AJNR

Automated Idiopathic Normal Pressure Hydrocephalus Diagnosis via Artificial Intelligence–Based 3D T1 MRI Volumetric Analysis

Joonhyung Lee, Dana Kim, Chong Hyun Suh, Suyoung Yun, Kyu Sung Choi, Seungjun Lee, Wooseok Jung, Jinyoung Kim, Hwon Heo, Woo Hyun Shim, Sungyang Jo, Sun Ju Chung, Jae-Sung Lim, Ho Sung Kim, Sang Joon Kim and Jae-Hong Lee

AJNR Am J Neuroradiol published online 9 September 2024
http://www.ajnr.org/content/early/2025/01/02/ajnr.A8489

Published January 2, 2025 as 10.3174/ajnr.A8489

Automated Idiopathic Normal Pressure Hydrocephalus Diagnosis via Artificial Intelligence–Based 3D T1 MRI Volumetric Analysis

Joonhyung Lee, Dana Kim, Chong Hyun Suh, Suyoung Yun, Kyu Sung Choi, Seungjun Lee, Wooseok Jung, Jinyoung Kim, Hwon Heo, Woo Hyun Shim, Sungyang Jo, Sun Ju Chung, Jae-Sung Lim, Ho Sung Kim, Sang Joon Kim, and Jae-Hong Lee

ABSTRACT

BACKGROUND AND PURPOSE: Idiopathic normal pressure hydrocephalus (iNPH) is reversible dementia that is underdiagnosed. The purpose of this study was to develop an automated diagnostic method for iNPH using artificial intelligence techniques with a T1-weighted MRI scan.

MATERIALS AND METHODS: We quantified iNPH, Parkinson disease, Alzheimer disease, and healthy controls on T1-weighted 3D brain MRI scans using 452 scans for training and 110 scans for testing. Automatic component measurement algorithms were developed for the Evans index, Sylvian fissure enlargement, high-convexity tightness, callosal angle, and normalized lateral ventricle volume. XGBoost models were trained for both automated measurements and manual labels for iNPH prediction.

RESULTS: A total of 452 patients (200 men; mean age, 73.2 [SD, 6.5] years) were included in the training set. Of the 452 patients, 111 (24.6%) had iNPH. We obtained area under the curve (AUC) values of 0.956 for automatically measured high-convexity tightness and 0.830 for Sylvian fissure enlargement. Intraclass correlation values of 0.824 for the callosal angle and 0.924 for the Evans index were measured. By means of the decision tree of the XGBoost model, the model trained on manual labels obtained an average cross-validation AUC of 0.988 on the training set and 0.938 on the unseen test set, while the fully automated model obtained a cross-validation AUC of 0.983 and an unseen test AUC of 0.936.

CONCLUSIONS: We demonstrated a machine learning algorithm capable of diagnosing iNPH from a 3D T1-weighted MRI that is robust to the failure. We propose a method to scan large numbers of 3D T1-weighted MRIs with minimal human intervention, making possible large-scale iNPH screening.

ABBREVIATIONS: AD = Alzheimer disease; AUC = area under the curve; DESH = disproportionately enlarged subarachnoid space hydrocephalus; HC = healthy controls; ICV = intracranial volume; iNPH = idiopathic normal-pressure hydrocephalus; PD = Parkinson disease; ROC = receiver operating characteristic; SENSE = sensitivity encoding

Idiopathic normal-pressure hydrocephalus (iNPH) is a neurologic condition that shows the typical triad of the following symptoms: gait disturbance, cognitive impairment, and urinary incontinence.[1] The iNPH is often reversible via CSF shunt surgery,[2,3] and when patients are accurately selected, >90% showed clinical improvement,[4] which strongly implies that appropriate diagnostic procedures are needed.[5] Although it is one of the reversible forms of dementia and is treatable in its early stages,[6] iNPH often goes undiagnosed and untreated because it is difficult to clinically differentiate iNPH from other similar neurodegenerative diseases.[7-9] This explanation is supported by data from the Hydrocephalus Association, which estimates that 80% of patients with normal pressure hydrocephalus are not identified, often being incorrectly diagnosed with Alzheimer disease (AD) or Parkinson disease (PD).[10]

Received March 8, 2024; accepted after revision July 24.

From the NAVER Cloud Inc (J.L.), Seoul, Republic of Korea; University of Ulsan College of Medicine (D.K.,), Seoul, Republic of Korea; Department of Radiology and Research Institute of Radiology (C.H.S., H.H., W.H.S., H.S.K., S.J.K.), and Department of Neurology (S.J., S.J.C., J.-S.L., J.-H.L.), Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea; Department of Radiology (S.Y.), Busan Paik Hospital, Inje University College of Medicine, Busan, Republic of Korea; Department of Radiology (K.S.C.), Seoul National University Hospital, Seoul, Republic of Korea; and VUNO Inc (J.L., S.L., W.J., J.K.), Seoul, Republic of Korea.

J. Lee and D. Kim contributed equally to this work.
C.H. Suh and S. Yun contributed equally to this work.
Work was performed at VUNO Inc.

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute, funded by the Ministry of Health & Welfare, Republic of Korea (HR20C0026) and the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education (NRF-2022R1I1A1A01072397).

Please address correspondence to Chong Hyun Suh, MD, PhD, Department of Radiology and Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, 88, Olympic-ro 43-gil, Songpa-gu, Seoul, Republic of Korea (05505); e-mail: chonghyunsuh@amc.seoul.kr

Indicates article with online supplemental data.

http://dx.doi.org/10.3174/ajnr.A8489

SUMMARY

PREVIOUS LITERATURE: Disproportionately enlarged subarachnoid space hydrocephalus, a representative radiologic finding of iNPH, has interobserver variability.

KEY FINDINGS: We were able to provide automatic component measurement algorithms for measuring the callosal angle and the Evans index and determining the presence of high-convexity subarachnoid space tightness and Sylvian fissure enlargement, which showed good diagnostic performance.

KNOWLEDGE ADVANCEMENT: We demonstrated a fully automated machine learning algorithm capable of diagnosing iNPH from a single 3D T1-weighted MRI that is robust to the failure of any of its component algorithms.

The commonly used diagnostic imaging biomarkers for iNPH are disproportionately enlarged subarachnoid space hydrocephalus (DESH), callosal angle, and the Evans index.[1,11-14] In a previous meta-analysis,[12] the sensitivity and specificity of the callosal angle for diagnosing iNPH were 91% and 93%, respectively, and those of the Evans index were 96% and 83%. However, the pooled prevalence of DESH in iNPH was only 44% (95% CI, 34%–54%), despite its high positive predictive value.[11] Moreover, in a previous study in which a brain morphometry–based nomogram was created on the basis of visual assistance, DESH alone was inappropriate for diagnosing iNPH due to its high interobserver variability.[15]

Regarding the diagnosis of iNPH, previous work using machine learning has helped measure only the ventricle volume or callosal angle.[9,15-18] To the best of our knowledge, studies aiming to assess DESH or diagnose iNPH on the basis of deep learning have not been reported. We hypothesize that by developing an algorithm for measuring DESH, the Evans index, and the callosal angle using machine learning, we will be able to obtain a biomarker for iNPH diagnosis that is more objective and reproducible than with human ratings alone. Therefore, this study aimed to develop an algorithm for measuring DESH, the Evans index, and the callosal angle using a combination of machine learning and deep learning techniques and to propose an automated iNPH diagnostic method with only a single T1-weighted MRI. We expect to achieve an appropriate algorithm and prediction model using FreeSurfer (http://surfer.nmr.mgh.harvard.edu) for data-preprocessing and a deep learning model for segmentation, which was developed with healthy controls (HC).

MATERIALS AND METHODS

VUNO (Seoul, Republic of Korea) provided technical support for analyzing totally automatic ventricle segmentation and by providing VUNO Med-DeepBrain for the analysis. In addition, 1 author (J.L.) was an employee of VUNO, and 2 other authors (S.L., W.J.) are currently employees of VUNO. However, VUNO did not have any role in the study design, data collection, or interpretation. The other authors did not have any conflicts of interest. Two institutions (Asan medical center and Seoul National University Hospital) were involved in this study.

This retrospective, observational, multi-institutional study was approved by the appropriate institutional review board (Asan medical center and Seoul National University Hospital), which waived the requirement for written informed consent. We followed the Standards for Reporting Diagnostic Accuracy (STARD) guidelines,[19] the Enhancing the Quality and Transparency of Health Research (EQUATOR) reporting guidelines,[20] and conformed to the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) standards.[21]

Training and Testing Data Sets

We quantified iNPH, PD, AD, and HC on T1-weighted 3D brain MRI scans using 452 scans from the Asan Medical Center (AMC) for training and 110 scans from Seoul National University Hospital (SNUH) for testing. AD was diagnosed in accordance with the clinical diagnostic guidelines of the National Institute on Aging-Alzheimer's Association workgroups,[22] and PD was diagnosed in accordance with the UK Parkinson Disease Society Brain Bank clinical diagnostic criteria.[23] Possible, probable, or definite iNPH was diagnosed in accordance with the Japanese guidelines.[1] Satisfaction with these possible, probable, and definite criteria was confirmed by 1 neurologist (S. J.). The patients in this study overlapped with those in a previous study,[15] but the purpose of analyzing the data differed. The diagnostic criteria for iNPH are described below.

Possible iNPH was diagnosed on the basis of the following criteria:

1) Presence of >1 symptom in the triad of gait disturbance, cognitive impairment, and urinary incontinence
2) Presence of clinical symptoms that cannot be completely explained by other neurologic or non-neurologic diseases
3) Absence of any obvious preceding disease that may be the cause of ventricular dilation.

Probable iNPH was diagnosed if a patient had all the following 3 features:

1) Met the requirement for possible iNPH
2) Had a CSF pressure of ≤ 200 mm H_2O and normal CSF content
3) Had one of the following 2 findings:

- Neuroimaging features of narrowing of the sulci and subarachnoid space over the high-convexity/midline surface (DESH) with gait disturbance
- Improvement of symptoms after a CSF tap test or drainage test.

Definite iNPH was diagnosed when objective improvement of symptoms was observed after CSF shunt surgery.

Of the 452 patients in the training data set, 111 had iNPH (possible: 29/111, 26%, probable: 63/111, 57%, definite: 19/111, 17%), whereas 28 (possible: 2/28, 7%, probable: 26/28, 93%, definite: 0, 0%) of the 110 patients in the testing data set had iNPH. Evans index, Sylvian fissure enlargement, high-convexity tightness, callosal angle, and normalized lateral ventricle volume were used for diagnosis. Manual labels are available for all these indicators except normalized lateral ventricle volume.

MRI Protocol

MRI data were obtained using various MRI machines at multiple institutions. MRI evaluations at AMC (the training data set) were performed on a 3T unit (Ingenia CX; Philips Healthcare) using a 32-channel sensitivity encoding (SENSE) head coil. High-resolution anatomic 3D volume images were obtained in the sagittal plane using a 3D gradient-echo T1-weighted sequence. The detailed parameters were as follows: TR, 9.6 ms; TE, 4.6 ms; flip angle, 8°; field of view, 224 × 224 mm; section thickness, 1 mm with no gap; matrix size, 224 × 224; and total scan time, 6 minutes 11 seconds.

MRI evaluations at SNUH (the unseen test data set) were performed on a 3T unit (Magnetom Skyra; Siemens) using a 32-channel SENSE head coil for those with AD and HC. High-resolution anatomic 3D volume images were obtained in the sagittal plane using a T1-weighted MPRAGE sequence. The detailed parameters were as follows: TR, 1600 ms; TE, 1.89 ms; flip angle, 9°; field of view, 250 × 250 mm; section thickness, 1 mm with no gap; matrix size, 256 × 256; and total scan time, 6 minutes 48 seconds. For the PD group, MRI on a 3T unit (Discovery MR750w; GE Healthcare) using a 32-channel SENSE head coil was performed. High-resolution anatomic 3D volume images were obtained in the sagittal plane using a T1-weighted fast-spoiled gradient recalled sequence. The detailed parameters were as follows: TR, 8.5 ms; TE, 3.2 ms; flip angle, 12°; field of view, 256 × 256 mm; section thickness, 1 mm with no gap; matrix size, 256 × 256; and total scan time, 3 minutes 30 seconds. For the iNPH group, MRI data were obtained using various MRI machines at multiple institutions.

Manual Measurement

For the reference standard, the MRI findings were retrospectively reviewed in consensus by 2 radiologists (S.Y. and C.H.S., with 5 and 10 years of clinical experience in neuroradiology, respectively) for the presence of tightness of the high-convexity subarachnoid space and Sylvian fissure enlargement. Any discrepancy between the 2 radiologists was resolved by a third radiologist (S.J.K., with 35 years of clinical experience in neuroradiology). The Evans index and the callosal angle were manually measured solely by 1 radiologist (S.Y., with 5 years of clinical experience in neuroradiology) using an in-house PACS. The reviewers were blinded to the final diagnosis and other predictors of each patient.

To investigate these MRI features of iNPH, we performed the following evaluations:

1) The tightness of the high-convexity subarachnoid space was defined as the narrow CSF space at the medial and/or high-convexity cortex sulci located above the body of the lateral ventricles in the coronal plane. The presence of tightness of the high-convexity subarachnoid space was visually evaluated and marked "Yes" or "No."
2) The presence of Sylvian fissure enlargement in the coronal plane was visually evaluated and marked "Yes" or "No."
3) The Evans index was calculated as the ratio of the maximum diameter of the frontal horns of the lateral ventricles to the maximum inner diameter of the skull in transverse sections.[24]
4) The callosal angle, the angle between the left and right corpus callosum, was measured in the coronal image perpendicular to the anterior/posterior commissure plane at the posterior commissure.[14]

Development of Automatic Component Measurement Algorithms

The scans used for the experiments were first preprocessed using FreeSurfer (Version 7.3.2)[25] and FSL (Version 6.0.5.2)[26] to minimize interscan variability. We found that the exact data-processing steps taken have a nontrivial effect on the performance of downstream algorithms. The FSL FMRIB Linear Image Registration Tool (FLIRT; http://www.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT) was first used to align the scan to the Montreal Neurological Institute 152 template,[27] normalize the brain size, and clip outlier intensity values. The registration process used spline interpolation to minimize any errors. In addition to helping downstream tools process the scans correctly, brain size normalization simplified the specification of brain region sizes, which have greater variability in the original scans. The resulting volumes were then processed with FreeSurfer automatic reconstruction stage 1 to obtain bias-field-corrected and intensity-normalized scans.

We applied a 3D Swin Transformer[28] (https://github.com/microsoft/Swin-Transformer) deep neural network trained on the parcellation outputs of >3000 healthy patients to obtain brain region segmentations for the processed scans. The neural network was trained using output from FreeSurfer, Version 7.3.2, with additional output for brainstem substructure segmentation. After merging classes present on both hemispheres, we obtained 56 brain region classes and 1 background class. The model used patches of size 64 × 64 × 64 sampled from the scan so that the input patches always contained brain tissue.

A separate neural network trained on manual labels was used to obtain the intracranial volume (ICV) segmentations.[29] A hole-filling mechanism was applied to the largest connected component of the model output to take advantage of the anatomic knowledge that the ICV is a single contiguous volume with no holes or disconnected regions.

The callosal angle, defined as the angle of the corpus callosum at the coronal slice containing the posterior commissure, was measured by taking the angle between the 2 sides of the minimum enclosing triangles of each ventricle (Fig 1*A*). The acpcdetect package[30] (https://github.com/tannerjared/MRI_Guide/blob/master/install_acpcdetect.md) was used to detect the posterior commissure, where a single coronal slice was taken to measure the callosal angle. For angle measurement, a triangle containing the corpus callosum and lateral ventricle was drawn on each hemisphere using OpenCV (https://opencv.org/).[31] The

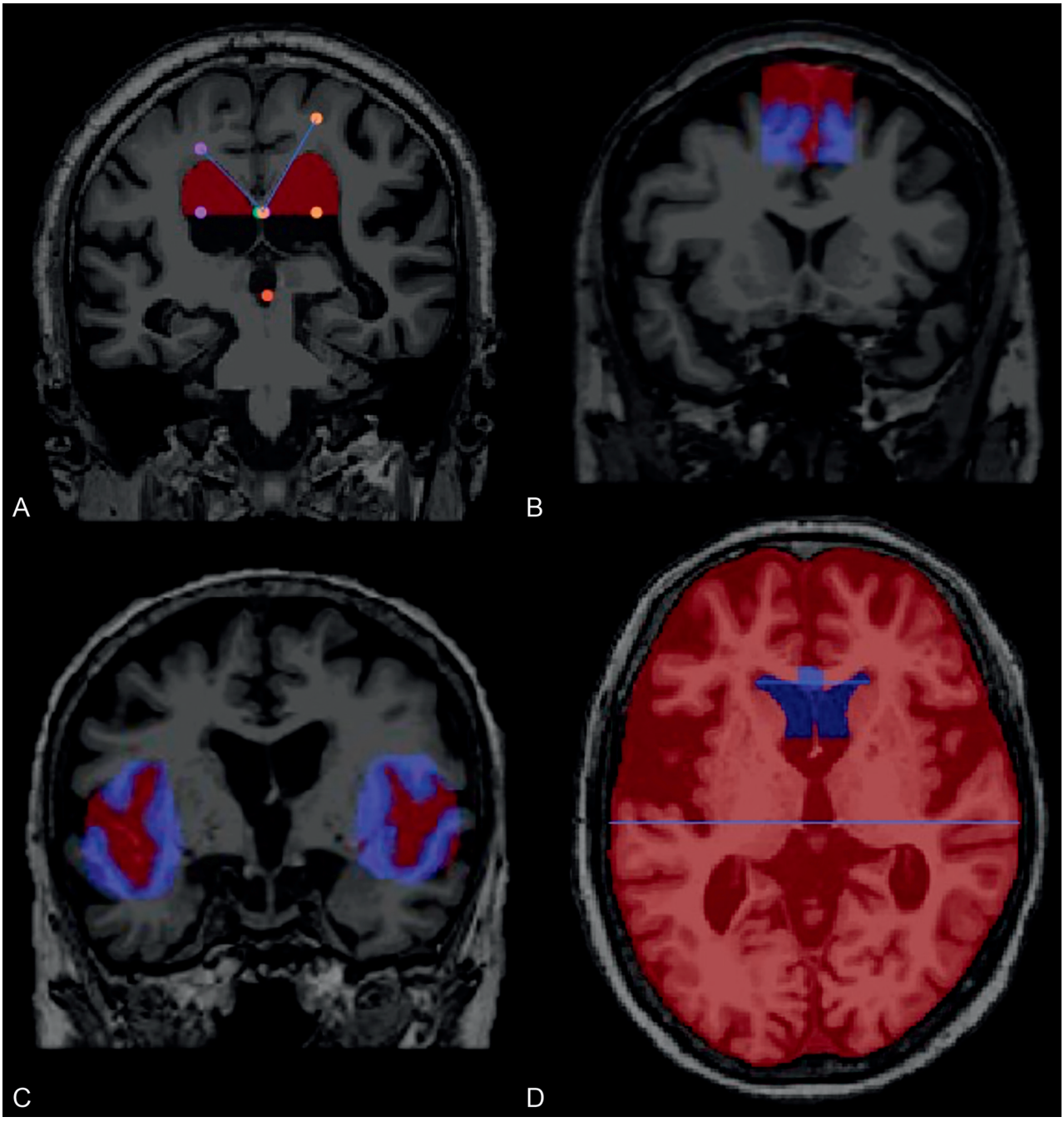

FIG 1. Callosal angle (*A*), high-convexity tightness (*B*), Sylvian fissure widening (*C*), and Evans index (*D*). Note that the results vary significantly depending on the preprocessing procedure. For the best results, we apply 7 *df* linear registration using FSL FLIRT, which applies intensity clamping, aligns the scan to the Montreal Neurological Institute 152 template, and normalizes the brain size. The resulting scans are then intensity-normalized and bias-field corrected with FreeSurfer, Version 7.3.2.

enclosing triangles of the callosal angle were then measured from the coronal slice at the posterior commissure by measuring the angle between the hypotenuses of the resulting triangles. To improve the quality of the measurement, we considered only the region between the peaks of the corpus callosum.

High-convexity subarachnoid space tightness and Sylvian fissure enlargement were both measured by the ratio of nonbrain regions in hand-crafted ROIs. For high-convexity tightness, as shown in Fig 1*B*, the upper three-quarters of the ICV region between the 2 peaks of the ventricles was used. The ventricle peaks were measured for each coronal slice separately. The ICV volume was calculated by summing the voxels for each axial slice within the area of the previously specified ventricular peaks. Sylvian fissure widening (Fig 1*C*) was measured by selecting the combined superior temporal, supramarginal, transverse temporal cortices, and insula regions around the Sylvian fissure from both hemispheres. The largest connected component of this combined region was extracted, and its 3D

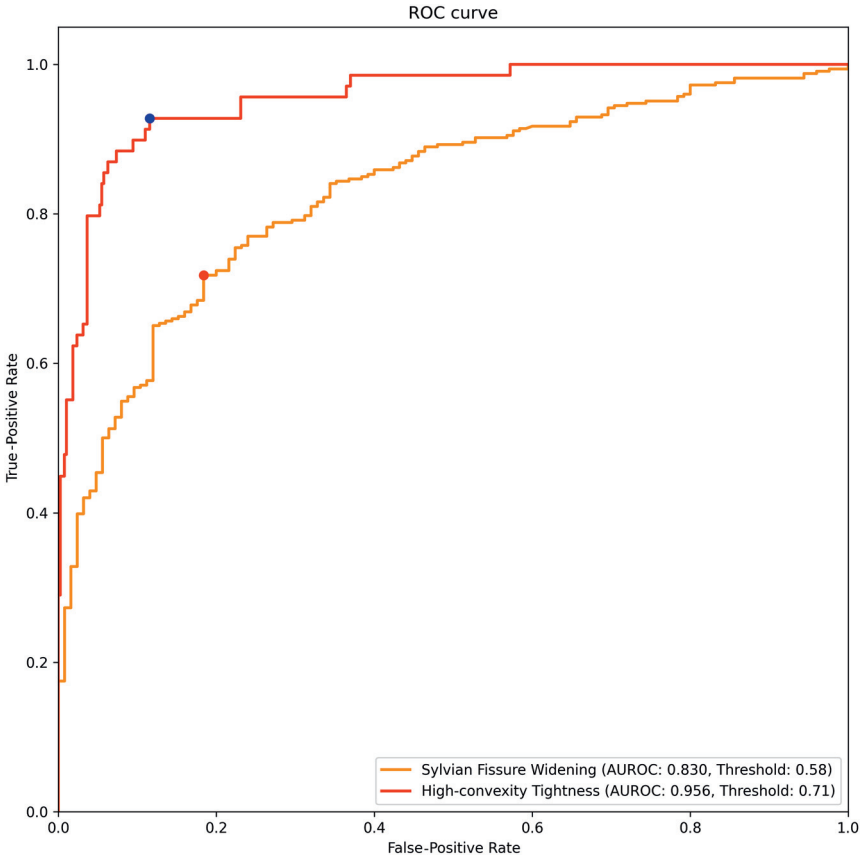

FIG 2. ROC curves for high-convexity tightness and Sylvian fissure widening measured by automatic methods. The volume of nontissue regions relative to the ROI volume gives prediction scores for the ROC curve.

convex hull defined the ROI. Within this ROI, the ratio of CSF to the total ROI volume was calculated as the Sylvian fissure widening metric.

The Evans index, a commonly used proxy for the relative volume of the lateral ventricles, was computed by localizing the frontal horns via the corpus callosum. The maximum horizontal distance between the ventricles at any coronal plane anterior to the frontal corpus callous was divided by the maximum horizontal ICV distance at the axial slice of the discovered plane as shown in Fig 1*D*.

We also include measurements for the normalized lateral ventricle volume,[15] defined as the ratio of the lateral ventricle volume to the total intracranial ventricle volume, which we found to be a strong predictor of iNPH. We speculate that the normalized lateral ventricle volume is an important predictor because it directly measures the volume of the CSF in the lateral ventricles as opposed to being a proxy metric used for ease of measurement. The discovery of the normalized lateral ventricle volume has the additional benefit of motivating the use of automated measurements because it is time-consuming to measure manually. Note also that the formulation of both the Evans index and normalized lateral ventricle volume allows them to be calculated even for the size-normalized images obtained via the proposed preprocessing steps, because they measure the relative, not absolute, sizes of brain regions.

Statistical Analyses

To measure the performance on each of the modalities, we compared the model prediction results with the manual labels. In addition, to enhance interpretability and for fair comparison with the manual labels, we indicated the presence or absence by discretizing each discrete feature using task-specific thresholds on the raw contiguous output values.

For the high-convexity tightness and Sylvian fissure enlargement, we measured the area under the receiver operating characteristic curve between the obtained volume ratios and manual labels using the Sci-kit learn Python package (https://scikit-learn.org/stable/index.html)[32] to obtain thresholds of 70% and 25% based on the cross-validated results from the training set for the performance each module and rounding by 5%. Test set data were excluded from the cross-validation measurements when selecting the thresholds. See Fig 2 for the cross-validation results. For the callosal angle and Evans index, both of which are continuous measurements, we measured the intraclass correlation with single fixed raters compared with the manual labels to measure the performance using the Pingouin[33] Python package (https://pypi.org/project/pingouin/).

The threshold for the normalized lateral ventricle volume was set to 5% after a preliminary investigation in which XGBoost models (https://xgboost.readthedocs.io/en/stable/) were found to produce the most accurate iNPH predictions when the threshold was set to this region. Following standard medical practice, the callosal angle threshold was set at 90°, while the Evans index used the thresholds of 0.25 and 0.3 for patients with mild and severe conditions, respectively. The manual labels were also categorized with the same threshold values for the callosal angle and Evans index.

Development of iNPH Prediction Models: Manual Measurement–Based and Fully Automated Models

The radiologic findings used to develop the 2 models are as follows: the tightness of the high-convexity subarachnoid space, Sylvian fissure enlargement, the Evans index, and the callosal angle. To develop the fully automated model, we also used the normalized lateral ventricle volume information.

For the final iNPH prediction, XGBoost[34] models were trained to make predictions based on the output features of the individual modules. Separate models were trained for both the discretized automated measurements and the manual labels to predict iNPH in individual patients.

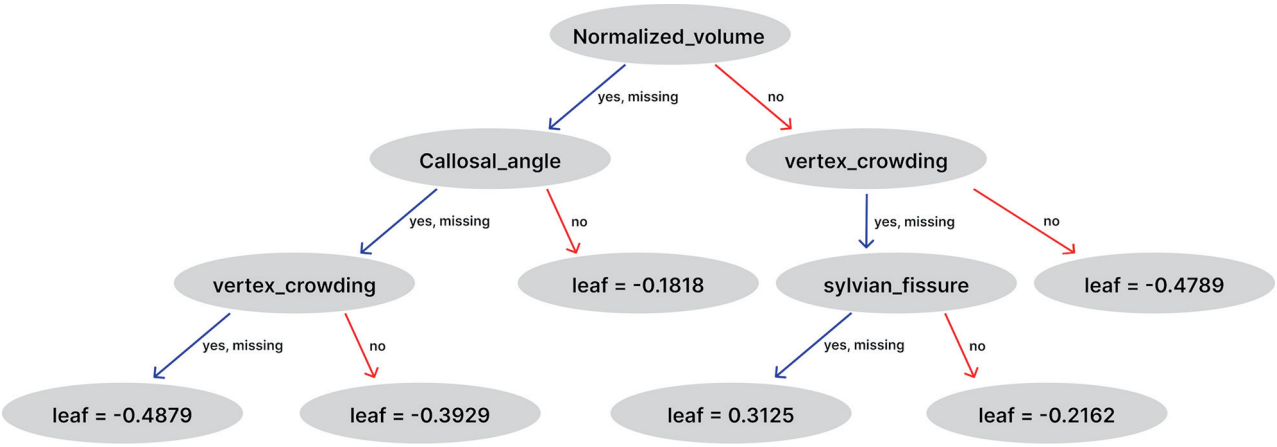

FIG 3. Decision tree of the XGBoost model with normalized lateral ventricle volume, callosal angle, high-convexity tightness (vertex region crowding), Sylvian fissure widening, and Evans index.

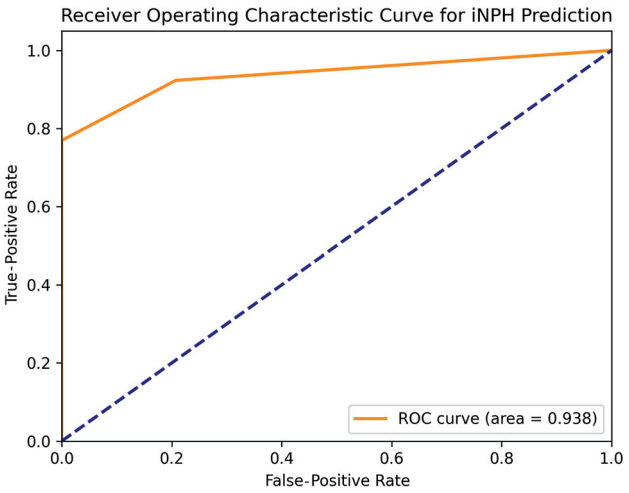

FIG 4. The ROC curve for our XGBoost model for the diagnosis of iNPH in the unseen test data set.

The performance of each prediction model was calculated using the receiver operating characteristic (ROC) curve with unseen test data set and summarized by calculating the area under the curve (AUC) of the ROC curve. To increase the robustness of the performance estimation, we performed 5-fold cross-validation.

RESULTS

Patient Demographics

A total of 546 Asian patients were selected for the training data set. Of these, 94 patients were excluded. Therefore, the training data set consisted of 452 patients (200 men, 252 women; mean age, 73.2 [SD, 6.5] years; Online Supplemental Data), with 111 (24.6%) cases of iNPH, 101 (22.3%) cases of AD, 103 (22.8%) cases of PD, and 137 (30.3%) HC. The unseen test data set included all 110 Asian patients (48 men and 62 women; mean age, 72.4 [SD, 7.7] years), with 28 (25.5%) cases of iNPH, 28 (25.5%) cases of AD, 26 (23.6%) cases of PD, and 28 (25.5%) HC.

Diagnostic performance of iNPH prediction models

Model	AUC: Unseen Test Data Set	Average Cross-Validation AUC: Training Data Set
Manual measurement–based model	0.938	0.988
Fully automated model	0.936	0.983

Diagnostic Performance of Component Measurement Algorithms

We obtained AUC values of 0.956 for automatically measured high-convexity tightness and 0.830 for Sylvian fissure enlargement (Fig 2). In addition, intraclass correlation values (single fixed raters) of 0.824 for the callosal angle and 0.924 for the Evans index were measured.

Diagnostic Performance of iNPH Prediction Models

Figure 3 depicts the decision tree of the XGBoost model when automatically measured features are provided as input. Figure 4 shows the ROC curve for our XGBoost model for the diagnosis of iNPH in the unseen test data set. The model trained on manual labels obtained an average cross-validation AUC of 0.988 on the training data set and 0.938 on the unseen test data set, while the fully automated model obtained a cross-validation AUC of 0.983 and a test AUC of 0.936 (Table). We obtained these results despite 4% of the scans failing at least one of the measurements, with the posterior commissure detection in the callosal angle measurement being the most common cause of failure. In addition, when the "possible" patients with iNPH were removed from both data sets, both models overfit the training data set.

In addition, we evaluated the AUC for discriminating between iNPH and AD, as well as between iNPH and PD. For distinguishing iNPH from AD, our model achieved an AUC of 0.90. When discriminating iNPH from PD, the model demonstrated an even higher AUC of 0.95. These results indicate that our model is highly effective in differentiating iNPH from PD.

DISCUSSION

In this study, we were able to provide automatic component measurement algorithms for measuring the callosal angle and Evans index and for determining the presence of high-convexity subarachnoid space tightness and Sylvian fissure enlargement, which showed good diagnostic performance. We were also able to provide a fully automated machine learning algorithm capable of diagnosing iNPH from a single 3D T1-weighted MRI. The automated iNPH prediction model also showed good discrimination ability. We internally validated the algorithm and tested with the unseen test data set, which showed an AUC of 0.983 for cross-validation and 0.936 for the unseen test data set. The fully automated iNPH prediction model has the possibility of helping alleviate the burden caused by a treatable disease that too often goes undiagnosed.

Unlike naïve deep learning–based predictions, our proposed method has the twin advantages of interpretability and robustness to failure. Because the decision tree was constructed using radiologic findings well known to neuroradiologists and clinicians, such as DESH, callosal angle, and the Evans index, disease prediction is possible through a process similar to how a neuroradiologist diagnoses iNPH. With interpretable features that can be confirmed visually, our work makes possible automated iNPH diagnosis at scale. Whereas the failure of any of the component algorithms in our method results in only a partial performance degradation, the failure of an end-to-end machine learning model could be highly inaccurate.

In this study, we also provide component measurement algorithms for each radiologic finding. These algorithms determine the presence of Sylvian fissure enlargement and high-convexity subarachnoid space tightness by setting thresholds. Therefore, a more objective evaluation than visual assessment is possible, and bias caused by human judgment can be reduced. By means of this algorithm, it may be possible to evaluate whether DESH is related to the shunt responsiveness in patients with iNPH; the answer is still unclear. Additionally, because this algorithm uses segmentation to obtain the volume ratio, further research on the size of this ratio, that is, the correlation between shunt responsiveness and DESH severity degree, will be possible.

There has been some effort to develop models that predict the probability of iNPH or predict the treatment response. According to a systematic review[35] that analyzed 22 articles published through 2022, the most used input data were MRI, as in our study. That review included various machine learning studies related to iNPH. Among these, there were only 4 articles that used MRI input data and presented iNPH probability or NPH status as output. Among these 4 articles, none presented testing data. As in our study, various studies attempting to predict iNPH using segmentation[36] are ongoing. However, the distinctiveness of our developed model lies in its independence from the black box nature of artificial intelligence. Instead, it focuses on detecting well-known iNPH findings, allowing the user to make judgments in a manner similar to that of a neuroradiologist.

There are several limitations in this model. First, 4% of the scans failed for at least one of the component measurement algorithms. The most error-prone was the callosal angle measurement, with 75% of failures due to this modularity. The posterior commissure was often incorrectly located in the cerebellum by acpcdetect, and the triangle measurements also had difficulty with patients showing signs of severe iNPH. Adding rules that catch anatomically infeasible results could aid in error detection. Second, the current implementation requires 2 hours of preprocessing for each scan, making it impractical for real-time deployment. Preprocessing could be accelerated by training a separate neural network to perform both intensity clipping and bias field correction instead of relying on pre-existing image-processing tools. In addition, the method must be verified across a greater variety of scanning protocols, especially accelerated acquisition methods, because learning-based methods are highly sensitive to the MRI acquisition protocols used to acquire the scans. Many scans may be available that have been collected for purposes other than iNPH detection, using acquisition settings that may be optimized for other diagnoses. This possibility is important to reduce the costs associated with acquiring an additional scan specifically for iNPH measurement and would allow retrospective studies for scans acquired with different settings. Third, we used patients with AD, PD, HC, and patients with iNPH as subjects. We included AD, PD, and aged HC because they may have clinical symptoms and MRI findings similar to those of iNPH. However, this inclusion led to an imbalanced data set. Nevertheless, we believe we have demonstrated the robustness of our model through the use of an unseen test data set. Last, a single neuroradiologist measured the callosal angle and Evans index, possibly causing bias. However, because the measurements are clearly defined values and the neuroradiologist who measured them also had sufficient clinical experience, the impact on internal validity is expected to be minimal.

CONCLUSIONS

We demonstrated a machine learning algorithm capable of diagnosing iNPH from a single 3D T1-weighted MRI scan that is robust to the failure of any of its component algorithms. We propose a method to scan large numbers of 3D T1-weighted MRI scans with minimal human intervention, making possible large-scale iNPH screening.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

1. Nakajima M, Yamada S, Miyajima M, et al; The research committee of idiopathic normal pressure hydrocephalus. **Guidelines for management of idiopathic normal pressure hydrocephalus: endorsed by the Japanese Society of Normal Pressure Hydrocephalus.** *Neurol Med Chir (Tokyo)* 2021;61:63–97 CrossRef Medline
2. Toma AK, Papadopoulos MC, Stapleton S, et al. **Systematic review of the outcome of shunt surgery in idiopathic normal-pressure hydrocephalus.** *Acta Neurochir (Wien)* 2013;155:1977–80 CrossRef Medline
3. Popal AM, Zhu Z, Guo X, et al. **Outcomes of ventriculoperitoneal shunt in patients with idiopathic normal-pressure hydrocephalus 2 years after surgery.** *Front Surg* 2021;8:641561 CrossRef Medline
4. Trungu S, Scollato A, Ricciardi L, et al. **Clinical outcomes of shunting in normal pressure hydrocephalus: a multicenter prospective observational study.** *J Clin Med* 2022;11:1286 CrossRef Medline
5. Nassar BR, Lippa CF. **Idiopathic normal pressure hydrocephalus: a review for general practitioners.** *Gerontol Geriatr Med* 2016;2:2333721416643702 CrossRef Medline

6. Little MO. **Reversible dementias.** *Clin Geriatr Med* 2018;34:537–62 CrossRef Medline
7. Fällmar D, Andersson O, Kilander L, et al. **Imaging features associated with idiopathic normal pressure hydrocephalus have high specificity even when comparing with vascular dementia and atypical parkinsonism**. *Fluids Barriers CNS* 2021;18:35 CrossRef Medline
8. Molde K, Söderström L, Laurell K. **Parkinsonian symptoms in normal pressure hydrocephalus: a population-based study**. *J Neurol* 2017;264:2141–48 CrossRef Medline
9. Serulle Y, Rusinek H, Kirov II, et al. **Differentiating shunt-responsive normal pressure hydrocephalus from Alzheimer disease and normal aging: pilot study using automated MRI brain tissue segmentation.** *J Neurol* 2014;261:1994–2002 CrossRef Medline
10. **What is normal pressure hydrocephalus**? https://www.hydroassoc.org/normal-pressure-hydrocephalus-2/. Accessed July 04, 2024
11. Park H, Park C, Suh C, et al. **Prognostic utility of disproportionately enlarged subarachnoid space hydrocephalus in idiopathic normal pressure hydrocephalus treated with ventriculoperitoneal shunt surgery: a systematic review and meta-analysis.** *AJNR Am J Neuroradiol* 2021;42:1429–36 CrossRef Medline
12. Park HY, Kim M, Suh CH, et al. **Diagnostic performance and interobserver agreement of the callosal angle and Evans' index in idiopathic normal pressure hydrocephalus: a systematic review and meta-analysis**. *Eur Radiol* 2021;31:5300–11 CrossRef Medline
13. Hashimoto M, Ishikawa M, Mori E, et al; Study of INPH on neurological improvement (SINPHONI). **Diagnosis of idiopathic normal pressure hydrocephalus is supported by MRI-based scheme: a prospective cohort study.** *Cerebrospinal Fluid Res* 2010;7:18 CrossRef Medline
14. Ishii K, Kanda T, Harada A, et al. **Clinical impact of the callosal angle in the diagnosis of idiopathic normal pressure hydrocephalus.** *Eur Radiol* 2008;18:2678–83 CrossRef Medline
15. Yun SY, Choi KS, Suh CH, et al. **Risk estimation for idiopathic normal-pressure hydrocephalus: development and validation of a brain morphometry–based nomogram**. *Eur Radiol* 2023;33:6145–56 CrossRef Medline
16. Miskin N, Patel H, Franceschi AM, et al; Alzheimer's Disease Neuroimaging Initiative. **Diagnosis of normal-pressure hydrocephalus: use of traditional measures in the era of volumetric MR imaging**. *Radiology* 2017;285:197–205 CrossRef Medline
17. Irie R, Otsuka Y, Hagiwara A, et al. **A novel deep learning approach with a 3D convolutional ladder network for differential diagnosis of idiopathic normal pressure hydrocephalus and Alzheimer's disease**. *Magn Reson Med Sci* 2020;19:351–58 CrossRef Medline
18. Borzage M, Saunders A, Hughes J, et al. **The first examination of diagnostic performance of automated measurement of the callosal angle in 1856 elderly patients and volunteers indicates that 12.4% of exams met the criteria for possible normal pressure hydrocephalus**. *AJNR Am J Neuroradiol* 2021;42:1942–48 CrossRef Medline
19. Cohen JF, Korevaar DA, Altman DG, et al. **STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration.** *BMJ Open* 2016;6:e012799 CrossRef Medline
20. Altman DG, Simera I, Hoey J, et al. **EQUATOR: reporting guidelines for health research.** *Lancet* 2008;371:1149–50 CrossRef Medline
21. Collins GS, Reitsma JB, Altman DG, et al. **Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement.** *Circulation* 2015;131:211–19 CrossRef Medline
22. McKhann GM, Knopman DS, Chertkow H, et al. **The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease**. *Alzheimers Dement* 2011;7:263–69 CrossRef Medline
23. Gibb W, Lees A. **The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease**. *J Neurol Neurosurg Psychiatry*. 1988;51:745–52 CrossRef Medline
24. Evans WA. **An encephalographic ratio for estimating ventricular enlargement and cerebral atrophy.** *Arch NeurPsych* 1942;47:931–37 CrossRef
25. Fischl B. **FreeSurfer.** *Neuroimage* 2012;62:774–81 CrossRef
26. Jenkinson M, Beckmann CF, Behrens TE, et al. **FSL.** *Neuroimage* 2012;62:782–90 CrossRef Medline
27. Fonov VS, Evans AC, McKinstry RC, et al. **Unbiased nonlinear average age-appropriate brain templates from birth to adulthood.** *Neuroimage* 2009;47:S102 CrossRef
28. Liu Z, Lin Y, Cao Y, et al. **Swin transformer: hierarchical vision transformer using shifted windows.** In: *Proceedings of the IEEE/CVF International Conference Computer Vision*. October 11–17, 2021:10012–22. Montreal, QC, Canada
29. Suh PS, Jung W, Suh CH, et al. **Development and validation of a deep learning-based automatic segmentation model for assessing intracranial volume: comparison with NeuroQuant, FreeSurfer, and SynthSeg.** *Front Neurol* 2023;14:1334962 CrossRef Medline
30. Ardekani BA, Bachman AH. **Model-based automatic detection of the anterior and posterior commissures on MRI scans.** *Neuroimage* 2009;46:677–82 CrossRef Medline
31. Bradski G. **The openCV library.** *Dr Dobb's* 2000;25:120–23. https://github.com/opencv/opencv/wiki/CiteOpenCV.
32. Pedregosa F, Varoquaux G, Gramfort A, et al. **Scikit-learn: machine learning in Python.** *The Journal of Machine Learning Research* 2011;12:2825–30
33. Vallat R. **Pingouin: statistics in Python.** *JOSS* 2018;3:1026 CrossRef
34. Chen T, Guestrin C. **Xgboost: ascalable tree boosting system.** In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. August 13–17, 2016: 785–94. New York
35. Pahwa B, Tayal A, Shukla A, et al. **Utility of machine learning in the management of normal pressure hydrocephalus (NPH): a systematic review.** *World Neurosurg* 2023 Jun 24 [Epub ahead of print] CrossRef Medline
36. Rau A, Kim S, Yang S, et al. **SVM-based normal pressure hydrocephalus detection.** *Clin Neuroradiol* 2021;31:1029–35 CrossRef Medline