



Brain MRI radiomics analysis may predict poor psychomotor outcome in preterm neonates

Youwon Shin¹ · Yoonho Nam^{2,3} · Taehoon Shin^{1,4} · Jin Wook Choi⁵ · Jang Hoon Lee⁶ · Da Eun Jung⁶ · Jiseon Lim³ · Hyun Gi Kim^{2,5} 

Received: 3 December 2020 / Revised: 8 February 2021 / Accepted: 24 February 2021 / Published online: 24 March 2021
© European Society of Radiology 2021

Abstract

Objectives This study aimed to apply a radiomics approach to predict poor psychomotor development in preterm neonates using brain MRI.

Methods Prospectively enrolled preterm neonates underwent brain MRI near or at term-equivalent age and neurodevelopment was assessed at a corrected age of 12 months. Two radiologists visually assessed the degree of white matter injury. The radiomics analysis on white matter was performed using T1-weighted images (T1WI) and T2-weighted images (T2WI). A total of 1906 features were extracted from the images and the minimum redundancy maximum relevance algorithm was used to select features. A prediction model for the binary classification of the psychomotor developmental index was developed and eightfold cross-validation was performed. The diagnostic performance of the model was evaluated using the AUC with and without including significant clinical and DTI parameters.

Results A total of 46 preterm neonates (median gestational age, 29 weeks; 26 males) underwent brain MRI (median corrected gestational age, 37 weeks). Thirteen of 46 (28.3%) neonates showed poor psychomotor outcomes. There was one neonate among 46 with moderate to severe white matter injury on visual assessment. For the radiomics analysis, twenty features were selected for each analysis. The AUCs of prediction models based on T1WI, T2WI, and both T1WI and T2WI were 0.925, 0.834, and 0.902. Including gestational age or DTI parameters did not improve the prediction performance of T1WI.

Conclusions A radiomics analysis of white matter using early T1WI or T2WI could predict poor psychomotor outcomes in preterm neonates.

Key Points

- Radiomics analysis on T1-weighted images of preterm neonates showed the highest diagnostic performance (AUC, 0.925) for predicting poor psychomotor outcomes.
- In spite of 45 of 46 neonates having no significant white matter injury on visual assessment, the radiomics analysis of early brain MRI showed good diagnostic performance (sensitivity, 84.6%; specificity, 78.8%) for predicting poor psychomotor outcomes.
- Radiomics analysis on early brain MRI can help to predict poor neurodevelopmental outcomes in preterm neonates.

Keywords Magnetic resonance imaging · Premature birth · Neurodevelopmental disorder · Infant · Radiomics

Youwon Shin and Yoonho Nam are co-first authors. This work origination is at the Department of Radiology, Ajou University School of Medicine, Ajou University Medical Center, Yeongtong-gu, Suwon, Republic of Korea.

✉ Hyun Gi Kim
catharina@catholic.ac.kr

¹ Division of Mechanical and Biomedical Engineering, Ewha Womans University, Seoul, Republic of Korea

² Department of Radiology, Eunpyeong St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea

³ Division of Biomedical Engineering, Hankuk University of Foreign Studies, Yongin-siGyeonggi-do, Republic of Korea

⁴ Department of Medicine, Case Western Reserve University, Cleveland, OH, USA

⁵ Department of Radiology, Ajou University School of Medicine, Ajou University Medical Center, Suwon, Republic of Korea

⁶ Department of Pediatrics, Ajou University School of Medicine, Ajou University Medical Center, Suwon, Republic of Korea

Abbreviations

AUC	Area under the curve
FA	Fractional anisotropy
GLCM	Gray level co-occurrence matrix
GLDM	Gray level dependence matrix
GLRLM	Gray level run length matrix
GLSZM	Gray level size zone matrix
GM	Gray matter
LoG	Laplacian of Gaussian-filtered
MRMR	Minimum redundancy maximum relevance
PDI	Psychomotor development index
PLIC	Posterior limbs of internal capsule
ROC	Receiver operating characteristics
ROI	Region of interest
T1WI	T1-weighted image
T2WI	T2-weighted image
WM	White matter

Introduction

Preterm birth is a leading worldwide health problem [1, 2]. For example, in the USA, more than 1.5% of live births are of neonates who weigh less than 1500 g. Preterm neonates are more likely to have poor neurodevelopmental outcomes such as cerebral palsy and other cognitive or behavioral impairments [3–7].

Previous studies have shown that intervention services have significant and positive effects on the cognitive development of high-risk preterm neonates [8]. Therefore, the prediction of neurodevelopmental outcomes is crucial for clinicians to identify those neonates who would benefit from developmental interventions [8]. Many neuroimaging studies have tried to correlate imaging markers and neurodevelopmental outcomes in preterm neonates [9–15]. In particular, gross and microstructural characteristics that can be observed on brain MR images have been actively utilized in earlier studies. Gross abnormalities in the white matter (WM) and gray matter (GM) structures of preterm neonates were found to predict their neurodevelopmental outcome at 2 years of age [9]. Abnormalities in WM volume and microstructural signals showed high predictive value for certain neurodevelopmental outcomes [14]. Still, mild to significant psychomotor outcomes were observed in up to 16% of cases which did not exhibit visible WM injury on MR images [16].

Radiomics has been increasingly used for image-based diagnosis and prognosis, enabling the extraction and analysis of quantitative features on medical images [17]. Many radiomics studies have helped radiologists discover hidden gems of information that could aid the diagnosis, treatment, and prognosis prediction of brain tumors. While radiomics is infrequently applied to fields other than oncology in the neuroradiology

field, recent studies showed that radiomics can potentially help classify ischemic group in infants [18], attention deficit hyperactivity disorder [19], or neurofibromatosis 1 subtypes [20].

We hypothesized that a radiomics analysis of WM in preterm neonates could help predict poor neurodevelopmental outcomes even without visible signal abnormalities. Thus, the objective of this study was to develop a model for predicting poor neurodevelopmental outcomes and validate its performance in preterm neonates with minimal or no visible WM abnormalities.

Materials and methods

Study participants

Study patients were selected from a prospective study performed at our institution, which was designed to analyze MR image data acquired from preterm neonates to identify imaging biomarkers that could predict neurodevelopmental outcomes. Our university institutional review board approved the study design and written informed consent was obtained from all patient guardians. All data were anonymized before the MR data were processed.

From September 2017 to December 2018, 60 neonates were enrolled. The inclusion criteria for enrollment were as follows. Preterm neonates delivered (*a*) at less than 28 weeks of gestational age, (*b*) with birth weight less than 1500 g, or (*c*) at less than 37 weeks of gestational age with suspected ischemic brain injury. The exclusion criteria were (*a*) neonates who had congenital malformation, infection, or metabolic diseases that would affect later development, (*b*) neonates who showed intracranial hemorrhage that was grade 3 or higher on ultrasound, or (*c*) neonates with any structural abnormality on a prior brain imaging study. All the enrolled preterm neonates underwent brain MRI before discharge, near or at term-equivalent age. Among enrolled 60 preterm neonates, 56 neonates successfully undergone neurodevelopmental assessment at a corrected age of 12 months. Ten neonates were excluded before radiomics analysis. Therefore, a total of 46 neonates were included for the analysis (Fig. 1).

Neurodevelopmental outcomes and clinical variables

One examiner evaluated the neurodevelopmental abilities of neonates using the Korean-Bayley Scales of Infant and Toddler Development, second edition (Bayley II) when each infant reached a corrected age of 12 months. Bayley II includes the psychomotor development of infants. The Psychomotor Development Index (PDI) assesses gross and fine motor skills. A delay in development

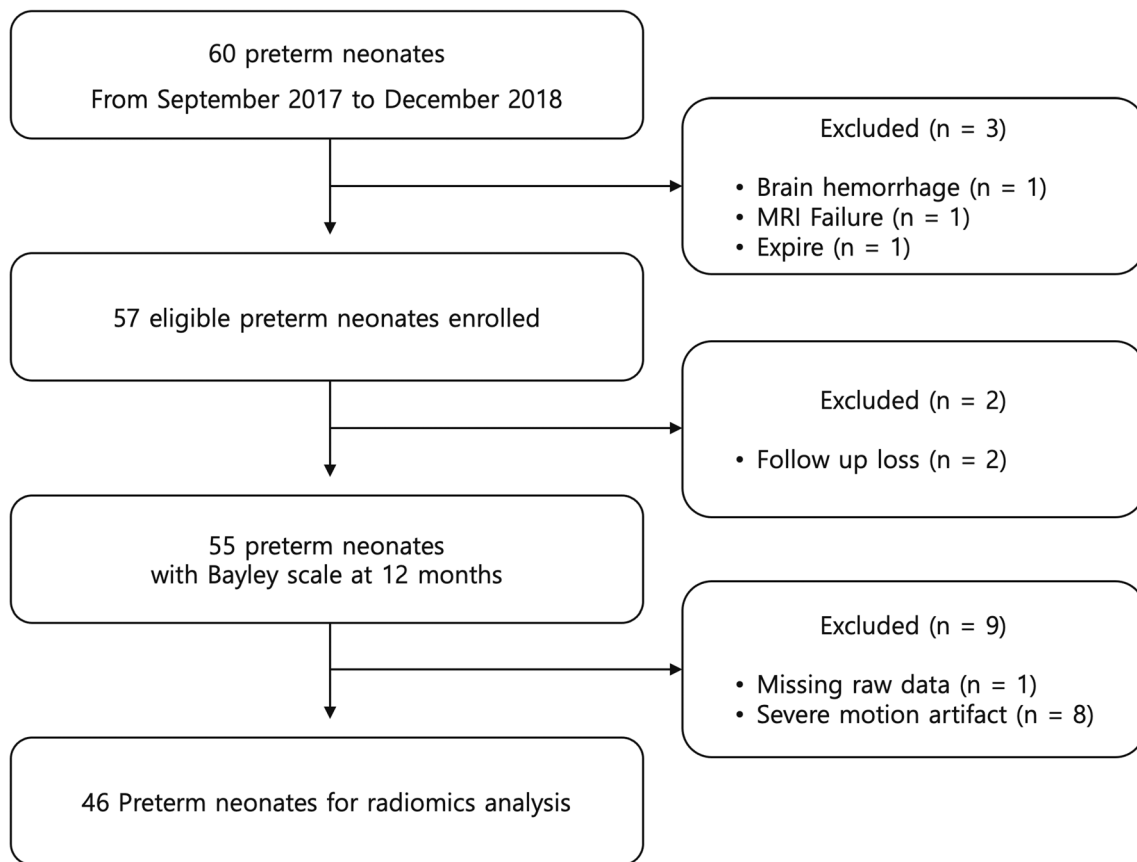


Fig. 1 Flowchart shows study population and exclusion criteria

was defined by a score that was more than 1 standard deviation below the normative mean (PDI score < 85). Clinical variables including birth history were collected from the electronic medical records. Questionnaires were complete by parents for the assessment of maternal education and socioeconomic level as those were associated with neurodevelopmental outcomes of preterm neonates [21, 22].

MRI protocols, regional segmentation, and radiologic assessment

Brain MRI was performed using a 3-T scanner (GE MRI 750w, GE Healthcare, Milwaukee, WI). The MRI examination included a 3D T1-weighted fast spoiled gradient-echo sequence, multiparametric quantitative MRI based on the multi-echo, multi-delay saturation recovery spin echo, and DTI. For the radiomics studies, we used synthetic T1-weighted images (T1WI) and T2-weighted images (T2WI) were generated from the multi-echo, multi-delay data. The imaging protocol is further described in Appendix 1. We obtained regional volumes (total brain, cortical GM, deep GM, and WM) using infant FreeSurfer [23, 24]. Regions of interests were drawn in both of the posterior limbs of internal capsules (PLICs) on a fractional anisotropy (FA) map derived

from DTI [25, 26]. All scan results were evaluated by two radiologists, a pediatric neuroradiologist with 10 years of experience and a neuroradiologist with 13 years of experience. Both radiologists were unaware of the neonates' perinatal history and ultrasonographic findings. White-matter abnormality was rated and classified into two categories: "none or mild abnormality" or "moderate to severe abnormality" [9]. The final classification was made in consensus.

Segmentation, feature extraction, and selection for radiomics analysis

The overall pipeline of the radiomics analysis is illustrated in Fig. 2. Segmentation was performed semi-automatically on the WM region by a radiologist. Radiomics features were defined according to the Pyradiomics Python package, version 3.0 [27]. A total of 11 image types were derived from the standard normalized original T1WI and T2WI and their filtered images. Filtering was done with wavelet-transformed (eight wavelet decompositions) and Laplacian of Gaussian-filtered (LoG) (two values of sigma) images. Radiomics features were extracted from both T1WI and T2WI, resulting in a total of 1906 features. Segmentation and feature extraction methods are described in detail in Appendix 2.

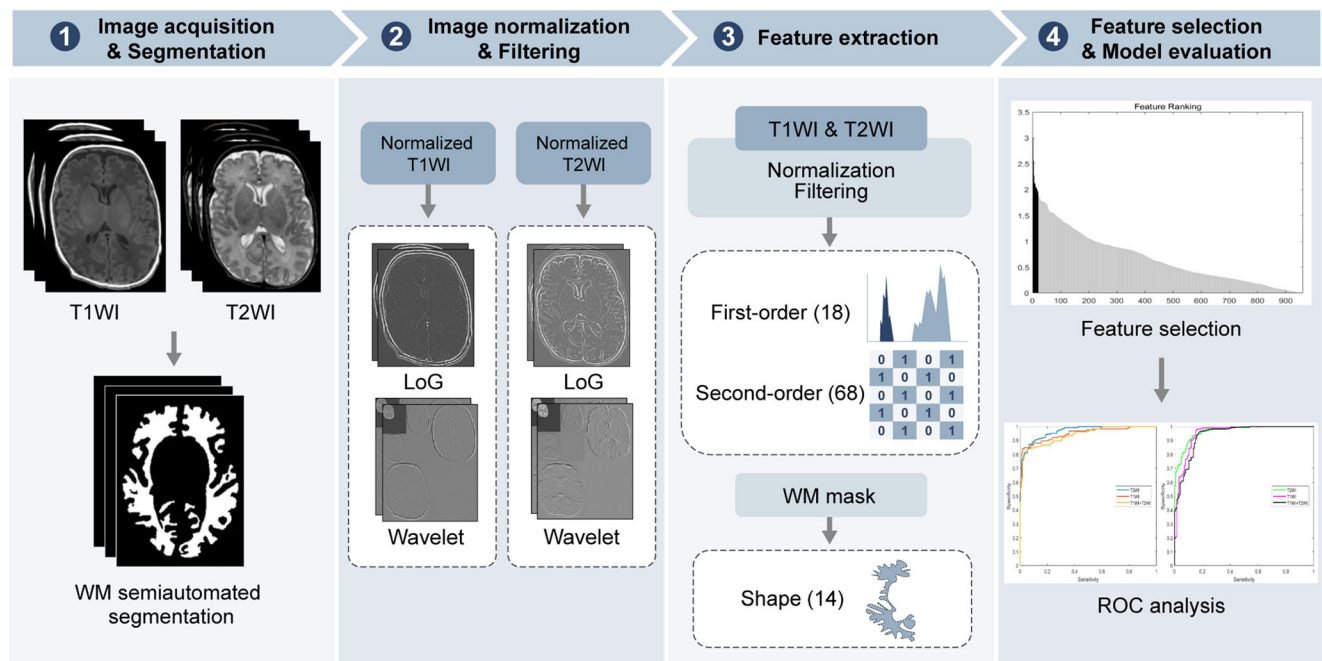


Fig. 2 The overall pipeline of the radiomics analysis. Part 1 includes image acquisition and segmentation. Masks for regions of interest for white matter were acquired based on T2WI using the semi-automated method. Part 2 includes image normalization and filtering. A total of 11 image types (2 Laplacian of Gaussian-filtered images, 8 decompositions of wavelet-transformed images, and the original image) which were all derived from normalized T1WI and T2WI. Part 3 includes extraction of

radiomics features. All of the derived image types were used to extract first-order and second-order features. The shape features were extracted from the binary white matter mask. Part 4 includes feature selection based on the minimum redundancy maximum relevance algorithm and the prediction performance of the model was evaluated by ROC analysis with the selected radiomics features

The minimum redundancy maximum relevance (MRMR) algorithm was used to select features of high relevance to train the prediction model. Specifically, the top 20 relevant features were used to build prediction models based on T1WI and T2WI, respectively. For another model based on a combination of T1WI and T2WI, the top 10 relevant features of each T1WI and T2WI were used.

Model construction and validation

A tree ensemble classifier was used as a prediction model for the binary classification of PDI. The random undersampling boost (RUSBoost) algorithm was used for ensemble aggregation due to its capability to handle imbalanced data with discrete class labels like the training data used in this study [28]. The prediction model was trained with radiomics features selected with the MRMR algorithm in 46 data sets. Eightfold cross-validation was done to evaluate the generalized performance of the prediction model. The receiver operating characteristic (ROC) curve analysis was used to explain the prediction performance of the selected radiomics features. The optimal area under the curve (AUC) was computed for each of the three cases (T1WI, T2WI, and combination of T1WI and T2WI). Clinical or imaging variables with statistical significance on multivariate analysis were additionally included and the prediction performance was computed.

Statistical analysis

Continuous variables were found to be nonparametric. Clinical and imaging characteristics of the neonates with and without delayed psychomotor development were compared using the Mann-Whitney U test. To find significant clinical and imaging variables, we performed univariate and multivariate analyses. Results were considered to be statistically significant if $p < .05$. SPSS version 25.0 (IBM) was used for analysis.

Results

Clinical and imaging characteristics of neonates

A total of 46 neonates (median gestational age, 29 weeks; 26 males) were included. The clinical and imaging characteristics of the neonates are summarized in Table 1. Of the 46 neonates, 13 (28.3%) showed delayed psychomotor development and 33 (71.7%) showed normal psychomotor development. Birth weight ($p = .002$), gestational age ($p = .001$), corrected gestational age at MRI scan ($p = .049$), regional brain volumes ($p = .006-.021$), and FA value in the left PLIC ($p = .035$) were significantly different between the neonates with and without abnormalities. Conversely, gender ($p = .749$), maternal

Table 1 Clinical and imaging characteristics of the preterm neonates

	Delayed psychomotor development ($N = 13$)	Normal psychomotor development ($N = 33$)	p value
Gender, male	8 (61.5%)	18 (54.5%)	.749
Birth weight, g	890 (480)	1320 (305)	.002
Gestational age, wk	28 (4)	30 (2)	.001
Corrected gestational age at MRI scan, wk	39 (4)	37 (2)	.049
Maternal level of education			.497
Primary/secondary	3 (23%)	12 (36%)	
Under/postgraduate	10 (77%)	21 (64%)	
Regional brain volume, ml			
Total brain	241 (51)	203 (75)	.021
Cortical GM	104 (16)	84 (40)	.006
Deep GM	16 (3)	14 (4)	.019
WM	97 (15)	84 (23)	.014
Diffusion tensor imaging			
Rt PLIC FA value	0.432 (0.08)	0.409 (0.03)	.262
Lt PLIC FA value	0.462 (0.06)	0.441 (0.03)	.035
Visual abnormality			.283
None or mild	12 (92%)	33 (100%)	
Moderate to severe	1 (8%)	0 (0%)	

Values are numbers (%) or medians (interquartile range), *PDI* psychomotor developmental index; *GM* gray matter; *WM* white matter; *PLIC* posterior limb of internal capsule; *FA* fractional anisotropy

education level ($p = .497$), and FA value in the right PLIC ($p = .262$) were not significantly different between the neonates with and without delayed psychomotor development. The imaging characteristics of the neonates assessed by the two radiologists did not significantly differ between the two groups ($p = .283$). Among the neonates with delayed psychomotor development, there was one neonate (1/13) with moderate to severe WM abnormality. Among the neonates with normal psychomotor development, there was no neonate with moderate to severe WM abnormality.

Feature selection

Table 2 shows the top 10 relevant image features extracted from T1WI and T2WI. Small dependence low gray level emphasis (gray level dependence matrix [GLDM]) obtained from the wavelet transformed with high-low-high-pass filter T1WI and cluster shade (gray level co-occurrence matrix [GLCM]) obtained from the LoG T2WI ranked as the most relevant feature in each case. In T1WI, seven features were obtained from the wavelet-transformed images, two features from the LoG images and one feature from the original images. In T2WI, one feature was obtained from the original images, one feature from the LoG images and eight features from the wavelet-transformed images. Figure 3 demonstrates the heatmap of the top 10 features selected from T1WI and T2WI, respectively, arranged by the binary class of PDI.

Model performance

Figure 4a shows the ROC curves and AUC values for the prediction of poor outcomes in preterm neonates based on T1WI, T2WI, and combination of T1WI and T2WI. The prediction model performed best with T1WI, providing an AUC of 0.925 (95% confidence interval [CI]: 0.85–1.0), sensitivity of 84.6%, specificity of 78.8%, and accuracy of 80.4%. The prediction model based on T2WI resulted in an AUC of 0.834 (95% CI: 0.72–0.95), sensitivity of 76.9%, specificity of 75.6%, and accuracy of 76.1%. The prediction model based on the combination of T1WI and T2WI resulted in an AUC of 0.902 (95% CI: 0.82–0.99), sensitivity of 92.3%, specificity of 81.8%, and accuracy of 84.8%. The confusion matrix evaluated by each prediction model is shown in Supplementary Figure S1.

Results from multivariate and univariate analyses using clinical and imaging variables are shown in the Supplementary Tables. The significant variable on multivariate analysis was gestational age. Regional brain volumes were not a significant factor predicting poor psychomotor outcomes. Higher FA value in left PLIC resulted statistically insignificant but high odds ratio (134.5). We added gestational age, FA value in the left PLIC, or both in the prediction model based on T1WI. Adding gestational age or FA values did not result in higher prediction performance (Fig. 4b). The prediction model based on T1WI with gestational age resulted in an AUC of 0.900 (95% CI: 0.81–0.99), sensitivity of 84.6%,

Table 2 List of significant MR radiomics features to classify psychomotor development in preterm neonates

Order	T1-weighted image			T2-weighted image		
	Image type	Feature type	Radiomics feature	Image type	Feature type	Radiomics feature
1	Wavelet-HLH	GLDM	Small dependence low gray level emphasis	LoG	GLCM	Cluster shade (3D, Sigma-5mm)
2	Wavelet-HLH	GLSZM	Low gray level zone emphasis	Wavelet-LHL	First order	Root mean squared
3	Wavelet-HLH	GLSZM	Small area low gray level emphasis	Original	Shape	Maximum 2D diameter column
4	Wavelet-HLH	GLRLM	Short run low gray level emphasis	Wavelet-HHL	GLSZM	Large area high gray level emphasis
5	Wavelet-HLH	GLRLM	Low gray level run emphasis	Wavelet-LHL	First order	Median
6	Wavelet-HLH	GLDM	Low gray level emphasis	Wavelet-HHL	GLDM	Large dependence high gray level emphasis
7	Wavelet-HLH	GLRLM	Long run low gray level emphasis	Wavelet-LHL	First order	90 percentile
8	LoG	First order	Maximum (3D, Sigma-5mm)	Wavelet-LHL	First order	Mean
9	Original	Shape	Maximum 2D diameter column	Wavelet-LLH	First order	Root mean squared
10	LoG	GLRLM	Run entropy (3D, Sigma-5mm)	Wavelet-HLL	GLDM	Dependence entropy

LoG Laplacian of Gaussian-filtered images; *H* high-pass filter; *L* low-pass filter; *GLCM* gray level co-occurrence matrix, *GLRLM* gray level run length matrix; *GLSZM* gray level size zone matrix; *GLDM* gray level dependence matrix

specificity of 78.8%, and accuracy of 80.4%. The model based on T1WI with FA value in the left PLIC resulted in an AUC of 0.908 (95% CI: 0.82–0.99), sensitivity of 84.6%, specificity of 75.8%, and accuracy of 78.3%. The prediction model based on T1WI including both gestational age and FA value in the left PLIC resulted in an AUC of 0.897 (95% CI: 0.81–0.99), sensitivity of 84.6%, specificity of 78.8%, and accuracy of 80.4%. The confusion matrix evaluated by each prediction model is shown in Supplementary Figure S2.

Discussion

In this study, we demonstrated that a radiomics analysis of the WM of preterm neonates on MRI enables the prediction of poor neurodevelopmental outcomes. This was achieved through automated extraction of important voxel-wise

imaging features and training of a prediction model for binary classification. We also found that T1WI is more effective than T2WI for predicting poor neurodevelopment when data is fed into the developed classifier.

It should be noted that our radiomics analysis was performed on preterm neonates and most of the neonates did not present with moderate to severe WM injury (45 out of 46). The radiomics analysis was able to successfully predict the neurodevelopmental outcomes of these neonates. The rate of neurodevelopmental impairment among those with no to mild WM injury in our study (26.7%) was comparable to the rate reported in a previous study (15–27%) [29]. This suggests that although moderate to severe WM injury is a good predictive imaging marker for poor neurodevelopmental outcomes, about one-fourth of preterm neonates without significant WM injury will still eventually have problems with their neurodevelopment. This finding underscores the importance

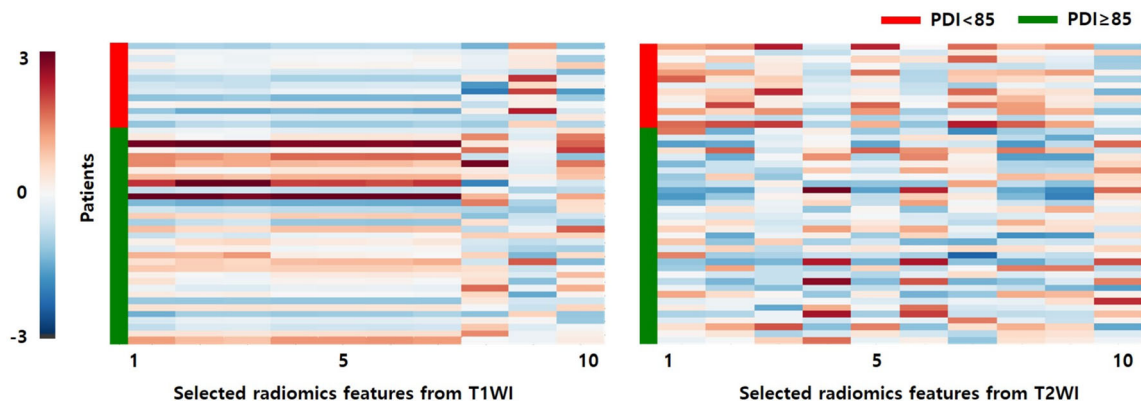


Fig. 3 Heatmap of the top 10 radiomics features of each T1WI and T2WI. Each column corresponds to radiomics features sorted by rank, and each row corresponds to the patients sorted by binary class. (PDI < 85: delayed psychomotor development, PDI ≥ 85: normal psychomotor development)

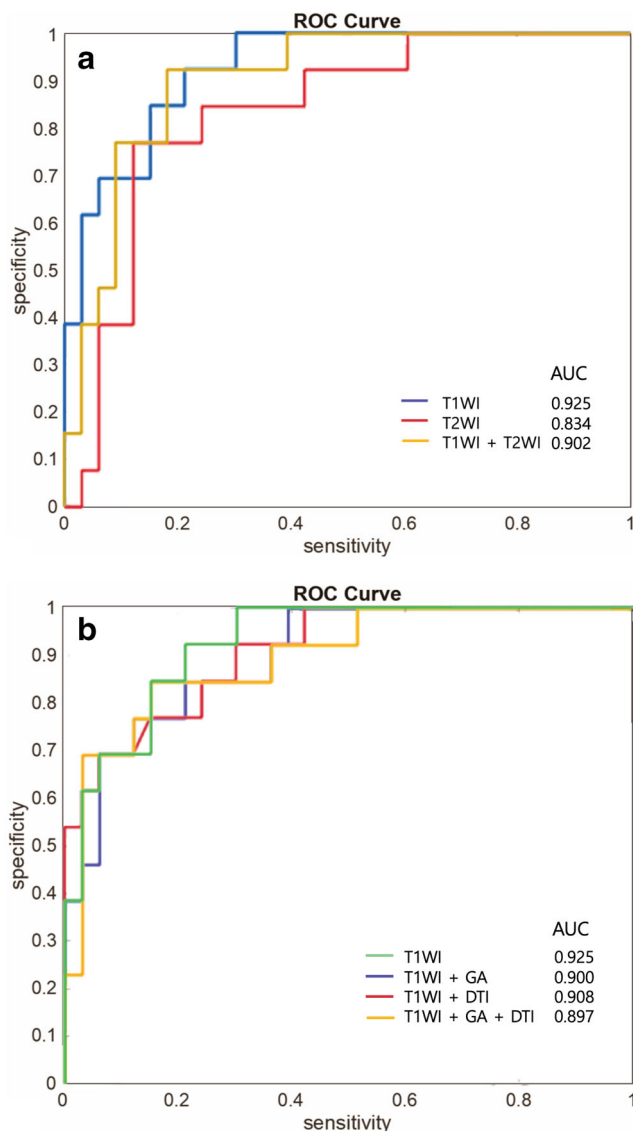


Fig. 4 The receiver operating characteristic curves and areas under the curve for three cases (a) T1WI, T2WI, and combination of T1WI and T2WI, and for four cases (b) T1WI, T1WI with gestational age, T1WI with the DTI parameter, and T1WI with both variables. GA, gestational age

of predicting neurodevelopmental outcomes in neonates without significant WM injury.

Previous studies which assessed WM to predict neurodevelopmental impairment in preterm neonates focused on abnormal signal intensity or volume loss [14, 15, 30]. The sensitivity of our model using T1WI (84.6%) was higher than that of previous studies (38–50%) using visual assessments for WM injury, but the specificity of our model was lower (78.8% vs 82–100%) [9, 15]. It should be noted that we included only one case with significant WM injury per visual assessment, which makes it difficult to make direct comparisons with previous studies that included various degrees of WM injuries. Nonetheless, the higher performance of our

prediction model in the absence of significant WM injury may be attributed to the highly quantified WM features which is a trait not possible with the conventional visual scoring system [9]. We believe that differences in the extracted radiomics features between normal and poor neurodevelopmental outcomes arose mainly from the microstructural changes caused by preterm birth.

Including gestational age or DTI parameters did not have additional benefits for predicting psychomotor outcomes in this study, emphasizing the value of radiomics feature analysis. A recent study showed that clinical parameters such as gestational age and social economic status had a higher value than imaging parameters of brain volume when predicting cognitive outcomes [31]. On the other hand, the same study showed brain volume as a significant factor when predicting psychomotor outcomes [31]. Although our assessment of regional volumes showed significant differences between the normal and poor psychomotor outcome groups, these parameters were insignificant on multivariate analysis. Similarly, although a DTI parameter was significantly different between normal and poor psychomotor outcome groups as in previous studies [25, 26], it was insignificant on multivariate analysis. The highest AUC for predicting poor psychomotor outcomes using regional brain volume or clinical variables in the past study was 0.64 [31]. We can postulate that radiomics features derived from WM may have higher predictive values than regional brain volume, DTI parameters, or clinical variables such as gestational age.

In our study, the significant radiomics features of T1WI and T2WI turned out to be GLDM, GLCM, and gray level size zone matrix (GLSZM). GLDM is a texture feature which quantifies gray level dependency that is defined as the number of connected voxels within a certain distance from the center voxel in an image [32]. GLCM is also a texture feature which reflects signal heterogeneity by calculating the joint probability distributions of pixel pairs [32]. Our results which showed significant difference in the GLCM of WM according to neurodevelopmental outcomes were in line with previous studies that showed the predictive value of WM heterogeneity in adults [33–35]. GLSZM represents the number of connected voxels with the same intensity values [32]. Many major features were significantly more relevant when obtained from wavelet-transformed images, especially from high-low-high (HLH) decomposition. We believe that texture analysis will show better performance when various features are combined.

There are some limitations to this study. First, the data set was of relatively small size ($n = 46$). Further validation with a larger data set is needed before clinical application is possible. Second, we only validated the model performance internally without external validation. Related future work should include test data from different institutions for validation and optimization of image acquisition

and analysis protocols. Third, the use of synthetic T1WI and T2WI instead of conventional T1WI or T2WI may limit the transferability of our model to other institutions. Although synthetic and true contrast images are thought to be similar in children [36], the exact signal intensity and distribution of images are different and this would affect the radiomics features. Therefore, signal intensity needs to be normalized to a fine degree during the entire procedure to validate our findings with conventional T1WI and T2WI. Finally, we did not subcategorize the study subjects according to WM injury before analysis. The subjects were prospectively enrolled and those with high-grade intracranial hemorrhages were excluded from the study population. Still, there was one subject with moderate to high-grade WM injury. This suggests that although few, some subjects with different degrees of WM injuries might be included in the radiomics analysis which may affect the radiomics features.

In conclusion, this study showed that radiomics-driven analysis of early brain MR images can predict poor psychomotor outcomes in preterm neonates. Our prediction model might be of potential use in quantitative analysis using T1WI and T2WI and in the absence of significant WM abnormalities.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00330-021-07836-7>.

Acknowledgements The authors thank Sun Mi Cho, Department of Psychiatry, Ajou University Hospital, for contributing data on neurodevelopmental outcomes.

Funding This study has received funding from the National Research Foundation of Korea.

Declarations

Guarantor The scientific guarantor of this publication is Seung Eun Jung, MD.

Conflict of interest The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Statistics and biometry No complex statistical methods were necessary for this paper.

Informed consent Written informed consent was obtained from all subjects (patients) in this study.

Ethical approval Institutional Review Board approval was obtained.

Study subjects or cohorts overlap Some study subjects have been previously reported in *European Radiology* titled “Texture analysis of deep medullary veins on susceptibility-weighted imaging in infants: evaluating developmental and ischemic changes.”

Methodology

- prospective
- prognostic study
- performed at one institution

References

1. Behrman RE, Butler AS (2007) Preterm birth: causes, consequences, and prevention. Washington, DC: National Academies Press (US)
2. Lee AC, Katz J, Blencowe H et al (2013) National and regional estimates of term and preterm babies born small for gestational age in 138 low-income and middle-income countries in 2010. *Lancet Glob Health* 1:e26–e36
3. Marlow N, Wolke D, Bracewell MA, Samara M (2005) Neurologic and developmental disability at six years of age after extremely preterm birth. *N Engl J Med* 352:9–19
4. Fanaroff AA, Stoll BJ, Wright LL et al (2007) Trends in neonatal morbidity and mortality for very low birthweight infants. *Am J Obstet Gynecol* 196:147.e1–147.e8
5. Saigal S, Doyle LW (2008) An overview of mortality and sequelae of preterm birth from infancy to adulthood. *Lancet* 371:261–269
6. Larroque B, Ancel PY, Marret S et al (2008) Neurodevelopmental disabilities and special care of 5-year-old children born before 33 weeks of gestation (the EPIPAGE study): a longitudinal cohort study. *Lancet* 371:813–820
7. Latal B (2009) Prediction of neurodevelopmental outcome after preterm birth. *Pediatr Neurol* 40:413–419
8. Ramey CT, Bryant DM, Wasik BH, Sparling JJ, Fendt KH, LaVange LM (1992) Infant Health and Development Program for low birth weight, premature infants: program elements, family participation, and child intelligence. *Pediatrics* 89:454–465
9. Woodward LJ, Anderson PJ, Austin NC, Howard K, Inder TE (2006) Neonatal MRI to predict neurodevelopmental outcomes in preterm infants. *N Engl J Med* 355:685–694
10. Woodward LJ, Clark CA, Bora S, Inder TE (2012) Neonatal white matter abnormalities an important predictor of neurocognitive outcome for very preterm children. *PLoS One* 7:e51879
11. Van't Hooft J, van der Lee JH, Opmeer BC et al (2015) Predicting developmental outcomes in premature infants by term equivalent MRI: systematic review and meta-analysis. *Syst Rev* 4:71
12. Roze E, Benders MJ, Kersbergen KJ et al (2015) Neonatal DTI early after birth predicts motor outcome in preterm infants with periventricular hemorrhagic infarction. *Pediatr Res* 78:298–303
13. Ullman H, Spencer-Smith M, Thompson DK et al (2015) Neonatal MRI is associated with future cognition and academic achievement in preterm children. *Brain* 138:3251–3262
14. Guo T, Duerden EG, Adams E et al (2017) Quantitative assessment of white matter injury in preterm neonates: association with outcomes. *Neurology* 88:614–622
15. Cayam-Rand D, Guo T, Grunau RE et al (2019) Predicting developmental outcomes in preterm infants: a simple white matter injury imaging rule. *Neurology* 93:e1231–e1240
16. Chau V, Synnes A, Grunau RE, Poskitt KJ, Brant R, Miller SP (2013) Abnormal brain maturation in preterm neonates associated with adverse developmental outcomes. *Neurology* 81:2082–2089
17. Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: images are more than pictures, they are data. *Radiology* 278:563–577
18. Kim HG, Choi JW, Han M, Lee JH, Lee HS (2020) Texture analysis of deep medullary veins on susceptibility-weighted imaging in infants: evaluating developmental and ischemic changes. *Eur Radiol* 30:2594–2603

19. Sun H, Chen Y, Huang Q et al (2018) Psychoradiologic utility of MR imaging for diagnosis of attention deficit hyperactivity disorder: a radiomics analysis. *Radiology* 287:620–630
20. Liu Y, Jordan JT, Bredella MA et al (2020) Correlation between NF1 genotype and imaging phenotype on whole-body MRI: NF1 radiogenomics. *Neurology* 94:e2521–e2531
21. Patra K, Greene MM, Patel AL, Meier P (2016) Maternal education level predicts cognitive, language, and motor outcome in preterm infants in the second year of life. *Am J Perinatol* 33:738–744
22. Benavente-Fernández I, Synnes A, Grunau RE et al (2019) Association of socioeconomic status and brain injury with neurodevelopmental outcomes of very preterm children. *JAMA Netw Open* 2:e192914
23. Zöllei L, Iglesias JE, Ou Y, Grant PE, Fischl B (2020) Infant FreeSurfer: an automated segmentation and surface extraction pipeline for T1-weighted neuroimaging data of infants 0–2 years. *Neuroimage* 218:116946
24. de Macedo Rodrigues K, Ben-Avi E, Sliva DD et al (2015) A FreeSurfer-compliant consistent manual segmentation of infant brains spanning the 0–2 year age range. *Front Hum Neurosci* 9:21
25. Arzoumanian Y, Mirmiran M, Barnes PD et al (2003) Diffusion tensor brain imaging findings at term-equivalent age may predict neurologic abnormalities in low birth weight preterm infants. *AJNR Am J Neuroradiol* 24:1646–1653
26. De Bruïne FT, Van Wezel-Meijler G, Leijser LM et al (2013) Tractography of white-matter tracts in very preterm infants: a 2-year follow-up study. *Dev Med Child Neurol* 55:427–433
27. van Griethuysen JJM, Fedorov A, Parmar C et al (2017) Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 77:e104–e107
28. Seiffert C, Khoshgoftaar TM, Van Hulse J, Napolitano A (2008) RUSBoost: Improving classification performance when training data is skewed. 2008 19th International Conference on Pattern Recognition. <https://doi.org/10.1109/ICPR.2008.4761297>
29. Spittle AJ, Cheong J, Doyle LW et al (2011) Neonatal white matter abnormality predicts childhood motor impairment in very preterm children. *Dev Med Child Neurol* 53:1000–1006
30. Hong HS, Kim SS, Park GY (2020) MRI findings to predict neurodevelopmental outcomes in preterm infants near term-equivalent age. *Investig Magn Reson Imaging* 24:30–37
31. Gui L, Loukas S, Lazeyras F, Hüppi PS, Meskaldji DE, Borradori Tolsa C (2019) Longitudinal study of neonatal brain tissue volumes in preterm infants and their ability to predict neurodevelopmental outcome. *Neuroimage* 185:728–741
32. Zwanenburg A, Leger S, Vallières M, Löck S (2019) Image biomarker standardisation initiative. *arXiv Prepr arXiv161207003*
33. Shu Z, Xu Y, Shao Y, Pang P, Gong X (2020) Radiomics from magnetic resonance imaging may be used to predict the progression of white matter hyperintensities and identify associated risk factors. *Eur Radiol* 30:3046–3058
34. Takahashi T, Murata T, Omori M et al (2004) Quantitative evaluation of age-related white matter microstructural changes on MRI by multifractal analysis. *J Neurol Sci* 225:33–37
35. Shao Y, Chen Z, Ming S et al (2018) Predicting the development of normal-appearing white matter with radiomics in the aging brain: a longitudinal clinical study. *Front Aging Neurosci* 10:393
36. Lee SM, Choi YH, Cheon JE et al (2017) Image quality at synthetic brain magnetic resonance imaging in children. *Pediatr Radiol* 47:1638–1647

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.