Supplemental Methods

Automatic ischemic lesion segmentation

To train this lesion segmentation model, a subset of 2822 randomly selected scans was manually labeled by two neurologists and one radiologist each with more than 5 years of experience in interpreting DWI and ADC scans. The acute ischemic lesions were defined as high signal intensities on DWI, with corresponding low signal intensities on ADC maps. Each scan was assessed by two different labelers. Inconsistent labels were reviewed by two senior stroke neurologists. Final adjudications were defined as the ground truth.

Next, A four leveled U-Net[1] segmentation model (Supplemental Figure IIB) was trained using the 2822 patients' manually labeled data with 75% of data being used for model training and the remaining 25% being reserved for segmentation performance evaluation. The model was designed to predict a 2D lesion mask using 2D axial slices of ADC and DWI scans provided as a dual-channel input. The DWI and ADC scans were first preprocessed by scaling the global (3D) image intensities to follow a standard normal distribution (mean of 0, and standard deviation of 1). Next, the scans were spatially resampled to have uniform axial dimensions of 256x256 pixels using bilinear interpolation and were center cropped to 224x224 pixels to remove the redundant background. The model was trained using the preprocessed 2D axial slices of ADC and DWI scans provided as a dual-channel input (input shape: 224x224x2) with an Adam optimizer for 100 epochs using a cross-entropy loss and a batch size of 32. The initial learning rate was set to $1x10^{-4}$ and it was reduced by $1x10^{-6}$ every epoch. After the completion of training, the model epoch with the lowest

loss on the internal validation set (10% of the training data) was selected as a final segmentation model. This model was then used for automatic lesion segmentation in the entire dataset. The complete 3D lesion mask for each patient was computed by concatenating the 2D lesion masks from all the axial slices. Following training, the model's segmentation accuracy was tested using the 25% reserved test data.

After ensuring good lesion segmentation performance based on a dice coefficient, the trained U-net model was used to predict the infarction lesion mask for all the patients. Furthermore, using the predicted lesion mask, the total number of infarct lesions was computed by counting the number of 3D disconnected lesion segments. Also, the infarct volume was obtained by multiplying the number of affected voxels in the predicted lesion map by the single voxel volume[2].

Identification of patients with brainstem infarction

Following the automatic segmentation of lesions, to identify the patients with brainstem infarction, the DWI b1000 scans were registered to the MNI152-T1-non-linear-asymmetric-1mm-isotropic template[3], using affine and non-rigid registration techniques by NiftyReg library[4]. Next, the automatically segmented lesion masks were spatially registered to the same atlas using the derived transformation parameters. Finally, using the registered masks, depending on the infarction lesion location, the patients with lesions only in the brainstem region were identified and were considered to investigate the relationship between functional, clinical, demographic, and neuroimaging features. The remaining patients' data (patients with lesions in regions other than brainstem) were used to develop a deep learning based

feature extraction model to extract abstract and compact neuroimaging features that influence the functional outcomes 3 months post-stroke. (Supplemental Figure I)

Deep neuroimaging features extraction

The model was designed to accept 3D, template registered, DWI scan and predicted ischemic lesion mask of the patient as an input, and produced a 512 dimensional feature vector that encoded abstract neuroimaging features as the output. As a preprocessing step, the registered DWI scans were standard normalized to have global (3D) image intensity mean of 0 and a standard deviation of 1. Next, the DWI scan and the lesion mask were center cropped to 3D dimensions of 160x200x160 pixels and were provided as a dual-channel input to a 3D ResNet model (input shape: 160x200x160x2).

The ResNet model used multiple 3D convolutional blocks with residual connections to continuously extract local and global contextual features, and global average pooling layers to fuse global high-dimensional image features. Lastly, a fully connected layer followed by a 2-unit SoftMax activation function was applied, giving the model's predictions in the form of probability of a patient having favorable or unfavorable functional outcomes at 3 months post-stroke.

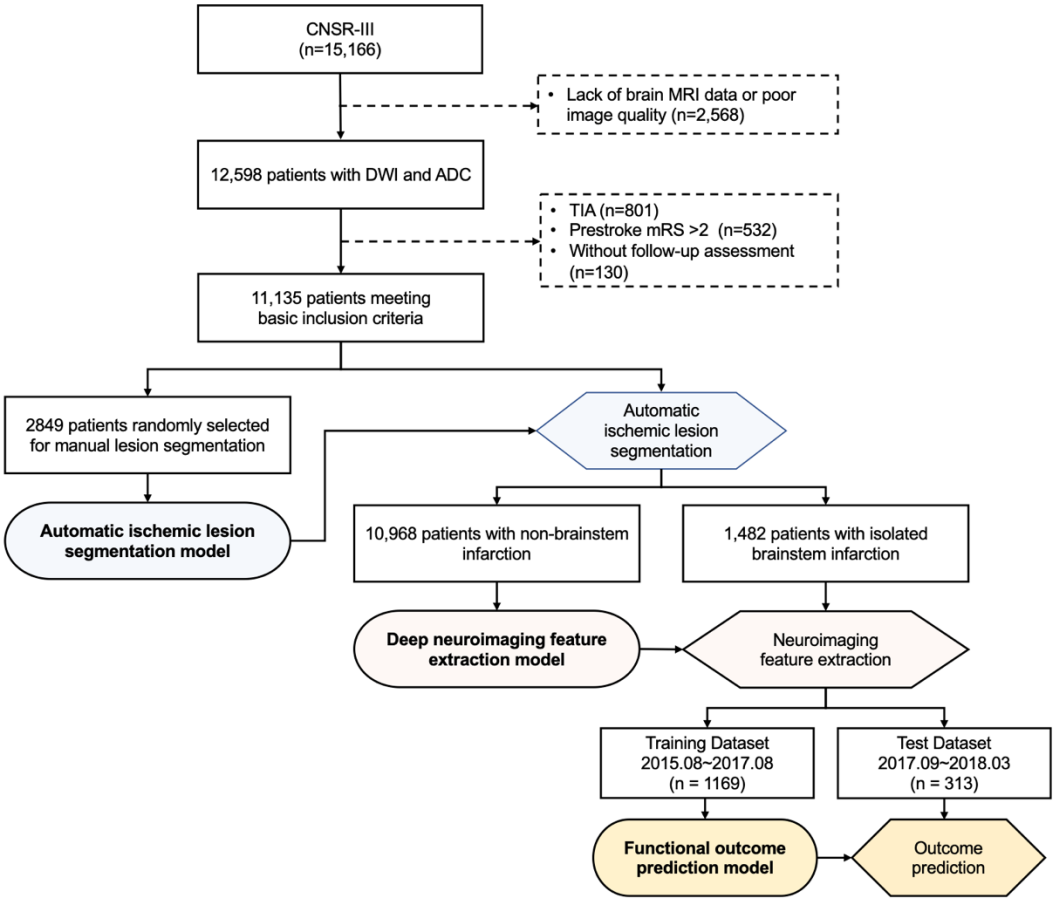
Next, the model was trained in an end-to-end manner using the imaging data from patients without brainstem infarction (n = 10,968) with Adam optimizer and focal loss for 100 epochs. During the training process, 10% of the training data was reserved as a validation set and the model with the lowest loss on the same was selected for

predictions on the test set. Also, Mix-up, a convex combination example-label pairs, was adopted in the training process to reduce the memorization of corrupt labels, increasing the robustness of the model ($\alpha = 0.6$) [5]. The training batch size was set to 16 and the L2 weight regularization was applied during the training process. The learning rate was reduced by half every 5 epochs. The value of initial learning rate (lr_init), and L2 regularization constant (l2) were optimized using a 3-fold cross-validation analysis with grid search ($\mathrm{lr\_init},\ \mathrm{l2} \subset \{\{1, 2, 5, 8\} \times \{10^{-3},\ 10^{-4},\ 10^{-5}\}\}$). Following the grid search, the top models with the highest area under the receiver operating curve (AUC) on their respective test folds were selected as final feature extraction models. In this manner, with the hyperparameter optimization, three feature extraction models (same architecture, different training hyperparameters) were trained using the data from patients without brainstem infarction and no data from patients with brainstem infarction was used in this training. Ultimately, using these trained feature extraction models, the 1536 (512 features/model * 3 models) features after the global average pooling layer were extracted as deep neuroimaging features for each patient. (Supplemental Figure II)

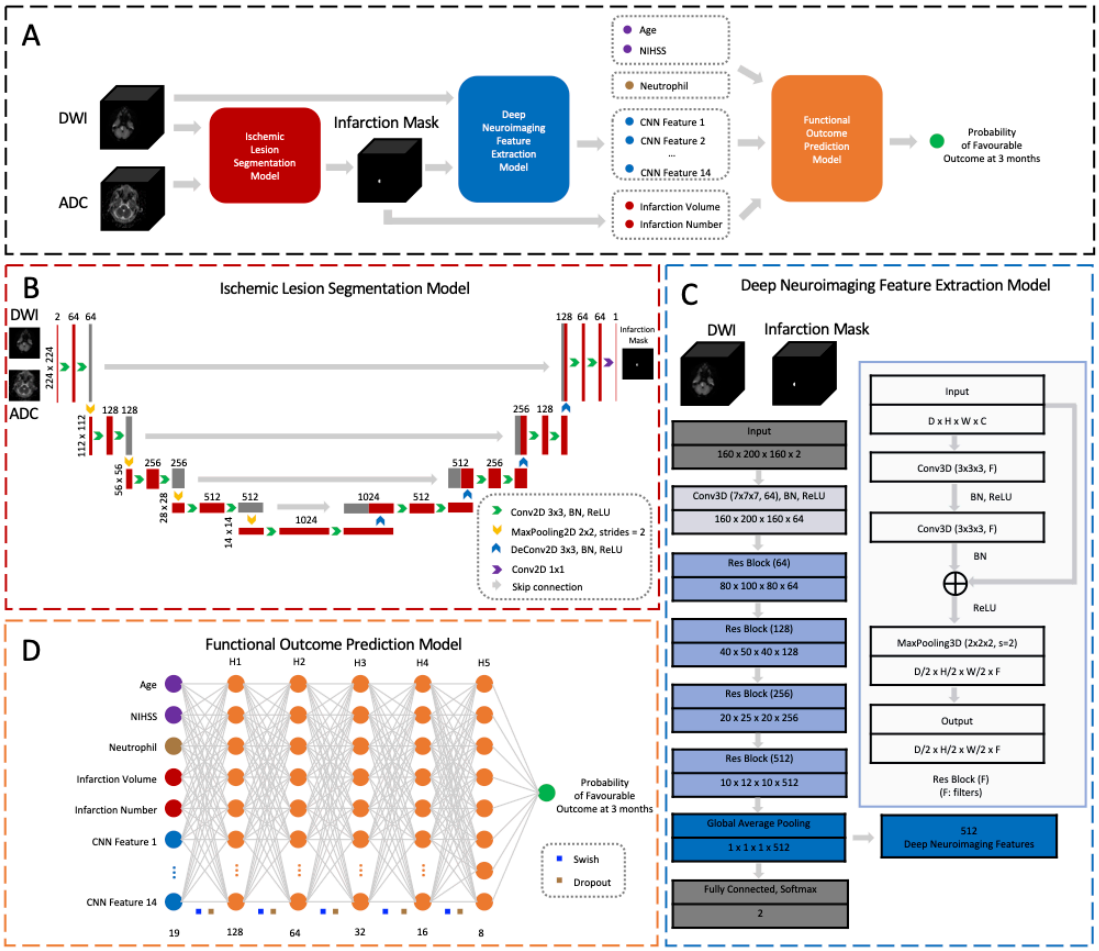
Supplemental Figures

Supplemental Figure I. Flowchart of patients included in the study.

CNSR-III = Third China National Stroke Registry; MRI = Magnetic resonance imaging; ADC = Apparent diffusion coefficient; DWI = Diffusion weighted imaging; TIA = Transient ischemic attacks; mRS = modified Rankin Scale.



Supplemental Figure II. The model architecture. (A) Study flow of prediction model. (B) U-Net for ischemic lesion segmentation. (C) Deep neuroimaging feature extraction model. (D) Functional outcome prediction model. DWI = Diffusion weighted imaging; ADC = Apparent diffusion coefficient; NIHSS = National Institutes of Health Stroke Scale.



Reference

1. Ronneberger Olaf FP, Brox Thomas. U-net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention – MICCAI. 2015:234-241
2. Zwanenburg A, Vallieres M, Abdalah MA, Aerts H, Andrearczyk V, Apte A, Ashrafinia S, Bakas S, Beukinga RJ, Boellaard R, et al. The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology. 2020;295:328-338
3. Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL, Brain Development Cooperative G. Unbiased average age-appropriate atlases for pediatric studies. Neuroimage. 2011;54:313-327
4. Modat M, Ridgway GR, Taylor ZA, Lehmann M, Barnes J, Hawkes DJ, Fox NC, Ourselin S. Fast free-form deformation using graphics processing units. Comput Methods Programs Biomed. 2010;98:278-284

5. Zhang H, Cisse M, Dauphin Y, Lopez-Paz D. Mixup: Beyond empirical risk minimization. Proceedings of the IEEE International Conference on Computer Vision (ICCV). 2017