Appendix E1

Case Selection

We identified potential patients with diseases involving deep gray matter structures from a search of institutional radiology archives (mPower, Nuance Communications, Burlington, Mass) of studies performed between January 2008 and January 2018 with search terms for each of the diagnostic entities and deep gray matter structures. Initial searches resulted in 348 potential patients (Fig 1). The diagnoses of potential patients were validated by consensus among two neuroradiologists (I.M.N., S.M.) through clinical chart review, pathology when available and analysis of other available imaging data, including follow-up imaging studies. Of these potential patients, 23 were excluded due to incorrect initial diagnoses (with 7 being reassigned to one of the other inclusion diagnoses) and 21 were excluded due to insufficient information to confirm diagnosis. To select the specific examination from each patient, the first diagnostic MRI performed was selected, with three being unavailable due to the MRI being performed prior to institutional review board (IRB)-approved date range. Potential cases were then further narrowed down by excluding cases without T1 imaging ($n = 24$; T1 images being required for the tissue segmentation and anatomic parcellation), severe motion artifacts ($n = 23$), cases with other secondary pathologic diagnoses other than a mild degree of chronic white matter small vessel ischemic disease ($n = 28$), and cases where the diagnostic findings were outside of the deep gray matter ($n = 21$), which resulted in a final sample of 212 cases (Table 1). An additional 178 cases were included for training of the CNNs from a related study of diseases involving the cerebral hemispheres (18), which initially consisted of 279 possible patients.

Two academic neuroradiologists (I.M.N., S.M.) classified the 36 diagnostic entities as 'common', 'moderately rare' or 'rare' in regard to the relative frequency in which they are diagnosed on brain MRI studies at our tertiary care center. The classification of prevalence was not specific to the diagnosis being present in deep gray matter (ie, if the disease was considered common but rarely found in deep gray matter, it was still considered a common disease).

Appendix E2

MRI Data Availability

Of the 212 deep gray matter cases, T1-weighted images were acquired in the axial plane in 86.7% of cases and in the sagittal plane in 13.3% of cases. Of the selected cases, 89.2% included axial T1 postcontrast, 98.6% axial FLAIR, 74.1% axial GRE, 99.1% DW, and 99.1% ADC images. Clinical MR images were collected from our health system PACS, which included outside studies submitted to our PACS for secondary interpretation between January 2008 and January 2018. The 390 MRI studies (212 deep gray matter cases and 178 used to supplement the CNNs) were obtained from 35 different physical scanners, 4 scanner manufacturers, and 16 unique scanner models (Table 2), noting that 97% of the data were acquired from Siemens (Munich, Germany) and GE Healthcare (Chicago, Ill) with fewer cases from Philips (Amsterdam, the Netherlands) and Toshiba Medical Systems (Otawara, Japan). There were 80

unique acquisition parameters for the T1 images alone (Table 2). For reference, the most commonly used acquisition parameters for the T1 image were as follows: repetition time (TR), 500 msec; echo time (TE), 17 msec; and in-plane resolution, 0.86 × 0.86 mm, with 5-mm thick slices, which was used in 25% of cases.

Appendix E3

Image Preprocessing, Registration, Tissue Segmentation, and Deep Gray Parcellation

We performed segmentation of brain tissues and deep gray matter structures on the T1 images using a customized image processing pipeline utilizing ANTs (Advanced Normalization Tools; http://picsl.upenn.edu/software/ants/) (20,21), FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl) (30), MATLAB (Mathworks, Natick, Mass), and customized Python scripts (Python version 3.7). First, the neck portion of the T1 MRI was removed using a custom algorithm based on Brain Extraction Tool (BET) (31). This step failed in 7% of the cases due to T1 acquisitions containing unexpectedly large nonbrain tissue in older protocols, requiring manual intervention. The T1 images were then up-sampled to reduce slice thickness to 1 mm using a multimodal patch-based superresolution technique (32). The ANTs cortical thickness analysis pipeline was then applied to segment the brain into cerebrospinal fluid, cerebral white matter, cortical and deep gray matter, cerebellum and brainstem. Then, the deep gray matter was further divided into eight subregions, consisting of left and right caudate, putamen, globus pallidus, and thalamus, by assigning each deep gray matter voxel to the closest subcortical label in the Automated Anatomic Labeling parcellation (33) that was warped to the space of each patient's up-sampled T1 scan. In addition, a mask of nearby structures including the ventricles, hypothalamus, optic chiasm and hippocampus in the OASIS template space (34) was warped to each patient and subtracted from the deep gray matter subregions to limit false positive extension of the deep gray matter into these structures. Morphometry correction was applied to generate the final segmentation of the subregions. Rigid registration (6 degrees of freedom) was performed to register each of the other modalities to the up-sampled T1 image of the same patient. In the end, the different tissue types and deep gray subregions were transformed and resampled to the native space of each MR sequence.

Appendix E4

Convolutional Neural Networks for Abnormal T1, FLAIR, and GRE Signal

For prediction of abnormal signal on T1, FLAIR and GRE MRI studies, we developed and trained customized 3D U-Nets (21–23) for each sequence, in a similar fashion as the FLAIR U-Net described in Duong et al (24) (Fig 3, *B*). To provide segmentation masks for the training data, all T1, FLAIR and GRE lesions were hand-segmented by a radiologist (J.D.R. or A.M.R.) using ITK-SNAP (http://www.itksnap.org) (35). The FLAIR U-Net was supplemented by an additional 178 hand-segmented FLAIR MRI studies (total $n = 387$) and the GRE U-Net was supplemented with an additional 92 segmented GRE MRI studies (total $n = 249$) of lesions of various etiologies involving the cerebral hemispheres (18).

Preprocessing for input into the CNNs included brain extraction, followed by intensity normalization by the mean and standard deviation. Next, images were resampled to 1 mm^3

isotropic resolution via linear interpolation. The images were then augmented using elastic transformations (36), including rotate, flip and skew, with small random affine transformations stacked on top of small random free-form deformations such that each imaging volume was augmented a total of three times. The augmented images were then split into 96 × 96 × 96-mm cubes ("3D patches") that were used as input into the CNN.

The network architecture was implemented with TensorFlow (37) (CUDA version 9.2.148) on an NVIDIA Titan Xp GPU (NVIDIA, Santa Clara, Calif; 12 GB memory). The 3D U-Net architecture consisted of four consecutive down-sampled blocks, followed by four consecutive up-sampled blocks (Fig 3, *B*). We used batch normalization for regularization and the rectified linear unit for nonlinearity. For down-sampling and up-sampling, the network used a stride-2 convolution and 2-stride deconvolution. A 3 × 3 × 3 convolutional kernel was applied across the network for each layer. In the down-sampling block, we applied a dilation factor of two in all convolutional layers. We applied a cross-link between corresponding up-sampling and down-sampling blocks as well as a residual connection between subsequent layers with number of features matched by a plain 1 × 1 × 1 convolution. After the final up-sampling block, three additional convolutional, rectified linear unit, batched-normalized (conv-ReLu-bn) layers were added, ending with a normalized exponential (softmax) head function. We utilized standard cross-entropy loss with an Adam optimizer (38) and a learning rate of 10^{-4}. A batch consisted of six patches. The networks were trained for 20–50 epochs over the course of 1–3 days, with training halted after training loss plateaued. No hyperparameter optimization was performed as training was only performed once for each network and the network weights were applied to the test cases only once. During training, the 3D patches were randomly sampled across the full brain volumes. To prevent sample imbalance, the number of patches that included abnormal signal was equal to the number that did not have abnormal signal. During testing, the brain volume was densely sampled with the cubes using a step size of 32 in each direction, resulting in a 64 pixel overlap between cubes. The overlapped segmentation predictions were then averaged.

Appendix E5

Abnormal Signal Detection for Enhancement and Restricted Diffusion

Detection of abnormal enhancement and restricted diffusion relied on analysis of the voxel intensities within deep gray matter subregions relative to the mean intensity of all deep and cortical gray matter voxels. To detect enhancement, voxels were required to be 2.5 standard deviations higher than mean signal on the subtraction map performed between the rigidly aligned T1 and T1-post images (Fig 3, *D*) and 1.5 standard deviations higher than the mean signal on the T1-post images. Voxels with restricted diffusion were detected by masking voxels with intensities that were 2.5 standard deviations above the mean gray matter intensity on the high *b* value DW image and 1.0 standard deviations below the mean gray matter intensity on the ADC images (Fig 3, *D*). These thresholds were based on minimizing false positives in the training patients.

Appendix E6

Thresholding of Abnormal Signal Maps in Deep Gray Subregions

To consider the signal derived from the U-Nets (T1, FLAIR, GRE) or voxel-wise methods (enhancement, diffusion) to be abnormal, modality- and subregion-specific thresholds were set for abnormal voxel percentage. The thresholds were chosen based on the training cases, such that they resulted in the highest accuracy for each feature within the training cases, and were then applied to the test cases. This fine-tuning of thresholds in the training sample allowed for the pipeline to be robust in the presence of a minimal amount of abnormal signal due to noise or normal variation, minimizing false positives. This process also resulted in ignoring abnormal GRE signal detected in the globi pallidi due to the frequency of physiologic mineral deposition resulting in reduced signal that would be considered abnormal in other regions of the brain. Thresholds varied across subregions between 2.0% and 4.5% of regional voxels for enhancement, 0.5% and 3.5% for restricted diffusion, 1.0% and 3.0% for increased or decreased on T1, 2.0% and 4.0% for increased FLAIR, and 3.0% and 3.5% for decreased on GRE. For GRE, the feature state "high" reflected cases where more than 20% of the voxels had abnormal susceptibility across all the deep gray structures excluding globus pallidus.

Appendix E7

Anatomic Subregion and Spatial Features

The four anatomic subregion features (caudate, putamen, globus pallidus, and thalamus) were considered to be involved if abnormal signal was present above the threshold for any of the modalities in either the right or the left side. The combined maps of abnormal signal across all modalities, masked by the deep gray subregions, were used to compute additional spatial pattern features, consisting of bilaterality and symmetry. Abnormalities were considered unilateral if there was a greater than 20-fold difference in abnormal signal volume between right- and left-sided subregions. Abnormalities were considered bilateral and symmetric if there was a less than fivefold difference in the ratio of abnormal signal detected in left versus right hemispheres within at least one set of paired deep gray subregions. These thresholds were determined based on optimizing performance of these features in the training sample relative to attending-derived reference standard features.

Appendix E8

Clinical Features

A review of electronic medical records was performed to obtain each patient's age, sex, acuity of the patient's clinical presentation and whether that patient was known to be immunocompromised at the time of imaging. These four clinical features were chosen based on their broad utility in helping develop a differential diagnosis for this set of diseases. Acute symptoms were defined as the predominant symptoms that necessitated the MRI occurring less than 7 days prior to imaging, subacute between 7 days and 3 months, and chronic and/or asymptomatic if the patient's predominant symptoms had been present for longer than 3 months or the MRI had been ordered based on further evaluation of an incidental finding or screening due to an underlying disease.

Appendix E9

Bayesian Network Construction and Analysis

The probabilities of each feature for each disease for the Bayesian network were determined by the consensus of four neuroradiologists and published statistics in a comprehensive neuroradiology textbook (39), when available. These probabilities were then tuned to approximate a weighted average of the expert-derived probabilities and frequency of feature states seen in the training sample. The full set of probabilities used for the analysis are displayed in Table E2. For the purposes of this experiment, the prior probabilities for each disease and normal were set to be equal to each other to approximate their relative frequency in the study population. The Bayesian network was implemented with custom Python scripts performing simple naïve Bayesian inference (https://github.com/rauscheck/radai). For instances where the feature could not be calculated, such as a missing sequence, the feature received an N/A and was not used as input to the Bayesian network.

Appendix E10

Attending-derived Reference Standard Imaging Features

To evaluate the performance of the image pipeline for detecting the presence of signal features (T1, FLAIR, susceptibility, enhancement, restricted diffusion), anatomic subregion (caudate, putamen, globi pallidi, thalami) and spatial features (bilateral, symmetric) we determined performance relative to the reference standard consensus of three radiologists (I.M.N., S.M., J.D.R.) evaluating each of the cases. Any discrepancies between the academic neuroradiologists regarding the imaging features and diagnoses were resolved through consensus after rereviewing the images. The prevalence of each of the features for each of the diagnostic entities is displayed in Table E1.

Appendix E11

Clinical Validation

To validate the performance of the AI system, four radiology residents, two neuroradiology fellows, two general radiologists, and two academic neuroradiology attending physicians reviewed the cases and provided their T3DDx from the 35 possible pathologic diagnoses plus normal. The academic neuroradiologists subsequently provided reference standard imaging features for each case to validate the performance of the image-processing portion of the pipeline. The test cases were copied into anonymized accession numbers and then displayed in a standard fashion in the clinical PACS (Sectra AB, Linköping, Sweden). The same four clinical features used in the Bayesian network were made available to each reader when viewing the cases, except that radiologists were provided with the exact age of the patient and the Bayesian network received a thresholded age as input. All radiologists were told that the frequency of all diagnoses was approximately equal in the study.

Appendix E12

Receiver Operating Characteristics

This ordinal scale of confidence for each diagnosis (top 1, top 2, top 3) was used to construct nonparametric ROC curves according to classic signal detection theory as applied to medical decision making (40), with the AUC used as a measure of criterion-independent performance. Nonparametric AUCs were calculated with 95% confidence intervals determined by 100 bootstrapping samples on this nonparametric ROC curve, using the MATLAB routine "paramROC" (http://www.mathworks.com/matlabcentral/fileexchange/39127-parametric-roc-curve). The significance of the difference between AUCs for different groups was tested using the DeLong test (25). The AUCs were averaged across each group for purposes of visualization in Figure 5, *C*[ID]FIG5[/ID].

References

30. Smith SM. Fast robust automated brain extraction. Hum Brain Mapp 2002;17(3):143–155https://doi.org/10.1002/hbm.10062.PubMed

31. Smith SM, Jenkinson M, Woolrich MW, et al. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 2004;23(Suppl 1):S208–S219https://doi.org/10.1016/j.neuroimage.2004.07.051.PubMed

32. Manjón JV, Coupé P, Buades A, Fonov V, Louis Collins D, Robles M. Non-local MRI upsampling. Med Image Anal 2010;14(6):784–792https://doi.org/10.1016/j.media.2010.05.010.PubMed

33. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 2002;15(1):273–289https://doi.org/10.1006/nimg.2001.0978.PubMed

34. Marcus DS, Fotenos AF, Csernansky JG, Morris JC, Buckner RL. Open access series of imaging studies: longitudinal MRI data in nondemented and demented older adults. J Cogn Neurosci 2010;22(12):2677–2684https://doi.org/10.1162/jocn.2009.21407.PubMed

35. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 2006;31(3):1116–1128https://doi.org/10.1016/j.neuroimage.2006.01.015.PubMed

36. Simard PY, Steinkraus D, Platt JC. Best practices for convolutional neural networks applied to visual document analysis. In: Proceedings of the Seventh International Conference on Document Analysis and Recognition, Edinburgh, UK, August 6, 2003. Piscataway, NJ: IEEE, 2003; 958–963https://doi.org/10.1109/ICDAR.2003.1227801.

37. Abadi M, Agarwal A, Barham P, et al. TensorFlow: Large-scale machine learning on heterogeneous systems. arXiv:1603.04467 [preprint]. https://arxiv.org/abs/1603.04467. Posted March 14, 2016. Accessed December 2019.

38. Kingma DP, Ba L. ADAM: A Method for Stochastic Optimization. Conference Paper at International Conference for Learning Representations, San Diego, 2015.

39. Osborn AG. Osborn's brain: imaging, pathology, and anatomy. Salt Lake City, Utah: Amirsys, 2013.

40. Obuchowski NA. ROC analysis. AJR Am J Roentgenol 2005;184(2):364–372https://doi.org/10.2214/ajr.184.2.01840364.PubMed

Table E1. Prevalence of imaging features for each of the diseases included in the study.

Disease	Chronicity	Immun.	T1	T2-FLAIR	R Diffusion	Suscept	Enhance	Caudate	Putamen	GP	Thala-mus	Bilateral	Symmet-ric
Abscess	100% acute	50%	50% inc, 100% dec	100% inc, 50% dec	100%	100%	100%	100%	100%	100%	50%	50%	0%
Anoxic Brain Injury: Acute	100% acute	0%	0% inc, 0% dec	100% inc, 0% dec	86%	0%	0%	100%	100%	63%	29%	100%	100%
Anoxic Brain Injury: Chronic	100% chronic	0%	100% inc, 50% dec	100% inc, 0% dec	0%	50%	0%	100%	100%	50%	50%	100%	100%
Anoxic Brain Injury: Subacute	100% subacute	0%	0% inc, 0% dec	100% inc, 0% dec	75%	0%	25%	100%	100%	75%	50%	100%	100%
Artery of Percheron Infarct	100% acute	0%	0% inc, 100% dec	100% inc, 0% dec	100%	0%	0%	0%	0%	0%	100%	100%	0%
Bilateral Thalamic Glioma	100% subacute	0%	0% inc, 100% dec	100% inc, 0% dec	0%	0%	10%	50%	0%	0%	100%	100%	50%
Calcium Deposition/ Fahr's disease	100% chronic	0%	100% inc, 29% dec	71% inc, 29% dec	0%	100%	0%	71%	86%	100%	71%	100%	100%
Carbon Monoxide: Acute	100% acute	0%	0% inc, 50% dec	100% inc, 0% dec	100%	0%	0%	0%	0%	100%	0%	100%	100%
Carbon Monoxide: Chronic	100% chronic	0%	100% inc, 0% dec	100% inc, 50% dec	0%	50%	0%	0%	0%	100%	0%	100%	100%
Carbon Monoxide: Subacute	100% subacute	0%	50% inc, 0% dec	100% inc, 0% dec	66%	100%	100%	0%	0%	100%	0%	100%	75%
Central Nervous System Lymphoma	9% acute, 82% sub, 9% chronic	27%	27% inc, 100% dec	100% inc, 0% dec	91%	0%	100%	82%	91%	91%	91%	55%	0%
Creutzfeldt-Jakob Disease	64% sub, 36% chronic	0%	0% inc, 18% dec	100% inc, 0% dec	100%	0%	0%	100%	100%	0%	73%	100%	91%
Cryptococcus	100% subacute	100%	50% inc, 50% dec	100% inc, 0% dec	100%	0%	100%	100%	100%	100%	100%	100%	50%
Deep Vein Thrombosis: Acute	100% acute	0%	0% inc, 100% dec	100% inc, 0% dec	50%	100%	0%	100%	100%	100%	100%	50%	0%
Deep Vein Thrombosis: Chronic	100% chronic	0%	67% inc, 33% dec	67% inc, 100% dec	0%	100%	0%	67%	33%	67%	33%	67%	0%
Deep Vein Thrombosis: Subacute	100% subacute	33%	33% inc, 100% dec	100% inc, 0% dec	67%	100%	33%	100%	100%	100%	100%	67%	33%
Encephalitis	33% acute, 67% subacute	50%	25% inc, 50% dec	100% inc, 0% dec	25%	0%	25%	100%	100%	100%	75%	100%	75%
Glioma: High Grade	18% acute, 82% subacute	0%	9% inc, 100% dec	100% inc, 27% dec	73%	71%	91%	9%	45%	55%	100%	0%	0%
Glioma: Low Grade	14% sub, 86% chronic	0%	0% inc, 71% dec	100% inc, 0% dec	0%	0%	29%	14%	29%	14%	100%	0%	0%
Hemorrhage: Acute	100% acute	0%	100% inc, 100% dec	100% inc, 100% dec	88%	100%	0%	88%	100%	100%	88%	25%	0%
Hemorrhage: Chronic	100% chronic	0%	83% inc, 100% dec	83% inc, 83% dec	0%	100%	0%	67%	100%	67%	50%	0%	0%

Disease												
Hemorrhage: Subacute	100% subacute	0%	100% inc, 50% dec	100% inc, 83% dec	67%	100%	0%	67%	50%	67%	67%	33%	0%
Infarct: Acute	100% acute	0%	10% inc, 70% dec	100% inc, 0% dec	100%	22%	0%	50%	50%	40%	40%	0%	0%
Infarct: Chronic	100% chronic	13%	25% inc, 88% dec	100% inc, 50% dec	0%	80%	0%	88%	75%	38%	13%	0%	0%
Infarct: Subacute	20% acute, 80% subacute	0%	70% inc, 70% dec	100% inc, 0% dec	70%	20%	90%	50%	60%	60%	30%	0%	0%
Manganese Deposition	22% acute, 22% sub, 56% chronic	33%	100% inc, 0% dec	0% inc, 0% dec	0%	11%	0%	0%	0%	100%	11%	100%	100%
Metastasis	36% acute, 18% sub, 45% chronic	45%	45% inc, 82% dec	100% inc, 0% dec	45%	38%	100%	55%	100%	55%	91%	91%	0%
Neuro Behçet's Disease	100% subacute	100%	0% inc, 100% dec	100% inc, 0% dec	0%	0%	0%	50%	50%	50%	100%	50%	0%
Neurofibromatosis type 1	100% chronic	0%	0% inc, 0% dec	100% inc, 0% dec	0%	0%	0%	0%	0%	100%	50%	0%	0%
Neurosarcoidosis	17% acute, 83% subacute	0%	0% inc, 50% dec	100% inc, 0% dec	17%	0%	83%	67%	67%	83%	100%	83%	17%
Nonketotic Hyperglycemia	50% acute, 50% subacute	0%	100% inc, 0% dec	0% inc, 0% dec	0%	100%	0%	50%	100%	50%	0%	100%	50%
Seizures	100% acute	0%	0% inc, 0% dec	100% inc, 0% dec	100%	0%	0%	100%	100%	0%	50%	100%	50%
Toxoplasmosis	44% acute, 56% subacute	82%	45% inc, 64% dec	100% inc, 9% dec	73%	80%	91%	73%	91%	100%	82%	55%	0%
Wernicke Encephalopathy	100% acute	0%	0% inc, 0% dec	100% inc, 0% dec	63%	0%	14%	0%	0%	0%	100%	100%	88%
Wilson's Disease	50% subacute, 50% chronic	0%	100% inc, 0% dec	50% inc, 0% dec	100%	0%	50%	50%	50%	100%	100%	100%	100%
Normal	33% acute, 33% sub, 33% chronic	10%	0% inc, 0% dec	0% inc, 0% dec	0%	0%	0%	0%	0%	0%	0%	0%	0%

Note.—dec = decreased, GP = globus pallidus, Immun = immunocompromised, inc = increased, R Diffusion = reduced diffusion, sub = subacute, Suscept = susceptibility, T1 = T1-weighted, T2-FLAIR = T2-weighted fluid-attenuated inversion recovery.

Table E2. Expert consensus probabilities (%) of imaging and clinical features for the 36 diagnostic entities included in the study.

	Sex		Age		Chronicity			Immunocompromised		T1_Increased		T1_Decreased		T2_FLAIR_Increased		T2_FLAIR_Decreased	
Disease	Male	Female	Young (<40)	Old (≥40)	Acute	Sub-acute	Chron-ic	Yes	No	Increased	Normal	Decreased	Normal	Increased	Normal	Decreased	Normal
Normal	50	50	50	50	34	33	33	10	90	10	90	10	90	10	90	10	90
Abscess	50	50	30	70	80	19	1	50	50	60	40	60	40	85	15	30	70
Artery_of_Percheron_acute	50	50	10	90	95	4	1	10	90	15	85	60	40	89	11	5	95

Bilateral_Thalamic_Glioma	50	50	50	50	5	60	35	10	90	5	95	68	32	78	22	5	95
Calcium_Deposition	50	50	20	80	5	15	80	10	90	75	25	25	75	50	50	30	70
Carbon_Monoxide_acute	50	50	40	60	95	4	1	10	90	10	90	45	55	89	11	11	89
Carbon_Monoxide_chronic	50	50	40	60	1	4	95	10	90	30	70	60	40	79	21	21	79
Carbon_Monoxide_subacute	50	50	40	60	30	69	1	10	90	70	30	30	70	89	11	11	89
Creutzfeldt_Jakob	50	50	5	95	1	60	39	10	90	5	95	47	53	75	25	10	90
Cryptococcus	50	50	50	50	30	65	5	95	5	20	80	60	40	85	15	6	94
Deep_Vein_Thrombosis_acute	25	75	30	70	95	4	1	10	90	25	75	60	40	90	10	15	85
Deep_Vein_Thrombosis_chronic	25	75	30	70	1	4	95	10	90	75	25	35	65	50	50	50	50
Deep_Vein_Thrombosis_subacute	25	75	30	70	2	95	3	10	90	50	50	60	40	85	15	10	90
Encephalitis	50	50	50	50	30	69	1	50	50	30	70	50	50	80	20	20	80
Glioma_High_Grade	70	30	30	70	20	60	20	10	90	25	75	85	15	90	10	21	79
Glioma_Low_Grade	50	50	50	50	5	15	80	10	90	10	90	50	50	89	11	5	95
Hemorrhage_acute	60	40	20	80	95	4	1	10	90	40	60	50	50	80	20	30	70
Hemorrhage_chronic	60	40	20	80	1	4	95	10	90	50	50	50	50	50	50	50	50
Hemorrhage_subacute	60	40	20	80	2	95	3	10	90	80	20	30	70	70	30	40	60
Hypoxic_Ischemic_Encephalopathy_acute	60	40	50	50	95	4	1	10	90	10	90	40	60	85	15	10	90
Hypoxic_Ischemic_Encephalopathy_chronic	60	40	50	50	1	4	95	10	90	75	25	35	65	85	15	10	90
Hypoxic_Ischemic_Encephalopathy_subacute	60	40	50	50	2	95	3	10	90	25	75	25	75	85	15	10	90
Infarct_acute	50	50	10	90	95	4	1	10	90	10	90	50	50	83	17	5	95
Infarct_chronic	50	50	10	90	1	4	95	10	90	35	65	55	45	60	40	45	55
Infarct_subacute	50	50	10	90	30	69	1	10	90	53	47	47	53	83	17	5	95
Lymphoma	50	50	20	80	15	70	15	40	60	25	75	75	25	85	15	30	70
Manganese_Deposition	50	50	10	90	20	20	60	30	70	94	6	5	95	20	80	20	80
Metastases	50	50	10	90	30	20	50	50	50	40	60	70	30	85	15	15	85
Neuro_Behcet_Disease	50	50	50	50	10	80	10	75	25	10	90	50	50	89	11	5	95
Neurofibromatosis1	50	50	90	10	10	10	80	10	90	15	85	10	90	84	16	5	95
Nonketotic_Hyperglycemia	50	50	20	80	45	45	10	10	90	90	10	10	90	10	90	10	90
Sarcoidosis	20	80	50	50	20	75	5	10	90	10	90	50	50	90	10	10	90
Seizures	50	50	50	50	90	9	1	10	90	10	90	25	75	85	15	10	90
Toxoplasmosis	50	50	50	50	48	48	4	80	20	50	50	60	40	80	20	20	80
Wernicke_Encephalopathy	50	50	20	80	95	4	1	10	90	5	95	7	93	85	15	10	90
Wilson_Disease	50	50	70	30	5	45	50	10	90	90	10	10	90	50	50	10	90

Disease	Enhancement Yes	Enhancement No	Diffusion Yes	Diffusion No	Susceptibility Yes	Susceptibility High	Susceptibility No	Bilateral Yes	Bilateral No	Symmetric Yes	Symmetric No	Caudate Yes	Caudate No	Globus_Pallidus Yes	Globus_Pallidus No	Putamen Yes	Putamen No	Thalamus Yes	Thalamus No
Normal	10	90	10	90	4	1	95	10	90	10	90	5	95	5	95	5	95	5	95
Abscess	95	5	95	5	40	20	40	30	70	5	95	50	50	50	50	50	50	50	50
Artery_of_Percheron_acute	10	90	80	20	10	0	90	95	5	50	50	5	95	5	95	5	95	95	5
Bilateral_Thalamic_Glioma	50	50	10	90	10	5	85	95	5	30	70	5	95	10	90	5	95	95	5
Calcium_Deposition	5	95	5	95	50	25	25	90	10	80	20	60	40	70	30	70	30	70	30
Carbon_Monoxide_acute	30	70	90	10	9	1	90	80	20	70	30	10	90	95	5	30	70	10	90
Carbon_Monoxide_chronic	5	95	5	95	40	10	50	80	20	70	30	40	60	95	5	40	60	25	75
Carbon_Monoxide_subacute	80	20	10	90	30	20	50	80	20	70	30	20	80	95	5	40	60	20	80
Creutzfeldt_Jakob	5	95	95	5	19	1	80	70	30	60	40	90	10	30	70	90	10	70	30
Cryptococcus	90	10	50	50	20	15	65	90	10	30	70	70	30	70	30	70	30	70	30
Deep_Vein_Thrombosis_acute	10	90	50	50	50	15	30	80	20	30	70	50	50	60	40	60	40	90	10
Deep_Vein_Thrombosis_ chronic	5	95	5	95	50	20	30	80	20	30	70	50	50	60	40	60	40	90	10
Deep_Vein_Thrombosis_ subacute	50	50	50	50	50	20	30	80	20	30	70	50	50	60	40	60	40	90	10
Encephalitis	30	70	35	65	15	10	75	80	20	50	50	80	20	80	20	80	20	60	40
Glioma_High_Grade	80	20	40	60	40	10	50	30	70	20	80	20	80	50	50	50	50	80	20
Glioma_Low_Grade	30	70	30	70	9	1	90	20	80	5	95	20	80	30	70	30	70	80	20
Hemorrhage_acute	1	99	60	40	20	79	1	30	70	10	90	60	40	80	20	80	20	80	20
Hemorrhage_chronic	1	99	5	95	69	30	1	20	80	10	90	60	40	80	20	80	20	80	20
Hemorrhage_subacute	20	80	50	50	39	60	1	20	80	10	90	60	40	80	20	80	20	80	20
Hypoxic_Ischemic_Encephalopathy_acute	5	95	70	30	9	1	90	80	20	60	40	80	20	60	40	90	10	50	50
Hypoxic_Ischemic_Encephalopathy_chronic	5	95	5	95	20	10	70	80	20	60	40	80	20	60	40	90	10	50	50
Hypoxic_Ischemic_Encephalopathy_subacute	5	95	70	30	15	5	80	80	20	60	40	80	20	60	40	90	10	50	50
Infarct_acute	20	80	85	15	24	1	75	20	80	10	90	50	50	50	50	50	50	50	50
Infarct_chronic	5	95	5	95	35	15	50	20	80	10	90	50	50	50	50	50	50	50	50
Infarct_subacute	80	20	40	60	45	5	50	20	80	20	80	50	50	50	50	50	50	50	50
Lymphoma	90	10	60	40	9	1	90	50	50	35	65	60	40	60	40	60	40	60	40
Manganese_Deposition	5	95	5	95	20	10	70	90	10	90	10	10	90	95	5	30	70	50	50
Metastases	80	20	50	50	45	5	50	60	40	30	70	50	50	60	40	60	40	60	40
Neuro_Behçet_Disease	20	80	20	80	9	1	90	60	40	35	65	50	50	50	50	50	50	75	25
Neurofibromatosis1	5	95	5	95	9	1	90	50	50	50	50	50	50	70	30	50	50	50	50
Nonketotic_Hyperglycemia	10	90	5	95	45	5	50	70	30	50	50	50	50	50	50	90	10	10	90
Sarcoidosis	90	10	5	95	9	1	90	90	10	30	70	60	40	60	40	60	40	70	30

Seizures	5	95	90	10	4	1	95	70	30	50	50	60	40	50	50	60	40	70	30
Toxoplasmosis	80	20	50	50	40	25	35	60	40	10	90	60	40	70	30	60	40	60	40
Wernicke_Encephalopathy	30	70	40	60	9	1	90	90	10	90	10	10	90	10	90	10	90	90	10
Wilson_Disease	30	70	60	40	20	10	70	90	10	90	10	50	50	90	10	70	30	70	30

Note.—T1 = T1-weighted, T2-FLAIR = T2-weighted fluid-attenuated inversion recovery.

Table E3. Accuracy for correct top 3 differential diagnosis for each of the 36 diagnostic entities.

Disease	Prevalence	Residents	General Rads	Neuro Fellows	Acad Attend	Automated
High Grade Glioma	Common	83%	67%	100%	100%	100%
Low Grade Glioma	Common	75%	33%	83%	83%	100%
Hemorrhage: Chronic	Common	100%	67%	83%	100%	67%
Infarct: Acute	Common	100%	83%	100%	100%	100%
Infarct: Subacute	Common	58%	67%	100%	100%	100%
Infarct: Chronic	Common	75%	100%	100%	100%	100%
Central Nervous System Lymphoma	Common	58%	67%	67%	100%	67%
Manganese Deposition	Common	50%	33%	67%	100%	100%
Metastasis	Common	67%	67%	67%	100%	100%
Normal	Common	100%	65%	95%	90%	100%
Abscess	Moderate	50%	100%	50%	75%	0%
Calcium Depositon/Fahr's disease	Moderate	73%	83%	83%	100%	100%
Creutzfeld-Jacob Disease	Moderate	50%	17%	67%	83%	100%
Hemorrhage: Acute	Moderate	83%	50%	100%	100%	100%
Hemorrhage: Subacute	Moderate	67%	67%	83%	83%	100%
Anoxic Brain Injury: Acute	Moderate	75%	100%	83%	100%	100%
Anoxic Brain Injury: Subacute	Moderate	42%	50%	33%	83%	67%
Toxoplasmosis	Moderate	42%	50%	17%	100%	67%
Wernickes Encephalopathy	Moderate	58%	25%	100%	100%	100%
Artery of Percheron Infarct	Rare	75%	75%	100%	100%	100%
Bilateral Thalamic Glioma	Rare	75%	50%	100%	100%	50%
Carbon Monoxide: Acute	Rare	75%	67%	67%	100%	100%
Carbon Monoxide: Subacute	Rare	25%	80%	100%	100%	33%
Carbon Monoxide: Chronic	Rare	38%	75%	100%	100%	100%
Cryptococcus	Rare	25%	25%	75%	75%	50%
Deep Vein Thrombosis: Acute	Rare	13%	25%	50%	50%	100%
Deep Vein Thrombosis: Subacute	Rare	17%	33%	33%	33%	100%

Deep Vein Thrombosis: Chronic	Rare	42%	17%	17%	33%	100%
Encephalitis	Rare	8%	0%	83%	33%	33%
Anoxic Brain Injury: Chronic	Rare	50%	75%	0%	75%	100%
Neuro Behçet's Disease	Rare	38%	50%	0%	25%	100%
Neurofibromatosis type 1	Rare	13%	0%	25%	50%	100%
Nonketotic Hyperglycemia	Rare	0%	25%	100%	100%	100%
Neurosarcoidosis	Rare	42%	0%	33%	83%	67%
Seizures	Rare	13%	0%	75%	75%	50%
Wilson's Disease	Rare	0%	25%	75%	75%	50%