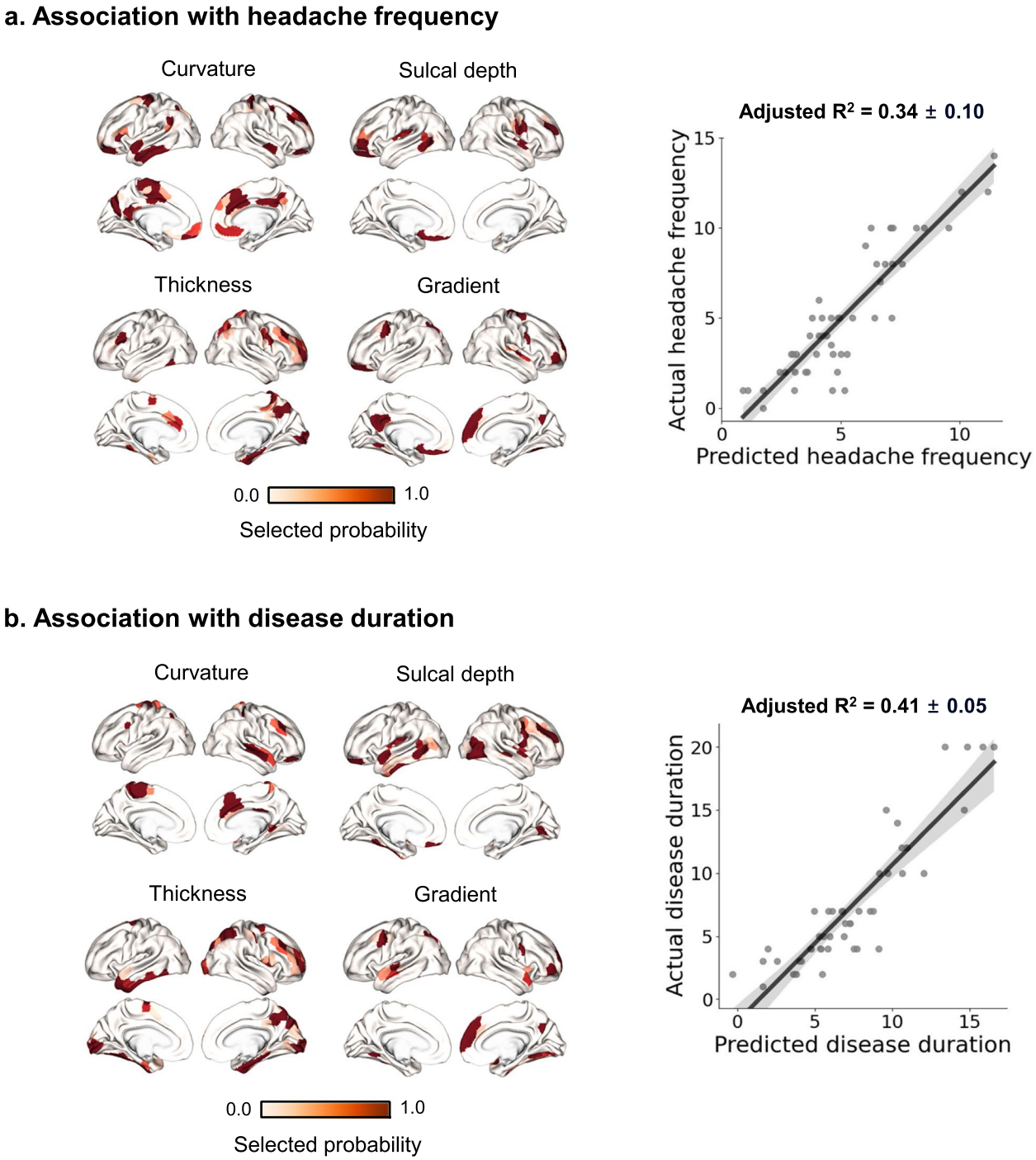

Fig. 5 Clinical score associations. **a** The selected probabilities of each feature that showed high correlations with headache frequency across L1 ratios and parcellations are visualized on brain surfaces (left). The scatter plot shows the correlation between the actual and predicted headache frequency (right). The gray area indicates the 95% confidence interval. **b** Results based on disease duration

In our analysis, both the morphological and functional features were identified as key markers for distinguishing patients with migraine from healthy controls. Prior studies have shown that disruptions in cortical morphology, such as reduced cortical thickness and altered curvature, can affect the brain folding landscape [56], potentially triggering a cascade of physiological events that lead to migraine symptoms [57]. Our findings of morphological alterations in the somatomotor and frontoparietal regions may reflect changes in the neuronal distribution

within these areas in patients with migraine. In addition to morphological features, functional features also play a critical role in the characterization of migraine [16, 24, 51, 58]. In the present study, we found that the functional connectivity gradient can serve as a robust marker for migraine diagnosis. The functional gradient is a reliable feature for assessing the brain hierarchy [42], while alterations in the gradient suggest disruptions in information processing along the hierarchical structure. Specifically, we observed decreased gradient values in the visual and somatomotor networks and increased values in the limbic network, indicating a more segregated network organization in patients with migraine. These changes may be associated with visual auras and the affective processing of pain, which is consistent with previous findings that noted disruptions in the sensory-limbic system in migraine [26]. When we associated the features with clinical scores, default mode and visual networks showed significant relationships. The findings are aligned with previous studies that reported associations between headache duration/frequency and abnormalities in volume and functional connectivity in regions, such as the anterior cingulate cortex, temporal lobe, frontal lobe, and limbic systems [59–62]. These results suggest that morphological and functional alterations may reflect the burden of recurrent headaches, alterations in sensory integration, and pain modulation in migraine. Taken together, these results indicate that the identified multimodal markers not only enhance diagnostic accuracy but also provide important insights into the topological disruptions underlying migraine.

Despite these promising results, several limitations of this study must be acknowledged. Using *TTestIndPower* from the Python package, we calculated that a sample size of 64 would be required to achieve sufficient power (0.8) while maintaining the type I error rate at 0.05. Since our dataset included 100 participants (50 patients with migraine and 50 healthy controls), the sample may not fully represent the population. However, the primary aim of our study was not to assess the absolute statistical significance of between-group differences but rather to identify robust feature combinations for classification and to observe the brain regions that prominently influenced classification. Although we implemented a rigorous cross-validation procedure and evaluated classification performance across a range of hyperparameter settings and parcellation schemes, the relatively small sample size remains a limitation in demonstrating the model's generalizability. In future studies, we aim to acquire larger and more diverse independent cohorts to validate our model. Additionally, incorporating other imaging modalities, such as diffusion MRI or positron emission tomography, could further enhance the diagnostic performance.

In conclusion, our multimodal diagnostic models demonstrated a robust and reliable performance in the diagnosis of migraine. This study presents a practical approach for migraine diagnosis by combining multimodal MRI with machine learning. In particular, the identification of morphological and functional connectivity features as key diagnostic markers highlights the importance of leveraging multimodal MRI to capture the complex neural changes associated with migraine. These findings could contribute to the development of more precise diagnostic tools for migraine. Future research could further explore the underlying biology of migraine by incorporating microscale genetic or neurotransmitter data to provide deeper insights into its pathophysiology through multiscale analysis.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s10194-024-01946-5.

Supplementary Material 1.

Code availability

All data supporting the findings of this study will be made available upon request from the corresponding author, Mi Ji Lee. The data are not publicly available due to Institutional Review Board restrictions. The codes for MRI preprocessing are available at https://github.com/CAMIN-neuro/FuNP; codes for gradient generation are available at https://github.com/MICA-MNI/BrainSpace; and codes for graph feature calculation are available at https://github.com/brainlife/BCT.

Authors' contributions

J.Y.N. and B.P. designed the study, analyzed the data, and wrote the manuscript. M.J.L. designed the study, acquired and interpreted the data, and critically revised the manuscript to ensure important intellectual content. E.N. and Y.J. aided with the experiments and reviewed the manuscript. M.J.L. and B.P. are the corresponding authors of this study and are responsible for the integrity of the data analysis. All authors have reviewed and approved the final version of the manuscript for publication.

Funding

Mi Ji Lee was supported by the National Research Foundation of Korea (NRF-2017R1A2B2009086 and RS-2024–00357394) and Korea Medical Device Development Fund grant funded by the Korea government (the Ministry of Science and Industry and Energy, the Ministry of Health and Welfare, the Ministry of Food and Drug Safety (1711198632 and RS-2023–00229484). Bo-yong Park received funding from the Institute for Information and Communications Technology Planning and Evaluation (IITP) funded by the Korea Government (MSIT) (No.2022–0-00448/RS-2022-II220448, Deep Total Recall: Continual Learning for Human-Like Recall of Artificial Neural Networks; RS-2021-II212068, Artificial Intelligence Innovation Hub) and the Institute for Basic Science (IBS-R015-D2).

Data availability

All data supporting the findings of this study will be made available upon request from the corresponding author, Mi Ji Lee. The data are not publicly available due to Institutional Review Board restrictions.

Declarations

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Data Science, Inha University, Incheon, Republic of Korea. [2]College of Medicine, Inha University, Incheon, Republic of Korea. [3]Department of Statistics and Data Science, Inha University, Incheon, Republic of Korea. [4]Department of Neurology, Seoul National University Hospital, Seoul, Republic of Korea. [5]Department of Neurology, Seoul National University College of Medicine, Seoul, Republic of Korea. [6]Department of Brain and Cognitive Engineering, Korea University, Seoul, Republic of Korea. [7]Center for Neuroscience Imaging Research, Institute for Basic Science, Suwon, Republic of Korea.

Received: 6 November 2024 Accepted: 30 December 2024
Published online: 09 January 2025

References
1. Silberstein SD (1995) Migraine Symptoms: Results of a Survey of Self-Reported Migraineurs. Headache. 35:387–396. https://doi.org/10.1111/J.1526-4610.1995.HED3507387.X
2. Kelman L (2006) The Postdrome of the Acute Migraine Attack. Cephalalgia 26:214–220. https://doi.org/10.1111/j.1468-2982.2005.01026.x
3. Schwedt TJ, Krauss MJ, Frey K, Gereau RW (2010) Episodic and chronic migraineurs are hypersensitive to thermal stimuli between migraine attacks. Cephalalgia 31:6–12. https://doi.org/10.1177/0333102410365108
4. Cucchiara B, Datta R, Aguirre GK et al (2014) Measurement of visual sensitivity in migraine: Validation of two scales and correlation with visual cortex activation. Cephalalgia 35:585–592. https://doi.org/10.1177/0333102414547782
5. Chong CD, Starling AJ, Schwedt TJ (2015) Interictal photosensitivity associates with altered brain structure in patients with episodic migraine. Cephalalgia 36:526–533. https://doi.org/10.1177/0333102415606080
6. Ashkenazi A, Mushtaq A, Yang I, Oshinsky ML (2009) Ictal and Interictal Phonophobia in Migraine—A Quantitative Controlled Study. Cephalalgia 29:1042–1048. https://doi.org/10.1111/j.1468-2982.2008.01834.x
7. Aurora SK, Winner P, Freeman MC, et al (2011) OnabotulinumtoxinA for Treatment of Chronic Migraine: Pooled Analyses of the 56-Week PREEMPT Clinical Program. Headache. 51:1358–1373. https://doi.org/10.1111/j.1526-4610.2011.01990.x
8. Granziera C, DaSilva AFM, Snyder J et al (2006) Anatomical Alterations of the Visual Motion Processing Network in Migraine with and without Aura. PLoS Med 3:e402. https://doi.org/10.1371/JOURNAL.PMED.0030402
9. Olesen J, Bes A, Kunkel R, et al (2013) The International Classification of Headache Disorders, 3rd edition (beta version). Cephalalgia 33:629–808. https://doi.org/10.1177/0333102413485658
10. Ashina M, Terwindt GM, Al-Karagholi MA-M, et al (2021) Migraine: disease characterisation, biomarkers, and precision medicine. The Lancet 397:1496–1504. https://doi.org/10.1016/S0140-6736(20)32162-0
11. Ho TW, Edvinsson L, Goadsby PJ (2010) CGRP and its receptors provide new insights into migraine pathophysiology. Nat Rev Neurol. 2010 6:10 6:573–582. https://doi.org/10.1038/nrneurol.2010.127
12. Cernuda-Morollón E, Larrosa D, Ramón C et al (2013) Interictal increase of CGRP levels in peripheral blood as a biomarker for chronic migraine. Neurology 81:1191–1196. https://doi.org/10.1212/WNL.0B013E3182A6CB72/SUPPL_FILE/FIGURE_E_1.TIF
13. Kamm K (2022) CGRP and Migraine: What Have We Learned From Measuring CGRP in Migraine Patients So Far? Front Neurol 13:930383. https://doi.org/10.3389/FNEUR.2022.930383/BIBTEX
14. Lee MJ, Lee SY, Cho S et al (2018) Feasibility of serum CGRP measurement as a biomarker of chronic migraine: a critical reappraisal. J Headache Pain 19:1–8. https://doi.org/10.1186/S10194-018-0883-X/FIGURES/5
15. Labastida-Ramírez A, Caronna E, Gollion C et al (2023) Mode and site of action of therapies targeting CGRP signaling. J Headache Pain 24:1–14. https://doi.org/10.1186/S10194-023-01644-8/TABLES/4
16. Yuan K, Zhao L, Cheng P et al (2013) Altered Structure and Resting-State Functional Connectivity of the Basal Ganglia in Migraine Patients Without Aura. J Pain 14:836–844. https://doi.org/10.1016/J.JPAIN.2013.02.010
17. Guarnera A, Bottino F, Napolitano A et al (2021) Early alterations of cortical thickness and gyrification in migraine without aura: a retrospective MRI study in pediatric patients. J Headache Pain 22:79. https://doi.org/10.1186/s10194-021-01290-y
18. Amin FM, De Icco R, Al-Karagholi MA-M et al (2021) Investigation of cortical thickness and volume during spontaneous attacks of migraine without aura: a 3-Tesla MRI study. J Headache Pain 22:98. https://doi.org/10.1186/s10194-021-01312-9
19. Magon S, May A, Stankewitz A et al (2018) Cortical abnormalities in episodic migraine: A multi-center 3T MRI study. Cephalalgia 39:665–673. https://doi.org/10.1177/0333102418795163
20. Chen Z, Chen X, Liu M, et al (2019) Volume of Hypothalamus as a Diagnostic Biomarker of Chronic Migraine. Front Neurol 10:00606. https://doi.org/10.3389/fneur.2019.00606
21. Coppola G, Petolicchio B, Di Renzo A et al (2017) Cerebral gray matter volume in patients with chronic migraine: correlations with clinical features. J Headache Pain 18:115. https://doi.org/10.1186/s10194-017-0825-z
22. Schwedt TJ, Chong CD, Wu T, et al (2015) Accurate Classification of Chronic Migraine via Brain Magnetic Resonance Imaging. Headache 55:762–777. https://doi.org/10.1111/HEAD.12584
23. Lee MJ, Park BY, Cho S et al (2019) Dynamic functional connectivity of the migraine brain: A resting-state functional magnetic resonance imaging study. Pain 160:2776–2786. https://doi.org/10.1097/J.PAIN.0000000000001676
24. Le WH, Zhou X, Chen YC et al (2019) Impaired intrinsic functional connectivity between the thalamus and visual cortex in migraine without aura. J Headache Pain 20:1–9. https://doi.org/10.1186/S10194-019-1065-1/FIGURES/3
25. Ke J, Yu Y, Zhang X, et al (2020) Functional Alterations in the Posterior Insula and Cerebellum in Migraine Without Aura: A Resting-State MRI Study. Front Behav Neurosci 14:567588. https://doi.org/10.3389/fnbeh.2020.567588
26. Lee CH, Park H, Lee MJ, Park B, yong, (2023) Whole-brain functional gradients reveal cortical and subcortical alterations in patients with episodic migraine. Hum Brain Mapp 44:2224–2233. https://doi.org/10.1002/HBM.26204
27. Messina R, Sudre CH, Wei DY, et al (2023) Biomarkers of Migraine and Cluster Headache: Differences and Similarities. Ann Neurol 93:729–742. https://doi.org/10.1002/ana.26583
28. Schramm S, Börner C, Reichert M et al (2023) Functional magnetic resonance imaging in migraine: A systematic review. Cephalalgia 43:03331024221128278. https://doi.org/10.1177/03331024221128278
29. Hougaard A, Gaist D, Garde E et al (2023) Lack of reproducibility of resting-state functional MRI findings in migraine with aura. Cephalalgia 43:03331024231212574. https://doi.org/10.1177/03331024231212574
30. Lee MJ, Park BY, Cho S et al (2019) Increased connectivity of pain matrix in chronic migraine: A resting-state functional MRI study. J Headache Pain 20:1–10. https://doi.org/10.1186/S10194-019-0986-Z/FIGURES/5
31. Cox RW (1996) AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 29:162–173. https://doi.org/10.1006/cbmr.1996.0014
32. Jenkinson M, Beckmann CF, Behrens TEJ et al (2012) FSL. Neuroimage 62:782–790. https://doi.org/10.1016/J.NEUROIMAGE.2011.09.015
33. Fischl B (2012) FreeSurfer. Neuroimage 62:774–781. https://doi.org/10.1016/J.NEUROIMAGE.2012.01.021
34. Avants BB, Tustison NJ, Song G et al (2011) A reproducible evaluation of ANTs similarity metric performance in brain image registration. Neuroimage 54:2033–2044. https://doi.org/10.1016/J.NEUROIMAGE.2010.09.025
35. Marcus DS, Harwell J, Olsen T, et al (2011) Informatics and data mining tools and strategies for the human connectome project. Front Neuroinform 5:. https://doi.org/10.3389/FNINF.2011.00004
36. Dale AM, Fischl B, Sereno MI (1999) Cortical Surface-Based Analysis: I Segmentation and Surface Reconstruction. Neuroimage 9:179–194. https://doi.org/10.1006/NIMG.1998.0395
37. Fischl B, Sereno MI, Dale AM (1999) Cortical Surface-Based Analysis: II: Inflation, Flattening, and a Surface-Based Coordinate System. Neuroimage 9:195–207. https://doi.org/10.1006/NIMG.1998.0396

38. Fischl B, Sereno MI, Tootell RBH, Dale AM (1999) High-Resolution Intersubject Averaging and a Coordinate System for the Cortical Surface. Hum Brain Mapping 8:272–284. https://doi.org/10.1002/(SICI)1097-0193(1999)8:4
39. Power JD, Barnes KA, Snyder AZ et al (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59:2142–2154. https://doi.org/10.1016/J.NEUROIMAGE.2011.10.018
40. Salimi-Khorshidi G, Douaud G, Beckmann CF et al (2014) Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. Neuroimage 90:449–468. https://doi.org/10.1016/J.NEUROIMAGE.2013.11.046
41. Schaefer A, Kong R, Gordon EM et al (2018) Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. Cereb Cortex 28:3095–3114. https://doi.org/10.1093/cercor/bhx179
42. Margulies DS, Ghosh SS, Goulas A et al (2016) Situating the default-mode network along a principal gradient of macroscale cortical organization. Proc Natl Acad Sci U S A 113:12574–12579. https://doi.org/10.1073/PNAS.1608282113/SUPPL_FILE/PNAS.201608282SI.PDF
43. Vos de Wael R, Benkarim O, Paquola C, et al (2020) BrainSpace: a toolbox for the analysis of macroscale gradients in neuroimaging and connectomics datasets. Commun Biol 2020 3:1 3:1–10. https://doi.org/10.1038/s42003-020-0794-7
44. Langs G, Golland P, Ghosh SS (2015) Predicting activation across individuals with resting-state functional connectivity based multi-atlas label fusion. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 9350:313–320. https://doi.org/10.1007/978-3-319-24571-3_38/COVER
45. Zou H, Hastie T (2005) Regularization and variable selection via the elastic net. J R Statist Soc B 67:301–320
46. Breiman L (2001) Random forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324/METRICS
47. Thomas Yeo BT, Krienen FM, Sepulcre J et al (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 106:1125. https://doi.org/10.1152/JN.00338.2011
48. Krawczyk B, Simić D, Simić S, Woźniak M (2013) Automatic diagnosis of primary headaches by machine learning methods. Cent Eur J Med 8:157–165. https://doi.org/10.2478/S11536-012-0098-5/MACHINEREADABLECITATION/RIS
49. Kwon J, Lee H, Cho S, et al (2020) Machine learning-based automated classification of headache disorders using patient-reported questionnaires. Scientific Rep 2020 10:1 10:1–8. https://doi.org/10.1038/s41598-020-70992-1
50. Subasi A, Ahmed A, Aličković E, Rashik Hassan A (2019) Effect of photic stimulation for migraine detection using random forest and discrete wavelet transform. Biomed Signal Process Control 49:231–239. https://doi.org/10.1016/j.bspc.2018.12.011
51. Wang Q, Gao Y, Zhang Y et al (2023) Decreased degree centrality values as a potential neuroimaging biomarker for migraine: A resting-state functional magnetic resonance imaging study and support vector machine analysis. Front Neurol 13:1105592. https://doi.org/10.3389/FNEUR.2022.1105592/BIBTEX
52. Rahman Siddiquee MM, Shah J, Chong C, et al (2022) Headache classification and automatic biomarker extraction from structural MRIs using deep learning. Brain Commun 5:. https://doi.org/10.1093/BRAINCOMMS/FCAC311
53. Duan S, Xia H, Zheng T et al (2023) Development and validation of non-invasive prediction models for migraine in Chinese adults. J Headache Pain 24:148. https://doi.org/10.1186/s10194-023-01675-1
54. Mitrović K, Petrušić I, Radojičić A, et al (2023) Migraine with aura detection and subtype classification using machine learning algorithms and morphometric magnetic resonance imaging data. Front Neurol 14:1106612. https://doi.org/10.3389/fneur.2023.1106612
55. Uludağ K, Roebroeck A (2014) General overview on the merits of multimodal neuroimaging data fusion. Neuroimage 102:3–10. https://doi.org/10.1016/j.neuroimage.2014.05.018
56. Messina R, Rocca MA, Colombo B et al (2013) Cortical abnormalities in patients with migraine: A Surfacebased analysis. Radiology 268:170–180. https://doi.org/10.1148/RADIOL.13122004/-/DC1
57. Burstein R, Noseda R, Borsook D (2015) Migraine: Multiple Processes, Complex Pathophysiology. J Neurosci 35:6619–6629. https://doi.org/10.1523/JNEUROSCI.0373-15.2015
58. Rao Y, Liu W, Zhu Y, et al (2023) Altered functional brain network patterns in patients with migraine without aura after transcutaneous auricular vagus nerve stimulation. Scientific Reports 2023 13:1 13:1–12. https://doi.org/10.1038/s41598-023-36437-1
59. Bashir A, Lipton RB, Ashina S, Ashina M (2013) Migraine and structural changes in the brain. Neurology 81:1260–1268. https://doi.org/10.1212/WNL.0b013e3182a6cb32
60. Rocca MA, Ceccarelli A, Falini A et al (2006) Brain Gray Matter Changes in Migraine Patients With T2-Visible Lesions. Stroke 37:1765–1770. https://doi.org/10.1161/01.STR.0000226589.00599.4d
61. Jin C, Yuan K, Zhao L, et al (2013) Structural and functional abnormalities in migraine patients without aura. NMR Biomed 26:58–64. https://doi.org/10.1002/nbm.2819
62. Schmitz N, Admiraal-Behloul F, Arkink EB, et al (2008) Attack Frequency and Disease Duration as Indicators for Brain Damage in Migraine. Headache 48:1044–1055. https://doi.org/10.1111/j.1526-4610.2008.01133.x

Publisher's Note