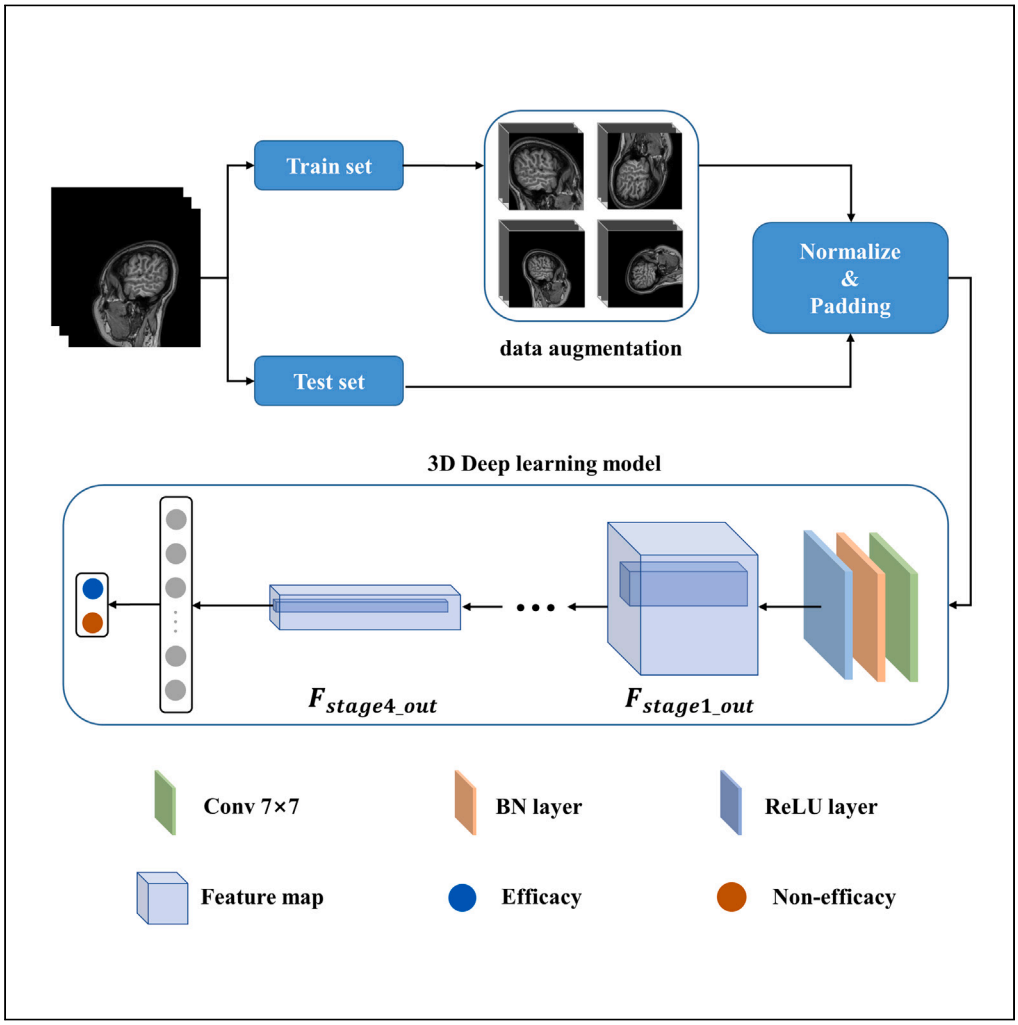
iScience

CellPress
OPEN ACCESS

Article

Predicting the efficacy of non-steroidal anti-inflammatory drugs in migraine using deep learning and three-dimensional T1-weighted images



Heng-Le Wei, Cunsheng Wei, Yibo Feng, ..., Xindao Yin, Junrong Li, Hong Zhang

ljry612@163.com (J.L.)
jnyyfsk@126.com (H.Z.)

Highlights

We developed a 3D deep learning model to predict NSAIDs efficacy for migraine

The 3D-ResNet18 model based on structural images outperforms conventional models

The deep learning model has the potential to be a recommender of treatment decision

Wei et al., iScience 26, 108107
November 17, 2023
https://doi.org/10.1016/j.isci.2023.108107

iScience

CellPress
OPEN ACCESS

Article

Predicting the efficacy of non-steroidal anti-inflammatory drugs in migraine using deep learning and three-dimensional T1-weighted images

Heng-Le Wei,[1,5] Cunsheng Wei,[2,5] Yibo Feng,[3] Wanying Yan,[3] Yu-Sheng Yu,[1] Yu-Chen Chen,[4] Xindao Yin,[4] Junrong Li,[2,*] and Hong Zhang[1,6,*]

SUMMARY

Deep learning (DL) models based on individual images could contribute to tailored therapies and personalized treatment strategies. We aimed to construct a DL model using individual 3D structural images for predicting the efficacy of non-steroidal anti-inflammatory drugs (NSAIDs) in migraine. A 3D convolutional neural network model was constructed, with ResNet18 as the classification backbone, to link structural images to predict the efficacy of NSAIDs. In total, 111 patients were included and allocated to the training and testing sets in a 4:1 ratio. The prediction accuracies of the ResNet34, ResNet50, ResNeXt50, DenseNet121, and 3D ResNet18 models were 0.65, 0.74, 0.65, 0.70, and 0.78, respectively. This model, based on individual 3D structural images, demonstrated better predictive performance in comparison to conventional models. Our study highlights the feasibility of the DL algorithm based on brain structural images and suggests that it can be applied to predict the efficacy of NSAIDs in migraine treatment.

INTRODUCTION

Migraine is a ubiquitous neurological disorder that is clinically characterized by recurrent, unilateral, pulsating headaches of moderate-to-severe intensity and a duration of 4–72 h.[1] Migraine is commonly accompanied by symptoms such as photophobia, phonophobia, and nausea/vomiting, which are aggravated by physical activity. According to the Global Burden of Disease Study 2019, migraine is the second leading cause of disability worldwide and the first leading cause of disability among young women.[2] The study has demonstrated that individuals with episodic migraine tend to chronicity with an annual progression rate of 3%.[3] Buse et al.[4] observed that the comorbidity rate increased as the frequency of headaches increased among people with migraine, incurring a considerable burden not only on the individuals but also on their families and society. Currently, non-steroidal anti-inflammatory drugs (NSAIDs) are recommended as first-line acute medications for patients with migraine.[5] Although these drugs may ameliorate disease symptoms, they do not cure the condition and are ineffective in a significant subset of patients, particularly in those with chronic migraine.[6] The most common risk factor for episodic migraine to progress to chronic migraine is medication overuse,[1] which can lead to severe multi-organ side effects, even life-threatening situations.[7] Further, limitations of effectiveness and optimization of migraine treatment with NSAIDs may prolong the disease course and even result in undesired side reactions.[8] Therefore, the development of a desired clinical approach that can effectively assess the efficacy of NSAIDs in migraine management is urgently warranted.

Increasing evidence on aberrant alterations suggests that marked functional and structural brain changes, central sensitization, and neuroinflammation are the mechanisms underlying migraine.[9,10] Although patients with primary headaches and no focal neurological signs do not undergo neuroimaging examination,[11] functional magnetic resonance imaging (fMRI) has become one of the most important techniques to noninvasively study human brain function and structure *in vivo*.[12,13] Increasing advances in fMRI technology have contributed to its application as a more powerful method to evaluate the correlation between subtle and spatially distributed signal patterns within the brain and clinical characteristics. Recently, several studies have applied machine learning methods to neuroimaging datasets to recognize and characterize neurological diseases.[14] Previous studies have suggested that neuroimaging data and clinical evidence could be of great value for classifying patients with migraine,[15] identifying migraine subtypes,[16] and predicting response to migraine treatment.[17] However, conventional machine learning models have exhibited insufficient performance and could not be able to satisfy clinical demands. Moreover, the characteristic

[1]Department of Radiology, The Affiliated Jiangning Hospital of Nanjing Medical University, Nanjing, Jiangsu 211100, China
[2]Department of Neurology, The Affiliated Jiangning Hospital of Nanjing Medical University, Nanjing, Jiangsu 211100, China
[3]Infervision Medical Technology Co., Ltd, Beijing, China
[4]Department of Radiology, Nanjing First Hospital, Nanjing Medical University, Jiangsu Province, Nanjing 210006, China
[5]These authors contributed equally
[6]Lead contact
*Correspondence: ljry612@163.com (J.L.), jnyyfsk@126.com (H.Z.)
https://doi.org/10.1016/j.isci.2023.108107

Table 1. Demographic data and questionnaire outcomes of patients with migraine

	Training set (n = 88)	Testing set (n = 23)	t/z/χ^2	p-value
Age (years)	31.50 (26.50, 40.50)	31.00 (24.00, 40.00)	−0.622	0.534
Sex (male/female)	19/69	2/21	1.977	0.160
Education (years)	12.00 (9.00, 16.00)	12.00 (9.00, 14.00)	−1.473	0.141
Aura (yes/no)	21/67	1/22	4.370	0.037
Family history (yes/no)	64/24	14/9	1.227	0.309
Photophobia (yes/no)	51/37	17/6	1.957	0.162
Phonophobia (yes/no)	45/43	10/13	0.428	0.513
Nausea/vomiting (yes/no)	57/31	12/11	1.231	0.267
Disease duration (years)	7.00 (3.00, 13.00)	4.00 (2.00, 12.50)	−1.117	0.264
VAS score	7.00 (5.50, 8.00)	7.00 (6.00, 8.00)	−0.482	0.630
Frequency (days/month)	4.00 (3.00, 6.00)	3.00 (2.50, 5.00)	−0.816	0.414
Attack time (hours)	12.00 (10.00, 24.00)	12.00 (10.00, 24.00)	−0.086	0.932
MIDAS score	21.00 (10.00, 46.50)	31.00 (23.00, 50.50)	−1.496	0.135
HIT-6 score	62.00 (51.50, 66.00)	57.00 (50.00, 68.00)	−0.310	0.757
GAD7 score	4.00 (3.00, 7.00)	4.00 (2.50, 8.50)	−0.139	0.889
PHQ-9 score	5.00 (3.00, 8.00)	6.00 (2.50, 10.50)	−0.179	0.858
PSQI score	7.50 (5.00, 11.00)	7.00 (6.00, 12.00)	−0.615	0.539
Pain medications			2.295	0.688
Ibuprofen	39	9		
Aspirin	19	4		
Acetaminophen	7	4		
Naproxen	14	3		
Celecoxib	9	3		

Continuous variables of normal distribution were presented as mean (standard deviation), whereas non-normal variables were presented as median (interquartile range). p-value <0.003 was considered significant after the Bonferroni correction. GAD-7, Generalized Anxiety Disorder 7-Item; HIT-6, Headache Impact Test 6-Item; MIDAS, Migraine Disability Assessment Scale; PHQ-9, Patient Health Questionnaire 9-Item; PSQI, Pittsburgh Sleep Quality Index; VAS, Visual Analogue Scale.

parameters based on the fMRI postprocessing can be influenced by differences in prior experience,[18,19] processing methods,[20,21] and application software.[22,23]

Deep learning (DL), a subset of machine learning, uses automatic complex multilayer neutral-network-architecture-based learning by converting input data into multiple kinds of abstractions.[24] DL has demonstrated promising performance in the field of medical image analysis.[25,26] In DL-based analysis of medical image patterns, convolutional neural network (CNN), one of the most common DL algorithms, can augment pattern recognition and characterization[27] and automatically learn how to extract valid features from the training samples for an assigned task by repetitive backpropagation adjustment of its weights without manual designation of the features as input information.[25] Owing to the ability of DL to detect abstract and complex patterns, it has been used in neuroimaging studies on psychiatric and neurological diseases with varying degrees of success.[26] Deep CNNs could complement MRI-based diagnosis by aggregating and processing large-scale information derived from neuroimaging data. Moreover, it is important to predict the efficacy of a specific therapy or personalized medicine for the future development of medical technologies. To date, few studies have utilized DL methods to predict the efficacy of NSAIDs in migraine treatment based on medical imaging analysis.[28,29] In addition, brain structure may be the neural substrate of brain functional changes, which precede changes in brain structure.[30] The extent of structural changes is comparatively lesser in magnitude when compared with functional changes, thereby suggesting that structural images possess a higher level of stability in accurately representing neuroimaging information of the brain.[31] Therefore, the present study aimed to develop and test a DL algorithm for predicting the efficacy of NSAIDs in migraine treatment using the baseline structural images.

RESULTS

Demographic characteristics and questionnaire outcomes

In total, 111 patients with migraine (62 responders and 49 non-responders to NSAIDs) were enrolled in this study. The statistical power of post hoc calculated was 0.98 much more than 0.80, indicating that the sample size included in this study met the requirements. All patients were divided into the training (n = 88) and testing (n = 23) sets in a 4:1 ratio. Detailed information about all subjects is summarized in Table 1. A

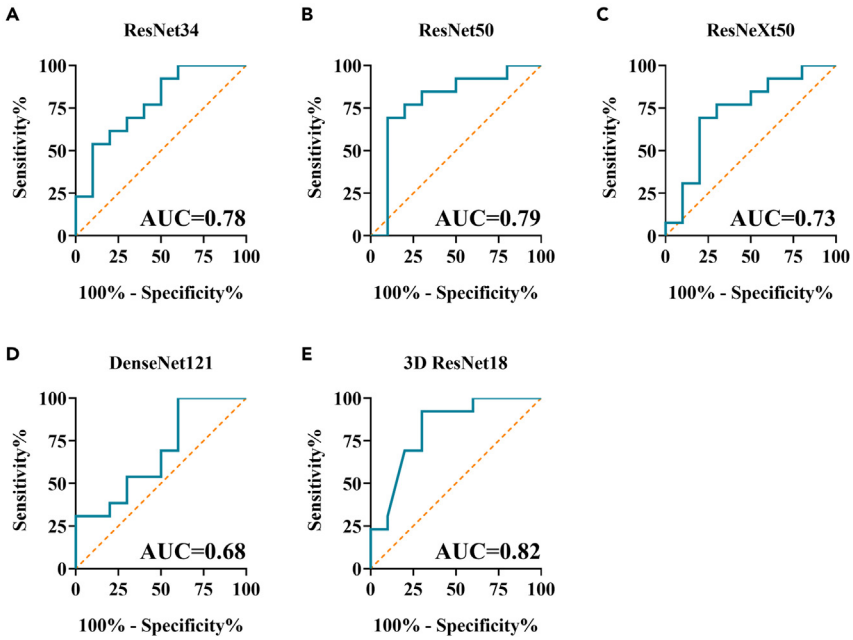


Figure 1. Receiver operating characteristic curves showing the performance of all the deep learning models for predicting NSAIDs efficacy in the testing set

(A–E) The areas under the curves of the ResNet34, ResNet50, ResNeXt50, DenseNet121, and 3D ResNet18 models (A–E) were 0.78, 0.79, 0.73, 0.68, and 0.82, respectively. NSAIDs: non-steroidal anti-inflammatory drugs.

statistical difference in aura symptom was observed between the two sets ($p < 0.05$, uncorrected). However, there were no statistically significant differences between the two sets after the Bonferroni correction.

Prediction performance of the different DL models

As shown in the receiver operating characteristic (ROC) curves (Figure 1), the conventional DL models trained with the 3D-T1WI sequence (i.e., ResNet34, ResNet50, ResNeXt50, and DenseNet121 models) had area under the curve (AUC) of 0.78, 0.79, 0.73, and 0.68, respectively, for predicting the efficacy of NSAIDs; these values were less than the AUC of our proposed model (0.82).

The accuracy, recall, precision and F1-score, positive likelihood ratio (PLR), negative likelihood ratio (NLR), and cutoff values were 0.65, 0.68, 0.71, 0.65, 2.13, 0.47, and 0.56; 0.74, 0.76, 0.77, 0.74, 3.13, 0.32, and 0.12; 0.65, 0.67, 0.68, 0.65, 2.02, 0.49, and 0.48; 0.70, 0.66, 0.73, 0.65, 1.95, 0.51, and 0.31; and 0.78, 0.77, 0.78, 0.78, 3.41, 0.29, and 0.50 in ResNet34, ResNet50, ResNeXt50, DenseNet121, and 3D ResNet18 models, respectively (Table 2).

Subgroup analysis

Ibuprofen was the most commonly used NSAIDs in migraine management. Considering this condition, the subgroup analysis was performed based on the Ibuprofen group and non-Ibuprofen group. The results showed that the AUC, accuracy, recall, precision and F1-score, PLR, and NLR values were 0.68, 0.78, 0.80, 0.83, 0.78, 4.00, and 0.25; and 0.89, 0.79, 0.79, 0.77, 0.78, 3.74, and 0.27, respectively, in Ibuprofen group and non-Ibuprofen group.

Moreover, there was a significant difference ($p < 0.05$, uncorrected) in aura symptom between the training and testing sets. Subgroup analysis was only performed in the patients without aura, because the sample size of patients with aura was too small to further conduct a reliable analysis. The results showed that the AUC, accuracy, recall, precision, F1-score, PLR, and NLR values were 0.83, 0.77, 0.76, 0.77, 0.77, 3.39, and 0.30, respectively.

DISCUSSION

The proposed DL model using features extracted from the 3D-T1WI sequence could feasibly evaluate the efficacy of NSAIDs in migraine treatment. The results demonstrated that our model using the 3D CNN algorithm could achieve a higher identification accuracy and outperform other conventional DL models. The superior classification performance of our single-mode DL model highlights its potential applications in clinical settings to optimize treatment decisions and improve clinical outcomes.

At present, CNN is the main approach used for detection and prediction tasks. The core elements of CNN take advantage of image properties and include local connections, shared weights, pooling, and the use of deep layers. Based on convolutional operation, the CNN input

Table 2. Classification performance of different models in testing set

Models	AUC (95% CI)	Accuracy	Recall	Precision	F1-score	PLR	NLR	Cutoff
ResNet34	0.78 (0.58, 0.97)	0.65	0.68	0.71	0.65	2.13	0.47	0.56
ResNet50	0.79 (0.58, 1.00)	0.74	0.76	0.77	0.74	3.13	0.32	0.12
ResNeXt50	0.73 (0.51, 0.95)	0.65	0.67	0.68	0.65	2.02	0.49	0.48
DenseNet121	0.68 (0.45, 0.90)	0.70	0.66	0.73	0.65	1.95	0.51	0.31
3D ResNet18	0.82 (0.64, 1.00)	0.78	0.77	0.78	0.78	3.41	0.29	0.50

Numbers in parentheses are 95% confidence intervals. AUC, area under the curve; Cis, confidence intervals; NLR, negative likelihood ratio; PLR, positive likelihood ratio.

includes highly correlated macroscopical information and representation. Therefore, it is well appropriate for processing structural images using spatial information to improve model performance.[32] For example, a recent study demonstrated that CNN models based on neuroimaging structural data have substantially improved performance compared with standard ML models.[33] In the present study, DL was used to discriminate between NSAIDs responders and non-responders. This package of techniques is self-acting and identifies patterns from raw brain imaging data without any assumptions about the predefined structure or content of the data. Nonlinear relationships and interrelationships between predictors are captured in the data. ResNet34, ResNet50, ResNeXt50, and DenseNet121 are commonly used DL approaches, and many studies have verified their promising performance for disease prediction.[34,35] In the present study, we constructed a model based on a 3D deep residual network with better classification performance, compared with the abovementioned conventional DL models.

Previous studies have reported that opioid receptor blockers injected into the core brain region of descending pain pathway could effectively reduce the analgesic effects of NSAIDs and have further suggested that NSAIDs act on the descending pain pathway via the endogenous opioid system to produce an antinociception effect.[36–38] Neuroimaging information may yield important representations in relation to treatment processes and outcomes. Although limited in number, some studies based on MRI techniques have revealed that changes in activity or morphology of the trigeminovascular system in patients with migraine were affected by different treatment modalities.[39–41] A study on postoperative analgesia with ibuprofen also suggested that increased cerebral blood flow (CBF) in the descending pain pathway is the potential neuropathologic mechanism underlying the occurrence and development of pain and that decreased CBF in the descending pathway after drug administration may represent the reduced activity of brain regions, thereby producing analgesic effects.[40] Further, based on baseline neuroimaging representations, some studies have described that traditional machine learning assessment models could play a role in predicting the efficacy of acupuncture treatment in patients with migraine over several weeks of follow-up.[39,41] These models achieved good prediction performance; however, they could not be widely implemented in clinical settings because they analyzed many risk factors obtained from detailed and time-consuming clinical history and neuroimaging characteristics obtained from complex postprocessing imaging methods.

In addition, classic fMRI neuroimaging with machine learning methods requires considerable prior experience in designing feature extractors from several postprocessing methods and transforming raw data into suitable internal representations or feature vectors, on a need basis to help learn subsystems to detect and classify patterns. However, this methodology of identifying features may be influenced by prior experience and be sensitive to irrelevant factors. As a result, it may not be flexible enough to reveal high-level differences or predict complex brain-behavior relationships.[24] In contrast, DL approaches use minimally pre-engineered features and multiple processing layers to learn potential representations of data with multiple abstraction levels and are highly flexible. Our model provides a method to analyze 3D structural images of the whole brain readily obtained via MRI. The model could be applied in clinical settings to rapidly and accurately predict the efficacy of NSAIDs in migraine treatment. The high AUCs of the DL models based on abstract information from raw medical images were similar to those in previous DL studies, in which cartilage lesions were detected in knee MRI,[42] pain progression in patients with knee osteoarthritis was predicted using knee radiographs,[43] and cardiovascular risks were predicted using chest computed tomography.[44] These studies further emphasize the favorable application prospects of DL-based methods for evaluating medical images. Taken together, these findings suggest that the DL algorithm is effective in discovering intrinsic features from high-dimensional data in high-quality medical images.

In this study, we combined brain structural images and DL algorithms to construct prediction models for NSAIDs efficacy in migraine management and obtained relatively good predictive effects, in terms of accuracy. The structural images have a higher level of stability to reflect neuroimaging information,[31] and the robust neuroimaging information will greatly benefit clinical practice. Moreover, DL can extract optimal features from the raw data with better learning capabilities, compared with traditional machine learning models. Therefore, our findings suggest that understanding how the intrinsic high-dimensional features of structural images are related to drug efficacy can contribute to providing insights into pathophysiological mechanisms of migraine. However, DL models inevitably involve black box models, which remain relatively uninterpretable when compared with conventional medical statistical methods.[45] It is difficult to back-construct higher-order representations of abstraction to the original structural dimension resulting in problems of interpreting the results, although the proposed models have been demonstrated to be feasible. These impose substantial restrictions on the interpretation of clinical significance in DL methods. Although these issues remain unsolved, intrinsic representative features derived from raw medical images can be automatically learned via different procedures to improve interpretability.[32] Our results suggest that feature extraction of different brain regions in combination with DL may have great clinical potential to improve the efficacy evaluation of migraine and can help clinicians make more informed decisions about therapeutic approaches for individual migraine patients.

In conclusion, we constructed a DL model that associated structural images of an individual with response to NSAIDs, facilitating the prediction of individual clinical outcomes. Therefore, our study provides a potentially feasible method to improve the selection of clinical treatment strategies for migraine.

Limitations of the study

The present study has some limitations that need to be addressed. First, this study used a cross-sectional design with a small sample size. Therefore, multiple classification models and larger prospective validation studies are warranted to improve prediction performance. Second, only the structural images of patients with migraine were evaluated. Importantly, the fusion model that combined multi-sequence MRI showed significantly higher performance in multiple classification tasks than models using a single sequence.[46] This suggests that multi-sequence MRI parameters can better understand migraine characteristics than a single sequence and can improve the identification performance of the model. Third, this study aimed to explore the relationship between 3D-T1 images and efficacy prediction using DL networks. Future studies should focus on integrating clinical and multimodal neuroimaging characteristics using DL and explore its capacity in enhancing the prediction performance of migraine treatment. Finally, this is a cross-sectional study that cannot reflect the causal relationship between structural changes and treatment response. A prospective cohort study for validating the temporal association is needed to obtain more reliable conclusions in the future.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Patient cohorts
  - Questionnaires
  - Outcomes measurement
  - Acquisition of 3D structural images
  - Construction of the DL model
  - Setting of experimental parameters
  - Model evaluation
- QUANTIFICATION AND STATISTICAL ANALYSIS

ACKNOWLEDGMENTS

The authors would like to thank the participants involved in this study. This work was supported by the Science and Technology Development Project of Nanjing Medical University (NMUB20220162).

AUTHOR CONTRIBUTIONS

H.L.W., C.W., and X.Y. supported the conception and design of this project. Y.S.Y. and Y.C.C. acquired data. Y.F. and W.Y. analyzed the data and prepared both figures. J.L. and H.Z. contributed to data quality control. H.L.W. wrote the main manuscript text. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: May 11, 2023
Revised: July 19, 2023
Accepted: September 27, 2023
Published: September 30, 2023

REFERENCES

1. Headache Classification Committee of the International Headache Society (IHS) (2018). The International Classification of Headache Disorders, 3rd edition. Cephalalgia, 1–211. https://doi.org/10.1177/0333102417738202.
2. GBD 2019 Diseases and Injuries Collaborators; Injuries Collaborators (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: A systematic analysis for the Global Burden of Disease Study 2019. Lancet *396*, 1204–1222. https://doi.org/10.1016/S0140-6736(20)30925-9.
3. Scher, A.I., Stewart, W.F., Ricci, J.A., and Lipton, R.B. (2003). Factors associated with the onset and remission of chronic daily headache in a population-based study. Pain *106*, 81–89. https://doi.org/10.1016/s0304-3959(03)00293-8.
4. Buse, D.C., Reed, M.L., Fanning, K.M., Bostic, R., Dodick, D.W., Schwedt, T.J., Munjal, S., Singh, P., and Lipton, R.B. (2020). Comorbid and co-occurring conditions in migraine and associated risk of increasing headache pain intensity and headache frequency: Results of the migraine in America symptoms and treatment (MAST) study. J. Headache Pain *21*, 23. https://doi.org/10.1186/s10194-020-1084-y.
5. Ashina, M. (2020). Migraine. N. Engl. J. Med. *383*, 1866–1876. https://doi.org/10.1056/NEJMra1915327.
6. Cho, S.J., Song, T.J., and Chu, M.K. (2017). Treatment update of chronic migraine. Curr. Pain Headache Rep. *21*, 26. https://doi.org/10.1007/s11916-017-0628-6.
7. Bindu, S., Mazumder, S., and Bandyopadhyay, U. (2020). Non-steroidal anti-inflammatory drugs (NSAIDs) and organ damage: A current perspective. Biochem. Pharmacol. *180*, 114147. https://doi.org/10.1016/j.bcp.2020.114147.
8. Cabassi, A., Tedeschi, S., Perlini, S., Verzicco, I., Volpi, R., Gonzi, G., and Canale, S.D. (2020). Non-steroidal anti-inflammatory drug effects on renal and cardiovascular function: From physiology to clinical practice. Eur. J. Prev. Cardiol. *27*, 850–867. https://doi.org/10.1177/2047487319848105.
9. Tolner, E.A., Chen, S.P., and Eikermann-Haerter, K. (2019). Current understanding of cortical structure and function in migraine. Cephalalgia *39*, 1683–1699. https://doi.org/10.1177/0333102419840643.
10. Lim, M., Jassar, H., Kim, D.J., Nascimento, T.D., and DaSilva, A.F. (2021). Differential alteration of fMRI signal variability in the ascending trigeminal somatosensory and pain modulatory pathways in migraine. J. Headache Pain *22*, 4. https://doi.org/10.1186/s10194-020-01210-6.
11. Jang, Y.E., Cho, E.Y., Choi, H.Y., Kim, S.M., and Park, H.Y. (2019). Diagnostic Neuroimaging in Headache Patients: A Systematic Review and Meta-Analysis. Psychiat Invest *16*, 407–417. https://doi.org/10.30773/pi.2019.04.11.
12. Liégeois, F., and Elward, R. (2020). Functional magnetic resonance imaging. Handb. Clin. Neurol. *174*, 265–275. https://doi.org/10.1016/B978-0-444-64148-9.00019-3.
13. Chow, M.S.M., Wu, S.L., Webb, S.E., Gluskin, K., and Yew, D.T. (2017). Functional magnetic resonance imaging and the brain: A brief review. World J. Radiol. *9*, 5–9. https://doi.org/10.4329/wjr.v9.i1.5.
14. Khosla, M., Jamison, K., Ngo, G.H., Kuceyeski, A., and Sabuncu, M.R. (2019). Machine learning in resting-state fMRI analysis. Magn. Reson. Imaging *64*, 101–121. https://doi.org/10.1016/j.mri.2019.05.031.
15. Kwon, J., Lee, H., Cho, S., Chung, C.S., Lee, M.J., and Park, H. (2020). Machine learning-based automated classification of headache disorders using patient-reported questionnaires. Sci. Rep. *10*, 14062. https://doi.org/10.1038/s41598-020-70992-1.
16. Frid, A., Shor, M., Shifrin, A., Yarnitsky, D., and Granovsky, Y. (2020). A biomarker for discriminating between migraine with and without aura: Machine learning on functional connectivity on resting-state EEGs. Ann. Biomed. Eng. *48*, 403–412. https://doi.org/10.1007/s10439-019-02357-3.
17. Fu, C., Zhang, Y., Ye, Y., Hou, X., Wen, Z., Yan, Z., Luo, W., Feng, M., and Liu, B. (2022). Predicting response to tVNS in patients with migraine using functional MRI: A voxels-based machine learning analysis. Front. Neurosci. *16*, 937453. https://doi.org/10.3389/fnins.2022.937453.
18. Smith, S.M., Miller, K.L., Salimi-Khorshidi, G., Webster, M., Beckmann, C.F., Nichols, T.E., Ramsey, J.D., and Woolrich, M.W. (2011). Network modelling methods for FMRI. Neuroimage *54*, 875–891. https://doi.org/10.1016/j.neuroimage.2010.08.063.
19. Bassett, D.S., and Sporns, O. (2017). Network neuroscience. Nat. Neurosci. *20*, 353–364. https://doi.org/10.1038/nn.4502.
20. Pervaiz, U., Vidaurre, D., Woolrich, M.W., and Smith, S.M. (2020). Optimising network modelling methods for fMRI. Neuroimage *211*, 116604. https://doi.org/10.1016/j.neuroimage.2020.116604.
21. Smitha, K.A., Akhil Raja, K., Arun, K.M., Rajesh, P.G., Thomas, B., Kapilamoorthy, T.R., and Kesavadas, C. (2017). Resting state fMRI: A review on methods in resting state connectivity analysis and resting state networks. NeuroRadiol. J. *30*, 305–317. https://doi.org/10.1177/1971400917697342.
22. Song, X.W., Dong, Z.Y., Long, X.Y., Li, S.F., Zuo, X.N., Zhu, C.Z., He, Y., Yan, C.G., and Zang, Y.F. (2011). REST: A toolkit for resting-state functional magnetic resonance imaging data processing. PLoS One *6*, e25031. https://doi.org/10.1371/journal.pone.0025031.
23. Jia, X.Z., Wang, J., Sun, H.Y., Zhang, H., Liao, W., Wang, Z., Yan, C.G., Song, X.W., and Zang, Y.F. (2019). RESTplus: An improved toolkit for resting-state functional magnetic resonance imaging data processing. Sci. Bull. *64*, 953–954. https://doi.org/10.1016/j.scib.2019.05.008.
24. Lecun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature *521*, 436–444. https://doi.org/10.1038/nature14539.
25. Chan, H.P., Samala, R.K., Hadjiiski, L.M., and Zhou, C. (2020). Deep learning in medical image analysis. Adv. Exp. Med. Biol. *1213*, 3–21. https://doi.org/10.1007/978-3-030-33128-3_1.
26. Gautam, R., and Sharma, M. (2020). Prevalence and Diagnosis of Neurological Disorders Using Different Deep Learning Techniques: A Meta-Analysis. J. Med. Syst. *44*, 49. https://doi.org/10.1007/s10916-019-1519-7.
27. Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S., and Dean, J. (2019). A guide to deep learning in healthcare. Nat. Med. *25*, 24–29. https://doi.org/10.1038/s41591-018-0316-z.
28. Wei, H.L., Yang, W.J., Zhou, G.P., Chen, Y.C., Yu, Y.S., Yin, X., Li, J., and Zhang, H. (2022). Altered static functional network connectivity predicts the efficacy of non-steroidal anti-inflammatory drugs in migraineurs without aura. Front. Mol. Neurosci. *15*, 956797. https://doi.org/10.3389/fnmol.2022.956797.
29. Wei, H.L., Xu, C.H., Wang, J.J., Zhou, G.P., Guo, X., Chen, Y.C., Yu, Y.S., He, Z.Z., Yin, X., Li, J., and Zhang, H. (2022). Disrupted functional connectivity of the amygdala predicts the efficacy of non-steroidal anti-inflammatory drugs in migraineurs without aura. Front. Mol. Neurosci. *15*, 819507. https://doi.org/10.3389/fnmol.2022.819507.
30. Jimenez-Marin, A., Rivera, D., Boado, V., Diez, I., Labayen, F., Garrido, I., Ramos-Usuga, D., Benito-Sánchez, I., Rasero, J., Cabrera-Zubizarreta, A., et al. (2020). Brain connectivity and cognitive functioning in individuals six months after multiorgan failure. Neuroimage. Clin. *25*, 102137. https://doi.org/10.1016/j.nicl.2019.102137.
31. Hwang, G., Hermann, B., Nair, V.A., Conant, L.L., Dabbs, K., Mathis, J., Cook, C.J., Rivera-Bonet, C.N., Mohanty, R., Zhao, G., et al. (2020). Brain aging in temporal lobe epilepsy: Chronological, structural, and functional. Neuroimage. Clin. *25*, 102183. https://doi.org/10.1016/j.nicl.2020.102183.
32. Yan, W., Qu, G., Hu, W., Abrol, A., Cai, B., Qiao, C., Plis, S.M., Wang, Y.P., Sui, J., and Calhoun, V.D. (2022). Deep learning in neuroimaging: Promises and challenges. IEEE Signal Proc. Mag. *39*, 87–98. https://doi.org/10.1109/MSP.2021.3128348.
33. Abrol, A., Fu, Z., Salman, M., Silva, R., Du, Y., Plis, S., and Calhoun, V. (2021). Deep learning encodes robust discriminative neuroimaging representations to outperform standard machine learning. Nat. Commun. *12*, 353. https://doi.org/10.1038/s41467-020-20655-6.
34. Zou, J., Gao, B., Song, Y., and Qin, J. (2022). A review of deep learning-based deformable medical image registration. Front. Oncol. *12*, 1047215. https://doi.org/10.3389/fonc.2022.1047215.
35. Gao, J., Jiang, Q., Zhou, B., and Chen, D. (2019). Convolutional neural networks for computer-aided detection or diagnosis in medical image analysis: An overview. Math. Biosci. Eng. *16*, 6536–6561. https://doi.org/10.3934/mbe.2019326.
36. Pirkulashvili, N., Tsiklauri, N., Nebieridze, M., and Tsagareli, M.G. (2017). Antinociceptive tolerance to NSAIDs in the agranular insular cortex is mediated by opioid mechanism. J. Pain Res. *10*, 1561–1568. https://doi.org/10.2147/JPR.S138360.
37. Tsiklauri, N., Pirkulashvili, N., Nozadze, I., Nebieridze, M., Gurtskaia, G., Abzianidze, E., and Tsagareli, M.G. (2018). Antinociceptive tolerance to NSAIDs in the anterior cingulate cortex is mediated via endogenous opioid mechanism. BMC Pharmacol. Toxicol. *19*, 2. https://doi.org/10.1186/s40360-017-0193-y.
38. Tsagareli, N., Tsiklauri, N., Kvachadze, I., and Tsagareli, M.G. (2020). Endogenous opioid and cannabinoid systems contribute to antinociception produced by administration of NSAIDs into the insular cortex of rats. Biomed. Pharmacother. *131*, 110722. https://doi.org/10.1016/j.biopha.2020.110722.
39. Yang, X.J., Liu, L., Xu, Z.L., Zhang, Y.J., Liu, D.P., Fishers, M., Zhang, L., Sun, J.B., Liu, P.,

Zeng, X., et al. (2020). Baseline brain gray matter volume as a predictor of acupuncture outcome in treating migraine. Front. Neurol. *11*, 111. https://doi.org/10.3389/fneur.2020.00111.

40. Hodkinson, D.J., Khawaja, N., O'Daly, O., Thacker, M.A., Zelaya, F.O., Wooldridge, C.L., Renton, T.F., Williams, S.C.R., and Howard, M.A. (2015). Cerebral analgesic response to nonsteroidal anti-inflammatory drug ibuprofen. Pain *156*, 1301–1310. https://doi.org/10.1097/j.pain.0000000000000176.
41. Yin, T., Sun, G., Tian, Z., Liu, M., Gao, Y., Dong, M., Wu, F., Li, Z., Liang, F., Zeng, F., and Lan, L. (2020). The spontaneous activity pattern of the middle occipital gyrus predicts the clinical efficacy of acupuncture treatment for migraine without aura. Front. Neurol. *11*, 588207. https://doi.org/10.3389/fneur.2020.588207.
42. Liu, F., Zhou, Z., Samsonov, A., Blankenbaker, D., Larison, W., Kanarek, A., Lian, K., Kambhampati, S., and Kijowski, R. (2018). Deep learning approach for evaluating knee MR images: Achieving high diagnostic performance for cartilage lesion detection. Radiology *289*, 160–169. https://doi.org/10.1148/radiol.2018172986.
43. Guan, B., Liu, F., Mizaian, A.H., Demehri, S., Samsonov, A., Guermazi, A., and Kijowski, R. (2022). Deep learning approach to predict pain progression in knee osteoarthritis. Skeletal Radiol. *51*, 363–373. https://doi.org/10.1007/s00256-021-03773-0.
44. Zeleznik, R., Foldyna, B., Eslami, P., Weiss, J., Alexander, I., Taron, J., Parmar, C., Alvi, R.M., Banerji, D., Uno, M., et al. (2021). Deep convolutional neural networks to predict cardiovascular risk from computed tomography. Nat. Commun. *12*, 715. https://doi.org/10.1038/s41467-021-20966-2.
45. Du, Y., Fu, Z., and Calhoun, V.D. (2018). Classification and Prediction of Brain Disorders Using Functional Connectivity: Promising but Challenging. Front. Neurosci. *12*, 525. https://doi.org/10.3389/fnins.2018.00525.
46. Xiang, Y., Zeng, C., Liu, B., Tan, W., Wu, J., Hu, X., Han, Y., Luo, Q., Gong, J., Liu, J., and Li, Y. (2022). Deep learning-enabled identification of autoimmune encephalitis on 3D multi-sequence MRI. J. Magn. Reson. Imag. *55*, 1082–1092. https://doi.org/10.1002/jmri.27909.
47. Nasreddine, Z.S., Phillips, N.A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L., and Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J. Am. Geriatr. Soc. *53*, 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x.
48. Aicher, B., Peil, H., Peil, B., and Diener, H.C. (2012). Pain measurement: Visual Analogue Scale (VAS) and Verbal Rating Scale (VRS) in clinical trials with OTC analgesics in headache. Cephalalgia *32*, 185–197. https://doi.org/10.1177/03331024111430856.
49. Shin, H.E., Park, J.W., Kim, Y.I., and Lee, K.S. (2008). Headache Impact Test-6 (HIT-6) scores for migraine patients: Their relation to disability as measured from a headache diary. J. Clin. Neurol. *4*, 158–163. https://doi.org/10.3988/jcn.2008.4.4.158.
50. Bigal, M.E., Rapoport, A.M., Lipton, R.B., Tepper, S.J., and Sheftell, F.D. (2003). Assessment of migraine disability using the migraine disability assessment (MIDAS) questionnaire: a comparison of chronic migraine with episodic migraine. Headache *43*, 336–342. https://doi.org/10.1046/j.1526-4610.2003.03068.x.
51. Seo, J.G., and Park, S.P. (2015). Validation of the Generalized Anxiety Disorder-7 (GAD-7) and GAD-2 in patients with migraine. J. Headache Pain *16*, 97. https://doi.org/10.1186/s10194-015-0583-8.
52. Seo, J.G., and Park, S.P. (2015). Validation of the Patient Health Questionnaire-9 (PHQ-9) and PHQ-2 in patients with migraine. J. Headache Pain *16*, 65. https://doi.org/10.1186/s10194-015-0552-2.
53. Mollayeva, T., Thurairajah, P., Burton, K., Mollayeva, S., Shapiro, C.M., and Colantonio, A. (2016). The Pittsburgh sleep quality index as a screening tool for sleep dysfunction in clinical and non-clinical samples: A systematic review and meta-analysis. Sleep Med. Rev. *25*, 52–73. https://doi.org/10.1016/j.smrv.2015.01.009.
54. Cheng, J., Tian, S., Yu, L., Gao, C., Kang, X., Ma, X., Wu, W., Liu, S., and Lu, H. (2022). ResGANet : Residual group attention network for medical image classification and segmentation. Med. Image Anal. *76*, 102313. https://doi.org/10.1016/j.media.2021.102313.
55. Yang, R., and Yu, Y. (2021). Artificial convolutional neural network in object detection and semantic segmentation for medical imaging analysis. Front. Oncol. *11*, 638182. https://doi.org/10.3389/fonc.2021.638182.
56. Abedalla, A., Abdullah, M., Al-Ayyoub, M., and Benkhelifa, E. (2021). Chest X-ray pneumothorax segmentation using U-Net with EfficientNet and ResNet architectures. PeerJ. Comput. Sci. *7*, e607. https://doi.org/10.7717/peerj-cs.607.
57. Faul, F., Erdfelder, E., Lang, A.G., and Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods *39*, 175–191. https://doi.org/10.3758/bf03193146.
58. Chen, S.Y., Feng, Z., and Yi, X. (2017). A general introduction to adjustment for multiple comparisons. J. Thorac. Dis. *9*, 1725–1729. https://doi.org/10.21037/jtd.2017.05.34.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Python (version 3.7.15)	Python Software	https://www.python.org
G*Power (version 3.1.9)	G*Power Software	https://www.psychologie.hhu.de

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Hong Zhang (jnyyfsk@126.com).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- All data reported in this paper will be shared by the lead contact upon request.
- All original code has been deposited at github and is publicly available as of the date of publication (https://github.com/IVisonMed/migraine_mri_dl).
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

This study does not use experimental models.

METHOD DETAILS

Patient cohorts

In total, 111 patients who were ≥18 years, with a migraine duration of at least 1 year, and were diagnosed with migraine at a neurological outpatient clinic, were prospectively and continuously enrolled in the study. The patient diagnosis was based on the International Classification of Headache Disorders, 3rd edition (ICHD-3).[1] To control potential pharmacological and physiological effects, the inclusion criteria were as follows: (1) patients who had not taken symptomatic or prophylactic medications for the last two weeks before enrollment; (2) patients in the interictal phase were headache-free for at least 3 days before and 1 day after scanning, which was ascertained via a structured telephonic interview. The general exclusion criteria were as follows: (1) patients with comorbidities with other forms of headache and neuropsychological or neurological disorders; (2) history of previous brain injury or psychoactive medication use; (3) history of alcohol or drug abuse; (4) patients with cognitive impairment with the Montreal Cognitive Assessment (MoCA) scores of <25[47]; (5) pregnant or lactating women, and (6) any contraindications to MRI scanning. Written informed consent was obtained from all study patients to participate in the study. Ethical approval was obtained from the Institutional Review Board of Nanjing Medical University, Nanjing Jiangning Hospital Ethics Committee (2020-03-026-K01).

Questionnaires

All patients with migraine were asked to fill out comprehensive structured questionnaires before the scanning and were followed up via telephone interviews. The information primarily collected were demographic data (e.g., age, sex, and education level), migraine characteristics (e.g., aura symptom, family history, headache location, photophobia, phonophobia, nausea/vomiting, disease duration, frequency, attack duration, headache intensity, impact extent, and burden on quality of life), and neuropsychiatric assessment (e.g., anxiety, depression, and sleep disorder). The intensity, extent, and burden of headaches were assessed using the Visual Analogue Scale (VAS),[48] Headache Impact Test 6-Item (HIT-6),[49] and Migraine Disability Assessment Scale (MIDAS),[50] respectively. Anxiety, depression, and sleep quality symptoms were measured using the Generalized Anxiety Disorder 7-Item (GAD7),[51] Patient Health Questionnaire 9-Item (PHQ-9),[52] and Pittsburgh Sleep Quality Index (PSQI),[53] respectively.

Outcomes measurement

The outcome measure chosen to assess the clinical efficacy of treatment with NSAIDs was migraine intensity. Headache intensity was evaluated on a scale of 0–10, with 0 indicating no pain and 10 indicating the worst pain imaginable. Patients were asked to maintain a headache

diary for VAS scores before and 2 h after drug intake over the following 3 months. The evaluation criteria for therapeutic efficacy were as follows: (1) no pain after 2 h; (2) improvement of pain from moderate to severe pain to mild or no pain (or decrease in VAS score by 50%) after 2 h; (3) the curative effect is repeatable, with effects in at least more than two of the three attacks; and (4) no recurrence or need for medications within 24 h after successful treatment.

Acquisition of 3D structural images

Brain images were acquired using a 3.0 Tesla scanner (Philips, Ingenia) with an 8-channel head coil. T1 images were acquired in three dimensions with the following sequence: repetition time = 8.1 ms, echo time = 3.7 ms, slices = 170, thickness = 1 mm, flip angle = 8°; matrix = 256 × 256, field of view = 256 mm × 256 mm, and voxel size 1 mm × 1 mm × 1 mm. The structural sequence was completed in 5 min and 29 s. During the MRI scan, all subjects were asked to remain conscious, keep their heads steady, and close their eyes.

Construction of the DL model

A 3D CNN was constructed to evaluate the sensitivity of NSAIDs in patients with migraine. Three steps were implemented for preprocessing. First, all two-dimensional DICOM slices were concatenated to a 3D pixel matrix. Second, all input images were normalized and padded to the same size with a width × height × depth of 256 × 256 × 170 pixels. Third, image augmentations, including random Gaussian noise, rotation, scaling, and flipping, were introduced to suppress the overfitting. The 3D ResNet18 was introduced as the classification backbone because it has demonstrated excellent results in medical image classification,[54] object detection,[55] and lesion segmentation,[56] effectively alleviating the challenges of gradient disappearance and network degradation caused by an increase in network depth. It included four feature extraction stages and one classification stage. Each feature extraction stage was stacked with several residual units, each of which contains a convolution layer, pooling layer, and rectified linear unit (ReLU) activation layer. The output size in the four feature extraction stages were 56, 28, 14, and 7, respectively. The output of the feature map from the last feature extraction stage contained high-level semantic information, which was fed into the fully connected layer to predict the sensitivity of NSAIDs in patients with migraine (Below figure).

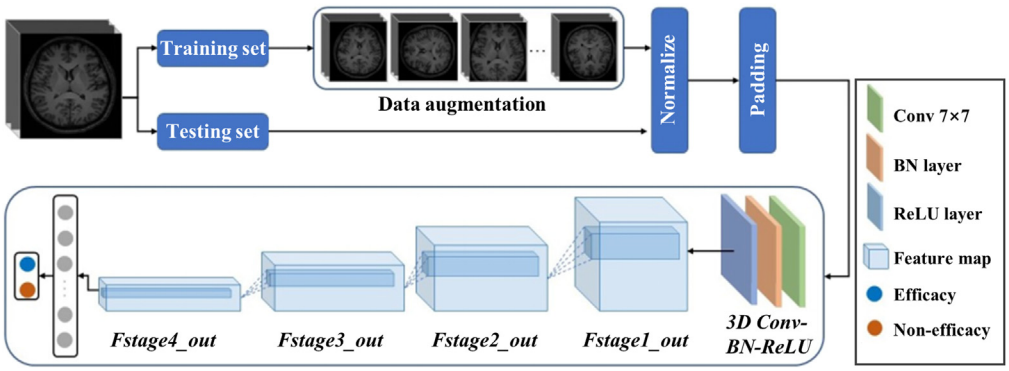


Illustration of the architecture of the deep learning model for predicting NSAIDs efficacy

BN, batch normalization; Conv, convolution; NSAIDs, non-steroidal anti-inflammatory drugs; ReLU, rectified linear unit.

Setting of experimental parameters

The classification model was optimized with AdamW and cross-entropy loss. The training batch size, learning rate, and epochs were 8, 0.0005, and 100, respectively. Two GTX 3090 graphics processing units were used, and the operating system was Ubuntu 20.04 with CUDA version 11.3. Model implementation was performed in Python 3.7.15 using PyTorch 1.11.0.

Model evaluation

The performance of the developed DL model was evaluated via ROC curve analysis by calculating the AUC. The confidence intervals (CIs) of the AUCs were calculated using the DeLong method. Further, the accuracy, recall, precision, F1-score, PLR, NLR, and cutoff values of all models were calculated.

QUANTIFICATION AND STATISTICAL ANALYSIS

All analyses were performed using SPSS software (SPSS version 24.0). Post hoc power analysis was performed using G*Power software (version 3.1.9, effect size = 0.8, α = 0.05, two-tail).[57] Data included both continuous and categorical variables. Continuous variables following normal distribution were presented as mean $\pm$ standard deviation and compared using the Student's *t* test, whereas those following nonnormal distribution were presented as median (interquartile range) and compared using the Mann-Whitney U test. Categorical variables were analyzed using the Chi-square test or Fisher's exact test. Due to the considerable number of statistical tests conducted, we employed a Bonferroni correction to address the issue of multiple testing.[58] Therefore, the significance level p = 0.05 was divided by 18, which provides a significance level corrected for multiple testing (p = 0.003).