**Supplementary material**

**Materials and Methods**

**Radiomics model building and DLCRN construction**

Considering the characteristics of collected data, we adopted the following principles to construct the models: (a) MRI images were shown in 3D formulation, and provided critical visual-spatial information from a practical point of radiology; (b) the models were parameter-efficient and easy to train with limited data; (c) the training strategy aimed at reducing over-fitting and highlighting the efficiency of data.

According to these principles, we adopted 3D ResNets derived from powerful deep convolutional neural networks 2D ResNets. ResNets have succeeded in 2D raw images, which elegantly use skip connection to avoid vanishing and exploding gradients, and efficiently extract high-level feature representation. Our 3D ResNets used the same topological structure but were 3D convolutions. The constructed 3D ResNet contained a convolutional layer, four 3D ResBlocks, a global pooling, and a fully-connected layer. The hyper-parameters of 3D ResNets were weight decay 5×10-4, momentum 0.9, and initial learning rate 10-2. When the validation error stopped diminution, the initial rate was one-tenth. In the training phase, we proposed a two-step training strategy. Firstly, the 3D ResNets, with U-net shape model, which formed an encode-decode image-segmentation algorithm in a pixel to pixel way, were pre-trained on BraTS2017. After the convergence of the pre-trained U-net, we removed the decoding architecture, which we called “backbone”, added a global pooling and a fully-connected layer to predict the HIE, and maintained the topology and weights of encoding network. For each neonate, the preprocessed MRI were fed to 3D ResNets. In the training phase, the data augment methods, including random flip, random scale, and Gaussian blur, were applied to avoid overfitting. When the input images were resized to 160×160×80 with trilinear interpolation and fed to the well-trained 3D ResNet classifiers, the probabilities for each sequence were generated to evaluate the risk of HIE. In this research, the probabilities called deep radiomics signatures correspond to “RadScore” in traditional radiomics and are considered as a quantitative assessment of radiology.

Software and system tools involved in the Methods were summarized here. Python (version 3.7.12) and PyTorch (version 1.6.0, https://pytorch.org/) were used to constructe the deep radiomics signatures and build the neural network. FSL (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/), SimpleITK (version 1.2.4), and Numpy (version 1.21.5) were used for MRI preprocessing and data augment. The deep radiomics signatures were developed and validated in the condition of Ubuntu (version 20.0), GTX1080Ti, and CUDA (version 11.4). The code of the constructed model will be released in the future.

**Supplementary Table 1.** Scanning protocols and parameters of brain magnetic resonance imaging examination from different manufacturers.

|  |  |  |
| --- | --- | --- |
| Center | Tongji Hospital | Maternal and Child Hospital of Hubei Province |
| Manufacturers | **SIEMENS** | **GE** | **United Imaging** | **SIEMENS** |
| Field strength | 1.5T | 1.5/3.0T | 3.0T | 1.5T |
| Sequences | T1WI | T2WI | T1WI | T2WI | T1WI | T2WI | T1WI | T2WI |
| Parameters |  |
| TR | 307-575 | 4500 | 360-735 | 4300 | 374 | 4618 | 614 | 2500 |
| TE | 6.9-22 | 98 | 6.5-19 | 129.4 | 8.7 | 101.1 | 22 | 126 |
| Flip angel | 120-150 | 120 | 90-111 | 90 | 115 | 125 | 90 | 150 |
| Slice thickness(mm) | 4-5 | 5 | 4-6 | 5 | 5 | 5 | 5 | 5 |
| Slice spacing(mm) | 4.4-6 | 5.1 | 4.5-7 | 5 | 5.5 | 5 | 6 | 6 |
| Head coil | 16 | 32 | 32 | 16 |

Note: T1WI = T1-weighted images; T2WI = T2-weighted images; TE = echo time; TR = repetition time.

**Supplementary Table 2.** Multivariate logistic regression results.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **B** | **SE** | **P value** |
| **Sex** | 0.09 | 0.41 | 0.83 |
| **Birthweight** | -2.26 | 0.63 | <0.01 |
| **Gestational age** | 0.25 | 0.17 | 0.16 |
| **Age** | -0.02 | 0.03 | 0.45 |
| **Apgar score** | -0.15 | 0.14 | 0.30 |
| **Neonatal asphyxia** | 21.51 | 8613.53 | 0.99 |
| **Fetal distress** | 20.87 | 11422.24 | 0.99 |
| **Dystonia** | -0.01 | 0.53 | 0.99 |
| **Abnormal sucking reflex** | 18.69 | 8461.93 | 0.99 |
| **Abnormal startle reflex** | -18.73 | 8461.93 | 0.99 |
| **T1 score** | 12.19 | 3.34 | <0.01 |
| **T2 score** | 9.32 | 3.06 | <0.01 |

**Supplementary Table 3.** Comparisons of the DLCRN model across manufacturers with

 different field strengths.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Field strength | Manufactures | AUC | 95%CI | Delong test |
| Training cohort |
| 3.0T | GE/United Imaging | 0.791 | 0.721-0.850 | 0.726 |
| 1.5T | GE/SIEMENS | 0.766 | 0.612-0.881 |
| Internal validation cohort |
| 3.0T | GE/United Imaging | 0.813 | 0.665-0.915 | 0.748 |
| 1.5T | GE/SIEMENS | 0.750 | 0.373-0.962 |

Note: AUC = area under receiver operating characteristics; CI = confidence interval; DLCRN = deep learning clinical-radiomics nomogram.

**Supplementary Figure 1.** The workflow of population acquisition, selection, grouping, and exclusion criteria.

