Original Article
Yonsei Med J 2022 Sep;63(9):856-863
https://doi.org/10.3349/ymj.2022.63.9.856
Yonsei Medical Journal
YMJ
pISSN: 0513-5796 · eISSN: 1976-2437

A Radiomics-Based Model with the Potential to Differentiate Growth Hormone Deficiency and Idiopathic Short Stature on Sella MRI

Taeyoun Lee[1], Kyungchul Song[2], Beomseok Sohn[1], Jihwan Eom[1,3], Sung Soo Ahn[1], Ho-Seong Kim[2], and Seung-Koo Lee[1]

[1]Department of Radiology and Research Institute of Radiological Science and Center for Clinical Imaging Data Science, Yonsei University College of Medicine, Seoul;
[2]Department of Pediatrics, Severance Children's Hospital, Endocrine Research Institute, Yonsei University College of Medicine, Seoul;
[3]Department of Computer Science, Yonsei University, Seoul, Korea.

Purpose: We hypothesized that a radiomics approach could be employed to classify children with growth hormone deficiency (GHD) and idiopathic short stature (ISS) on sella magnetic resonance imaging (MRI). Accordingly, we aimed to develop a radiomics prediction model for differentiating GHD from ISS and to evaluate the diagnostic performance thereof.

Materials and Methods: Short stature pediatric patients diagnosed with GHD or ISS from March 2011 to July 2020 at our institution were recruited. We enrolled 312 patients (GHD 210, ISS 102) with normal sella MRI and temporally split them into training and test sets (7:3). Pituitary glands were semi-automatically segmented, and 110 radiomic features were extracted from the coronal T2-weighted images. Feature selection and model development were conducted by applying mutual information (MI) and a light gradient boosting machine, respectively. After training, the model's performance was validated in the test set. We calculated mean absolute Shapley values for each of the selected input features using the Shapley additive explanations (SHAP) algorithm. Volumetric comparison was performed for GHD and ISS groups.

Results: Ten radiomic features were selected by MI. The receiver operating characteristics curve of the developed model in the test set was 0.705, with an accuracy of 70.6%. When analyzing SHAP plots, root mean squared values had the highest impact in the model, followed by various texture features. In volumetric analysis, sagittal height showed a significant difference between GHD and ISS groups.

Conclusion: Radiomic analysis of sella MRI may be able to differentiate between GHD and ISS in clinical practice for short-statured children.

Key Words: MRI, child, human growth hormone, growth disorders, pituitary gland

INTRODUCTION

Short stature is defined as a height below the third percentile or more than two standard deviations below the corresponding mean value for a given age, sex, and race. There are many possible causes of short stature, among growth hormone (GH) deficiency (GHD) is one of the most important.[1,2] GH is a polypeptide hormone produced by the pituitary gland and is involved in growth and cell reproduction. GHD can be caused by congenital or acquired factors.[2,3] The tools for the screening and evaluating GHD include auxology, radiographic assessment of bone age, measurement of insulin-like growth factor-I

Received: January 26, 2022 **Revised:** April 21, 2022
Accepted: June 7, 2022 **Published online:** August 18, 2022
Corresponding author: Beomseok Sohn, MD, Department of Radiology and Research Institute of Radiological Science and Center for Clinical Imaging Data Science, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea.
Tel: 82-2-2228-7400, Fax: 82-2-2227-8337, E-mail: BEOMSEOKSOHN@yuhs.ac
This study was presented at the Korean Congress of Radiology 2021, Seoul.
•The authors have no potential conflicts of interest to disclose.

(IGF-I), GH stimulation testing, brain magnetic resonance imaging (MRI), and, in certain cases, genetic testing.[1] Among these, the GH stimulation test is widely used for confirming GHD diagnosis, wherein responses to at least two GH stimulation tests with various pharmacological stimuli, such as insulin, clonidine, glucagon, arginine, L-dopa, or growth hormone-releasing hormone (GHRH), are below normal.[4,5] However, the GH stimulation test has some limitations: due to an arbitrary cutoff value, there are issues concerning validity and reproducibility, making it insufficient for GHD diagnosis.[6-8] Additionally, multiple venous samples and hospital admissions required to conduct the test can be burdensome to patients.[8] Without evidence of systemic, endocrine, nutritional, or chromosomal abnormalities, children with short stature are diagnosed with idiopathic short stature (ISS).[9,10] Proper classification and a precise diagnosis of short stature are essential for determining treatment policy and prognosis.

Currently, in many pediatric endocrinology clinics, evaluation of the pituitary gland by sella MRI is an essential diagnostic process for patients with GHD.[11] Neuroimaging is currently focused on the measurement of the pituitary gland size using various methods. Many studies have reported volumetric differences between GHD and ISS and have demonstrated anterior pituitary hypoplasia with GHD.[12-16] Kessler, et al.[17] reviewed pituitary volume (PV) differences among three groups (control, GHD, and ISS) and reported significant differences in PV and minimal increases in PV with age in the GHD group. However, there are GHD and ISS children who exhibit normal PV for their age.

Radiomics is a methodology that seeks to extract a large number of features from medical images using mathematical algorithms.[18] These 'radiomic features' can detect molecular profiles or disease characteristics that the human eye cannot. As radiomics applies advanced computational methods to automatically extract and analyze hundreds of quantitative radiomic features from tissues of interest based on medical imaging data, its utility for the acquisition of quantitative information has been widely investigated.[18,19] Many studies have been performed on adult pituitary adenomas; however, the application of a radiomic approach to the pituitary of children is lacking.[20-26]

We hypothesized that a radiomics approach could be helpful in classifying children with GHD and ISS on normal sella MRI. To our knowledge, no previous study has investigated sella MRI radiomics in children. The purpose of this study was to develop a radiomics-based prediction model from normal sella MRI to classify GHD and ISS in short stature children.

MATERIALS AND METHODS

Patient population

This retrospective study was approved by our Institutional Review Board (Registration number: 4-2021-0594), and the requirement for informed consent was waived. All experiments were performed in accordance with relevant guidelines and regulations (including Declaration of Helsinki). We reviewed the electronic medical records of patients consecutively diagnosed with short stature, either GHD or ISS, at our hospital (tertiary care center) from March 2011 to July 2020. In total, 386 patients who underwent appropriate endocrinologic and radiologic evaluation using GH provocation test and sella MRI were reviewed. GHD was attributed to the patients whose height was below the 3rd percentile for age and sex and whose peak GH levels were less than 10.0 ng/mL after two GH stimulation tests using insulin, arginine, and/or L-dopa.[27,28] ISS was defined as a height below the 3rd percentile for age and sex without any other identifiable causes, including endocrine, nutritional, skeletal, or genetic abnormalities.[27,29] Among 386 patients, the patients with other medical diseases affecting growth, with a brain or sella anomaly on MRI, or severe motion artifact/process error were excluded. The training and test-validation sets were prepared by temporally dividing the patients into two subgroups, with a proportion of 7:3. The date of separation was July 4, 2014.

Image acquisition

The patients were scanned on various 3.0 T MRI units [Discovery MR750/750, GE Healthcare (Chicago, IL, USA); or Ingenia/Ingenia CX, Philips Medical Systems (Amsterdam, Nederland)] under general anesthesia. The imaging protocols included pre-contrast sagittal T1-weighted, sagittal T2-weighted, high-resolution coronal T2-weighted, and contrast-enhanced T1-weighted imaging (T1C). The sequence parameters of the high-resolution coronal T2-weighted images (T2WI) are as follows: TR/TE=2447/80 ms; slice thickness=1.5 mm; intersection gap=0 mm; field of view (FOV)=20×20 cm; pixel spacing=0.391×0.391 mm.

Image processing and radiomic feature extraction

The Digital Imaging and Communication in Medicine images of high-resolution T2WI were converted to NIfTI files. The images were resampled with 1×1×1 mm resolution, and low-frequency intensity nonuniformity was corrected using N4 bias correction.[30] The pituitary glands were semi-automatically segmented on each slice of the coronal T2WI by a neuroradiologist with 5 years of experience (B.S) who was blinded to each children's clinical diagnosis using open-source software (Medical Image Processing, Analysis, and Visualization; Center for Information Technology, National Institutes of Health, Bethesda, MD, USA). The segmentation methods included signal intensity thresholding, region growing, and edge detection. The outermost boundaries of the pituitary gland sliced by slice on coronal T2WI were delineated. Entire pituitary glands were included in regions of interest. To evaluate the robustness of segmentation, 31 patients (10% of the final cohort) were randomly selected, and segmentation was performed again by the neuroradiologist and a third-year radiology resident. The dice coefficient between segmentation masks was evaluated. Subsequently,

the radiomic features were extracted using Pyradiomics 2.1.0 (http://www.radiomics.io/pyradiomics.html), with 128 fixed bin counts.[31]

Machine Learning and Statistical Analysis

The mutual information (MI) method was used for feature selection, and the light gradient boosting machine (LightGBM) method was used to train a classification model.[32,33] Ten-fold cross validation was performed internally during the train process. Model validation was performed in the test set. The receiver operating characteristics (ROC) curve (AUC) were drawn for the model with the model performance in the test set. Model performance in the test set was assessed with an area under AUC, sensitivity, specificity, accuracy and F1 score. All processes of the investigation are summarized in Fig. 1. To demonstrate the clinical usefulness of the radiomics model, we additionally built a clinical model with only sex, age, and shape variables. We compared AUC values between the radiomics and clinical models.

To examine which parameters played important roles in the radiomics model with the best performance, we calculated mean absolute Shapley values for each of the selected input features using the Shapley additive explanations (SHAP) algorithm.[34]

Volumetric compare between the GHD and ISS groups

For additional analysis, we compared volumes and diameters between GHD and ISS group. All processes up to this point were performed using Python 3 with ScikitLearn library v0.21.2 and the R software (version 3.5.1; R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

The electronic medical records of 386 patients were reviewed. Patients with other medical diseases affecting growth, such as small for gestational age, Turner syndrome, or chronic kidney disease, were excluded (n=55). The patients with a brain or sella anomaly on sella MRI, such as empty sella syndrome or Rathke's cleft cyst (n=16), severe motion artifact (n=2) on MRI, or process error (n=1), were also excluded. Finally, 312 patients were enrolled. In this final cohort, 210 patients were confirmed to have GHD and 102 to have ISS by provocative GH tests. The training and test sets were prepared by temporally dividing the patients into training and test sets, at a proportion of 7:3. In total, 218 patients were included in the training set, in which 147 (67.4%) had GHD and 71 (32.6%) had ISS. The test set included 94 patients, among which 63 (67.0%) and 31 (33.0%) were confirmed to have GHD and ISS, respectively. The patient characteristics are summarized in Table 1.

The dice scores of segmentation masks from randomly selected subgroups were 0.88 [IQR 0.80-0.91] for inter-rater reliability assessment and 0.86 [IQR 0.77-0.89] for intra-rater assess-

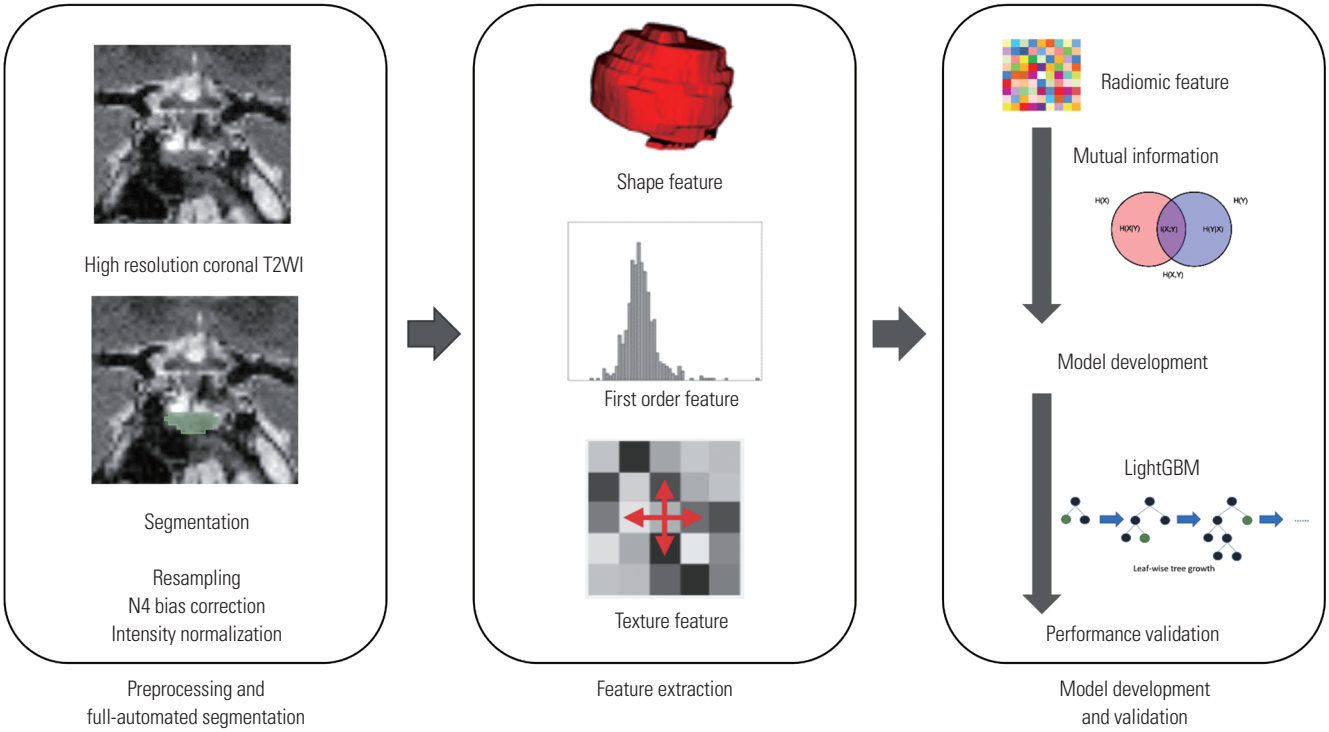

Fig. 1. Radiomics pipeline. After resampling, N4 bias correction, image intensity normalization, semi-automated segmentation of the pituitary gland on high-resolution coronal T2WI was applied. After radiomic feature extraction, feature selection and prediction model development/validation were performed. T2WI, T2-weighted image; LightGBM, light gradient boosting machine.

https://doi.org/10.3349/ymj.2022.63.9.856

ment. A total of 110 radiomic features was extracted per patient, including shape (16 features), first-order (19 features), gray level co-occurrence matrix (GLCM; 24 features), gray level dependence matrix (GLDM; 14 features), gray level size zone matrix (GLSZM; 16 features), gray level run length matrix (GLRLM; 16 features), and neighboring gray tone difference matrix (NGTDM; five features).

After MI feature selection, 10 features were selected. Training and ten-fold cross-validation was performed with LightGBM to the training set, and ROC curves were drawn with the test set (Fig. 2). Model performance in test set was as follows: AUC, 0.705 [95% confidence interval (CI) 0.588–0.816]; accuracy, 70.6%; sensitivity, 81.4%; specificity, 46.9%; and F1 score, 0.792.

In the development of the clinical model, sex and five shape parameters were selected, including axial diameter, height, volume, 3D longest diameter, and 3D shortest diameter. The AUC of the clinical model in test set was 0.473 (95% CI 0.348–0.598). The accuracy of the clinical model was 63.7%. The ensemble of radiomics model and clinical model did not show performance improvement.

The mean absolute Shapley values for each of the selected radiomic features were calculated to visualize feature importance in the developed radiomics model. In our model, 10 radiomic features were selected by MI feature selection: root mean squared (RMS) (first order), high gray level zone emphasis (GLSZM), small area low gray level emphasis (GLSZM), skewness (first order), long run high gray level emphasis (GLRLM), coarseness (NGTDM), joint energy (GLCM), short run high gray level emphasis (GLRLM), high gray level emphasis (GLDM), and high gray level run emphasis (GLRLM). With the SHAP algorithm, the importance of these 10 features were visualized (Fig. 3). When analyzing the dot SHAP summary and decision

Table 1. Patient Demographics for Each Dataset

Demographics	Total (n=312)	Training set (n=218)	Test set (n=94)	*p* value
Age (month)		87.0±35.3	89.9±34.3	0.49†
Sex				0.31*
Male	186	134 (61.5)	52 (55.3)	
Female	126	84 (38.5)	42 (44.7)	
Diagnosis				0.94*
GHD	210	147 (67.4)	63 (67.0)	
ISS	102	71 (32.6)	31 (33.0)	

Data are presented as mean±standard deviation or n (%).
*Calculated using chi-square test; †Calculated using t-test.

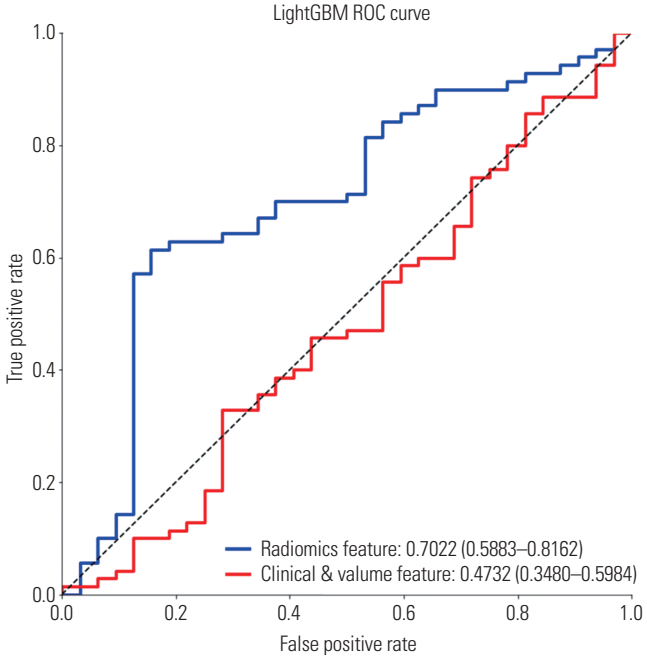

Fig. 3. Receiver operating characteristic (ROC) curves (AUC) from the radiomics model and clinical model in the test set. LightGBM, light gradient boosting machine.

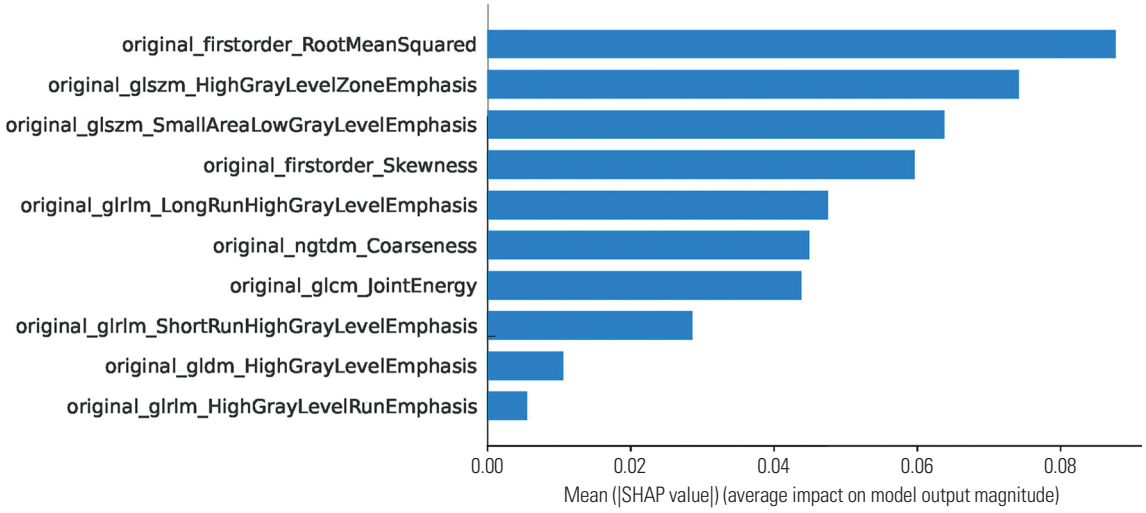

Fig. 2. Feature importance according to mean absolute Shapley additive explanations (SHAP) values for the prediction of growth hormone deficiency in the light gradient boosting machine model from the test set.

plots, high values for small area low gray level emphasis, skewness, and joint energy had influence in predicting GHD. Conversely, low values for high gray level zone emphasis, long run high gray level emphasis, and coarseness had an effect on predicting ISS (Fig. 4). We could obtain force plots for each patient through SHAP analysis: for example, one force plot is shown in Fig. 5. The low value for joint energy pushed this case to the left side; however, the low values for high gray level zone emphasis and coarseness and the high value for skewness push the prediction to right side, which lead to high possibility of GHD.

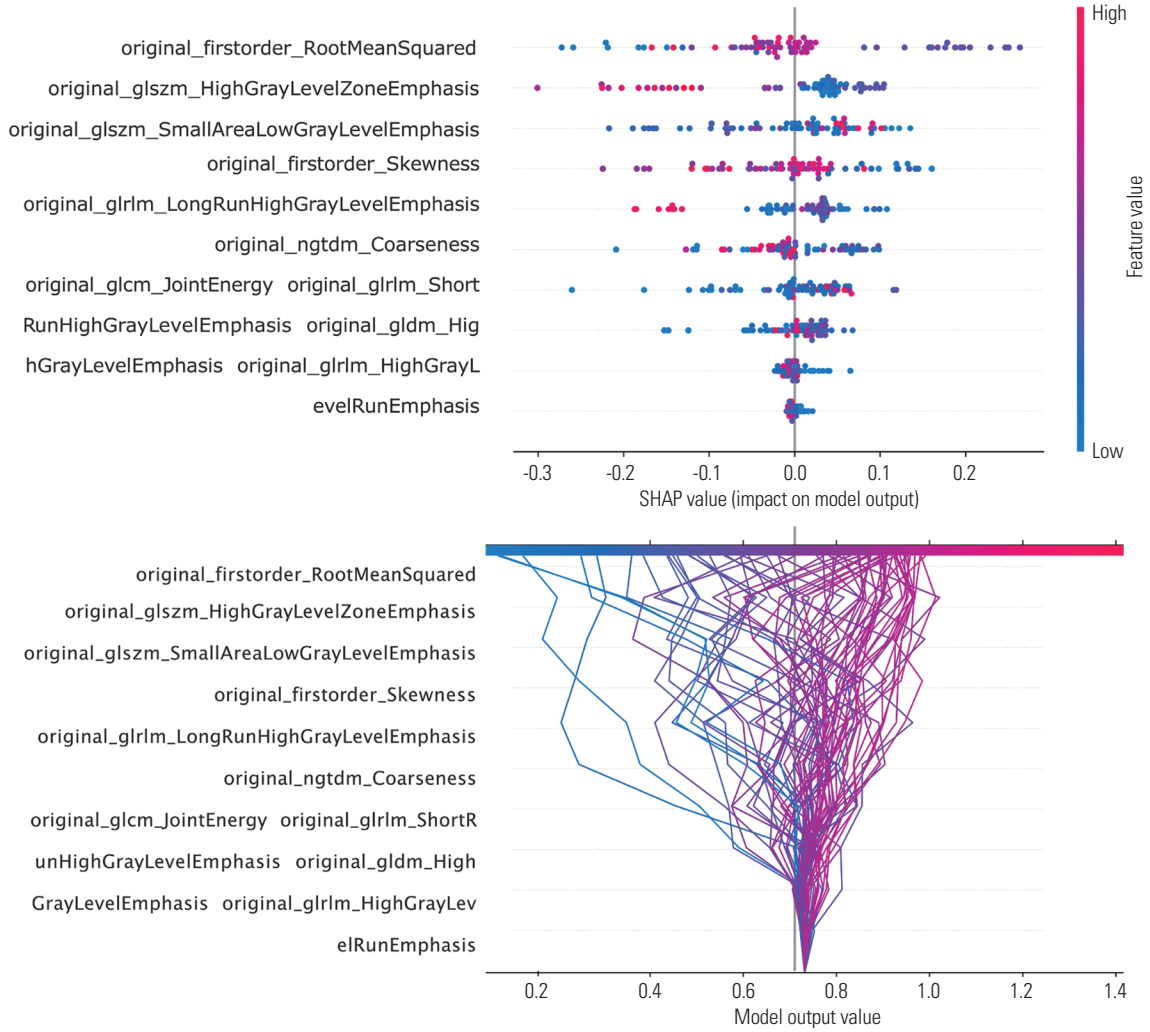

Fig. 4. Feature importance presented by mean absolute Shapley additive explanations (SHAP) value dot plot (upper) and decision plot (lower) for the prediction of growth hormone deficiency or idiopathic short stature in the developed model. Color shows whether the parameter was high or low for that row of the patient dataset. Horizontal location shows whether the effect of that value elicited a higher or lower prediction.

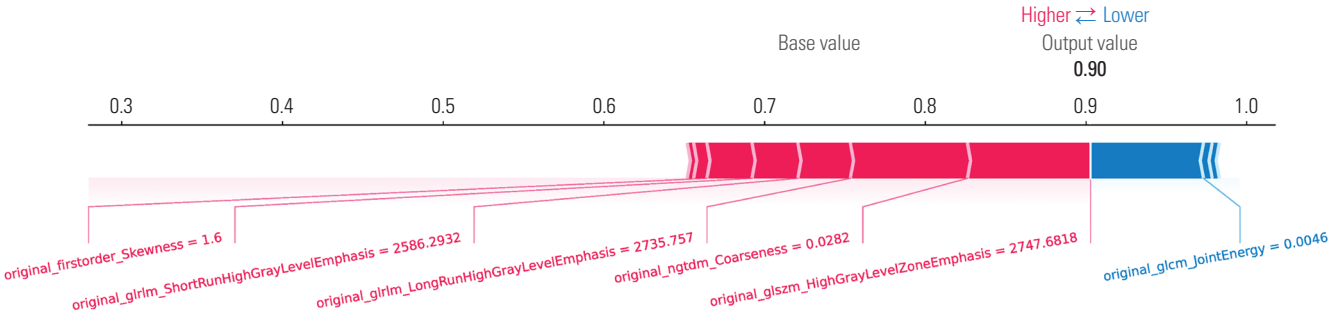

Fig. 5. Representative force plot of one case. The low value for joint energy pushes the prediction of this case to the left side (high possibility of ISS). However, low values for high gray level zone emphasis and coarseness and the high value for skewness push the prediction to the right side more strongly, which indicates to high possibility of GHD. This case was confirmed as GHD. ISS, idiopathic short stature; GHD, growth hormone deficiency.

https://doi.org/10.3349/ymj.2022.63.9.856

Table 2. Comparison of Size Parameters for the Pituitary Gland Between the GHD and ISS Groups

Group	Mean±SD	*p* value
Axial diameter		0.262
GHD	11.89±1.967	
ISS	12.27±4.045	
Coronal diameter		0.235
GHD	12.17±1.901	
ISS	12.56±4.019	
Sagittal height		0.014
GHD	6.12±1.022	
ISS	6.45±1.262	
Volume		0.050
GHD	161.45±58.26	
ISS	176.47±72.87	

GHD, growth hormone deficiency; ISS, idiopathic short stature.
Data were calculated using t-tests.

Finally, volumetric and diameter analysis between the GHD and ISS groups was performed. The mean (standard deviation) volume values were 176.47 mm^3 (72.87 mm^3) and 161.45 mm^3 (58.26 mm^3) in the GHD and ISS group, respectively (p=0.050). The mean (standard deviation) sagittal height values were 6.45 mm (1.26 mm) and 6.12 mm (1.02 mm), respectively (p=0.014). Other results are listed in Table 2.

DISCUSSION

Patients with GHD sometimes have a structural abnormality on sella MRI, resulting in decreased GH secretion. However, many cases do not have abnormal MRI findings. In an investigation in our hospital, approximately 72% of children with GHD had normal MRI findings.[27] In these normal MRIs, there may be microscopic organic lesions that cannot be detected by the human eye. A previous investigation revealed that GHD patients with morphologically normal pituitary glands had impaired perfusion of adenohypophysis on perfusion MRI.[35] The present study was conducted under the hypothesis that GH secretion impairment may be due to microscopic abnormalities that are not visible under conventional MRI resolution and could potentially be detected with a radiomics approach.

To differentiate between GHD and ISS, children must be hospitalized. Additionally, provocation tests require approximately 10 venous blood samples and include various pharmacologic stimuli, such as insulin, clonidine, glucagon, arginine, levodopa, clonidine, or GHRH.[8,36] The insulin test poses a risk of hypoglycemia.[37] Levodopa and arginine can induce vomiting as an adverse effect.[38,39] As such, diagnosis based on other exams that obviate the need for stimulation tests could save time and costs and ease the patient's burden. One investigation has aimed to diagnose and classify GHD with one blood sample using transcriptomics and random forest analysis.[40] However, sella MRI-based approaches to classify children into GHD and ISS are very rare. Most radiomics-based pituitary gland studies of sella MRI have been attempted in adults with pituitary tumors, and none have been tested in children. The present study is the first radiomics analysis performed on a sample of patients with short stature diagnosed with GHD and ISS.

In this study, we used LightGBM, a tree-based machine learning method. Tree-based machine learning methods are relatively easy to understand, and they do not need feature scaling, such as standardization or normalization. Therefore, when there are feature variations, these models have the advantage of impeding the domination of high-range features. In addition, the LightGBM model has a relatively fast learning speed.[33] Therefore, we believe that the LightGBM method can be effectively used in radiomics-based machine learning studies.

In this study, we investigated the SHAP values of our model. The mean absolute Shapley values were introduced here to create an explainable radiomics model and to solve its notorious 'black box' nature.[34] In our study, RMS values were the most important feature and low values thereof predicted ISS. RMS is the square-root of the mean of all the squared intensity values. It is another measure of the magnitude of image values. This feature is volume-confounded such that a small volume of pituitary gland would lead to small RMS values.[31] We noted a tendency for ISS patients to have smaller RMS volume than GHD patients, and therefore, this may be related to the volume of the pituitary gland. The second important feature was high gray level zone emphasis (HGLZE), which reflected a greater proportion of higher gray-level values and size zones.[31] A gray level zone is defined as the number of connected voxels that share the same gray level intensity.[31] High values of HGLZE suggest relatively homogenous texture and were predictive of ISS in our model. We suspect that this may be related to the invisible, microscopic structure of the pituitary gland. In addition, our model showed that low skewness values were the fourth important feature to predict GHD. Possibly, this is related to tiny, invisible Rathke's cleft cyst, which has low voxel intensity and leads to low skewness.

In the comparison between the GHD and ISS groups for volume and diameter, the ISS group had significantly shorter sagittal heights than the GHD group. Also, the ISS group showed a tendency to have smaller pituitary gland volume than the GHD group. This result is different from previous literature.[17] We measured volume with three dimensional ROI, which is probably more accurate than estimates obtained using the ellipsoid formula, which was used in previous literature.[17] Meanwhile, the clinical model developed with only volumetric and demographic data showed disappointing performance. We believe radiomics features may have potential and additional value over conventional parameters in discriminating between GHD and ISS.

The present research has several limitations. First, our study was conducted at a single institution, and validation in an ex-

ternal test set was not attempted. However, in this study, by setting the temporally split test set (so called 'temporal external validation set'), we attempted to correct for overfitting and obtain more objective results.[41] The random division of the patients induces falsely improved diagnostic performance as the distribution of disease exists in the same proportion in both training and validation sets. Temporal validation in the same institution is a better method for independent validation. Second, an AUC of approximately 0.705 and 70.6% accuracy were obtained in the test set. A sufficiently high value to replace the provocation test has not yet been obtained, and its reliability has not yet been verified. For now, radiomics models using sella MRI are rather unlikely to replace GH stimulation test. However, we showed the feasibility of a radiomics approach in sella MRI in children. Now we have to perform further validation study and increase the performance of the model to expand the role of sella MRI in this group of patients not only to exclude other structural cause, but also to predict GHD or ISS. Also, biologic correlates and evidence-based interpretation of the physiologic meaning of radiomics features were not sufficiently done in this investigation. Finally, only coronal T2WI were used, and we did not investigate the added value of the other sequences. Due to the slice thickness and size of FOV, we chose coronal T2WI. In order to incorporate T1WI in image processing, the protocols should be modified to improve spatial resolution by having smaller FOV and thinner thickness. Although coronal T2WI used in this study are thinner than other sequences, still, a 1.5-mm thickness is not thin enough to capture all information in pituitary glands. Therefore, the prediction model would be more powerful if more features from other sequences, including 3D sequences, were included. We are now planning future study with larger, multicenter populations including contrast enhanced T1WI to confirm, expand, and validate our results.

The present study has several strengths. First, we reviewed a relatively large sample size of patients (n=312) enrolled over nearly 10 years. Follow-up data, such as treatment response, could be possible as further study. Second, the application of a radiomics prediction model to pediatric sella MRI is a relatively novel approach. Previous investigations that have applied radiomics to the sella MRI of children are rare. Finally, by SHAP analysis, we attempted to develop an explainable radiomics model, and this allowed us to speculate on the possibility of microlesions in the pituitary gland.

In conclusion, in normal sella MRI of pediatric patients of short stature, we found it feasible to use a radiomics-based prediction model to differentiate between GHD and ISS.

ACKNOWLEDGEMENTS

This research was supported by a faculty research grant from Yonsei University College of Medicine (6-2021-0151).

AUTHOR CONTRIBUTIONS

Conceptualization: Beomseok Sohn. Data curation: Kyungchul Song, Ho-Seong Kim, and Beomseok Sohn. Formal analysis: Beomseok Sohn and Jihwan Eom. Funding acquisition: Beomseok Sohn. Investigation: Taeyoun Lee and Beomseok Sohn. Methodology: Sung Soo Ahn, Jihwan Eom, and Beomseok Sohn. Project administration: Beomseok Sohn. Resources: Sung Soo Ahn, Kyungchul Song, and Ho-Seong Kim. Software: Jihwan Eom. Supervision: Seung-Koo Lee. Validation: Beomseok Sohn and Jihwan Eom. Visualization: Beomseok Sohn and Jihwan Eom. Writing—original draft: Taeyoun Lee and Beomseok Sohn. Writing—review & editing: Taeyoun Lee, Kyungchul Song, Beomseok Sohn, Sung Soo Ahn, Ho-Seong Kim, and Seung-Koo Lee. Approval of final manuscript: all authors.

ORCID iDs

Taeyoun Lee https://orcid.org/0000-0003-1006-5759
Kyungchul Song https://orcid.org/0000-0002-8497-5934
Beomseok Sohn https://orcid.org/0000-0002-6765-8056
Jihwan Eom https://orcid.org/0000-0002-3793-4750
Sung Soo Ahn https://orcid.org/0000-0002-0503-5558
Ho-Seong Kim https://orcid.org/0000-0003-1135-099X
Seung-Koo Lee https://orcid.org/0000-0001-5646-4072

REFERENCES

1. Growth Hormone Research Society. Consensus guidelines for the diagnosis and treatment of growth hormone (GH) deficiency in childhood and adolescence: summary statement of the GH Research Society. J Clin Endocrinol Metab 2000;85:3990-3.
2. Melmed S. Pathogenesis and diagnosis of growth hormone deficiency in adults. N Engl J Med 2019;380:2551-62.
3. Webb SM, Rigla M, Wägner A, Oliver B, Bartumeus F. Recovery of hypopituitarism after neurosurgical treatment of pituitary adenomas. J Clin Endocrinol Metab 1999;84:3696-700.
4. Martha PM Jr, Gorman KM, Blizzard RM, Rogol AD, Veldhuis JD. Endogenous growth hormone secretion and clearance rates in normal boys, as determined by deconvolution analysis: relationship to age, pubertal status, and body mass. J Clin Endocrinol Metab 1992;74:336-44.
5. Hanew K, Utsumi A. The role of endogenous GHRH in arginine-, insulin-, clonidine- and l-dopa-induced GH release in normal subjects. Eur J Endocrinol 2002;146:197-202.
6. Rosenfeld RG. Is growth hormone deficiency a viable diagnosis? J Clin Endocrinol Metab 1997;82:349-51.
7. Bidlingmaier M. Problems with GH assays and strategies toward standardization. Eur J Endocrinol 2008;159 Suppl 1:S41-4.
8. Stanley T. Diagnosis of growth hormone deficiency in childhood. Curr Opin Endocrinol Diabetes Obes 2012;19:47-52.
9. Ranke MB. The Kabi Pharmacia International Growth Study: aetiology classification list with comments. Acta Paediatr Scand Suppl 1991;379:87-92.
10. Cohen P, Rogol AD, Deal CL, Saenger P, Reiter EO, Ross JL, et al. Consensus statement on the diagnosis and treatment of children with idiopathic short stature: a summary of the Growth Hormone Research Society, the Lawson Wilkins Pediatric Endocrine Society, and the European Society for Paediatric Endocrinology Workshop. J Clin Endocrinol Metab 2008;93:4210-7.
11. Naderi F, Eslami SR, Mirak SA, Khak M, Amiri J, Beyrami B, et al. Effect of growth hormone deficiency on brain MRI findings among

children with growth restrictions. J Pediatr Endocrinol Metab 2015;28:117-23.
12. Nagel BH, Palmbach M, Petersen D, Ranke MB. Magnetic resonance images of 91 children with different causes of short stature: pituitary size reflects growth hormone secretion. Eur J Pediatr 1997;156:758-63.
13. Arends NJ, V d Lip W, Robben SG, Hokken-Koelega AC. MRI findings of the pituitary gland in short children born small for gestational age (SGA) in comparison with growth hormone-deficient (GHD) children and children with normal stature. Clin Endocrinol (Oxf) 2002;57:719-24.
14. Bordallo MA, Tellerman LD, Bosignoli R, Oliveira FF, Gazolla FM, Madeira IR, et al. Neuroradiological investigation in patients with idiopathic growth hormone deficiency. J Pediatr (Rio J) 2004;80:223-8.
15. Maghnie M, Ghirardello S, Genovese E. Magnetic resonance imaging of the hypothalamus-pituitary unit in childrensuspected of hypopituitarism: who, how and when toinvestigate. J Endocrinol Invest 2004;27:496-509.
16. Ariza Jiménez AB, Martínez Aedo Ollero MJ, López Siguero JP. Differences between patients with isolated GH deficiency based on findings in brain magnetic resonance imaging. Endocrinol Diabetes Nutr (Engl Ed) 2020;67:78-88.
17. Kessler M, Tenner M, Frey M, Noto R. Pituitary volume in children with growth hormone deficiency, idiopathic short stature and controls. J Pediatr Endocrinol Metab 2016;29:1195-200.
18. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016;278:563-77.
19. Kickingereder P, Andronesi OC. Radiomics, metabolic, and molecular MRI for brain tumors. Semin Neurol 2018;38:32-40.
20. Rui W, Wu Y, Ma Z, Wang Y, Wang Y, Xu X, et al. MR textural analysis on contrast enhanced 3D-SPACE images in assessment of consistency of pituitary macroadenoma. Eur J Radiol 2019;110:219-24.
21. Zhang Y, Chen C, Tian Z, Cheng Y, Xu J. Differentiation of pituitary adenoma from Rathke cleft cyst: combining MR image features with texture features. Contrast Media Mol Imaging 2019;2019:6584636.
22. Fan Y, Chai Y, Li K, Fang H, Mou A, Feng S, et al. Non-invasive and real-time proliferative activity estimation based on a quantitative radiomics approach for patients with acromegaly: a multicenter study. J Endocrinol Invest 2020;43:755-65.
23. Galm BP, Buckless C, Swearingen B, Torriani M, Klibanski A, Bredella MA, et al. MRI texture analysis in acromegaly and its role in predicting response to somatostatin receptor ligands. Pituitary 2020;23:212-22.
24. Machado LF, Elias PCL, Moreira AC, Dos Santos AC, Murta Junior LO. MRI radiomics for the prediction of recurrence in patients with clinically non-functioning pituitary macroadenomas. Comput Biol Med 2020;124:103966.
25. Peng A, Dai H, Duan H, Chen Y, Huang J, Zhou L, et al. A machine learning model to precisely immunohistochemically classify pituitary adenoma subtypes with radiomics based on preoperative magnetic resonance imaging. Eur J Radiol 2020;125:108892.
26. Zhang Y, Chen C, Tian Z, Xu J. Discrimination between pituitary adenoma and craniopharyngioma using MRI-based image features and texture features. Jpn J Radiol 2020;38:1125-34.
27. Song KC, Jin SL, Kwon AR, Chae HW, Ahn JM, Kim DH, et al. Etiologies and characteristics of children with chief complaint of short stature. Ann Pediatr Endocrinol Metab 2015;20:34-9.
28. Yoon JY, Cheon CK, Lee JH, Kwak MJ, Kim HJ, Kim YJ, et al. Response to growth hormone according to provocation test results in idiopathic short stature and idiopathic growth hormone deficiency. Ann Pediatr Endocrinol Metab 2022;27:37-43.
29. Ahn J, Oh J, Suh J, Song K, Kwon A, Chae HW, et al. Next-generation sequencing-based mutational analysis of idiopathic short stature and isolated growth hormone deficiency in Korean pediatric patients. Mol Cell Endocrinol 2022;544:111489.
30. Avants BB, Tustison N, Song G. Advanced normalization tools (ANTS). Insight J 2009;2:1-35.
31. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. Cancer Res 2017;77:e104-7.
32. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc Series B Stat Methodol 1996;58:267-88.
33. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. Lightgbm: a highly efficient gradient boosting decision tree. In: von Luxburg U, Guyon I, Bengio S, Wallach H, Fergus R, Vishwanathan SVN, et al, editors. Advances in Neural Information Processing Systems 30. Red Hook: Curran Associates, Inc.; 2017. p.3147-55.
34. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: von Luxburg U, Guyon I, Bengio S, Wallach H, Fergus R, Vishwanathan SVN, et al, editors. Advances in Neural Information Processing Systems 30. Red Hook: Curran Associates, Inc.; 2017. p.4766-75.
35. Wang CY, Chung HW, Cho NY, Liu HS, Chou MC, Kao HW, et al. Idiopathic growth hormone deficiency in the morphologically normal pituitary gland is associated with perfusion delay. Radiology 2011;258:213-21.
36. Kim JH, Chae HW, Chin SO, Ku CR, Park KH, Lim DJ, et al. Diagnosis and treatment of growth hormone deficiency: a position statement from Korean Endocrine Society and Korean Society of Pediatric Endocrinology. Endocrinol Metab (Seoul) 2020;35:272-87.
37. Rhee N, Oh KY, Yang EM, Kim CJ. Growth hormone responses to provocative tests in children with short stature. Chonnam Med J 2015;51:33-8.
38. Porter BA, Rosenfield RL, Lawrence AM. The levodopa test of growth hormone reserve in children. Am J Dis Child 1973;126:589-92.
39. AbdelNabi R, Al Khalifah R. Growth hormone stimulation test with clonidine and arginine an unreported side effect. J Clin Transl Endocrinol: Case Rep 2020;15:100055.
40. Murray PG, Stevens A, De Leonibus C, Koledova E, Chatelain P, Clayton PE. Transcriptomics and machine learning predict diagnosis and severity of growth hormone deficiency. JCI Insight 2018;3:e93247.
41. Welch ML, McIntosh C, Haibe-Kains B, Milosevic MF, Wee L, Dekker A, et al. Vulnerabilities of radiomic signature development: the need for safeguards. Radiother Oncol 2019;130:2-9.