

Development of a predictive model of growth hormone deficiency and idiopathic short stature in children

MENGDI CONG¹, SHI QIU², RONGPIN LI¹, HAIYAN SUN¹, LINING CONG³ and ZHENZHOU HOU¹

¹Department of Computed Tomography and Magnetic Resonance, Children's Hospital of Hebei Province, Shijiazhuang, Hebei 050031; ²Key Laboratory of Spectral Imaging Technology Chinese Academy of Science, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, Shaanxi 710119; ³Department of Radiology, Children's Hospital of Hebei Province, Shijiazhuang, Hebei 050031, P.R. China

Received February 18, 2020; Accepted August 19, 2020

DOI: 10.3892/etm.2021.9925

Abstract. The aim of the present study was to develop predictive models using clinical features and MRI texture features for distinguishing between growth hormone deficiency (GHD) and idiopathic short stature (ISS) in children with short stature. This retrospective study included 362 children with short stature from Children's Hospital of Hebei Province. GHD and ISS were identified via the GH stimulation test using arginine. Overall, there were 190 children with GHD and 172 with ISS. A total of 57 MRI texture features were extracted from the pituitary gland region of interest using C++ language and Matlab software. In addition, the laboratory examination data were collected. Receiver operating characteristic (ROC) regression curves were generated for the predictive performance of clinical features and MRI texture features. Logistic regression models based on clinical and texture features were established for discriminating children with GHD and ISS. Two clinical features [IGF-1 (insulin growth factor-1) and IGFBP-3 (IGF binding protein-3) levels] were used to build the clinical predictive model, whereas the three best MRI textures were used to establish the MRI texture predictive model. The ROC analysis of the two models revealed predictive performance for distinguishing GHD from ISS. The accuracy of predicting ISS from GHD was 64.5% in ROC analysis [area under the curve (AUC), 0.607; sensitivity, 57.6%; specificity, 72.1%] of the clinical model. The accuracy of predicting ISS from GHD was 80.4% in ROC analysis (AUC, 0.852; sensitivity, 93.6%; specificity, 65.8%) of the MRI texture predictive model. In conclusion, these findings indicated that a texture predictive model using MRI texture features was superior

for distinguishing children with GHD from those with ISS compared with the model developed using clinical features.

Introduction

Growth hormone deficiency (GHD) is caused by a pituitary problem and is a condition in which the body does not produce enough GH. Idiopathic short stature (ISS) refers to short stature of unknown origin. The term ISS is defined as the absence of a dysfunction in the GH/insulin like growth factor (IGF) axis, or other identifiable disorder of the endocrine, genetic or organ systems in short children (1). GH stimulation tests remain the gold standard for determining the ability of a child to produce adequate amounts of GH (2). However, the tolerance of children and their families to the GH stimulation test is low because it requires frequent blood sampling. In addition, there are a series of adverse reactions and risks associated with the GH stimulation test, which makes diagnosis difficult (3). Since the treatment guidelines for children with GHD and ISS are different, and the dose of GH required for children with GHD differs compared with that required for children with ISS, it is necessary to distinguish between the two causes of short stature before treatment (4). The pituitary gland diseases should also be excluded before treatment; therefore, MRI examination is necessary. Furthermore, neuroimaging has become an important part of the diagnosis of GHD in children (5).

Previous studies have demonstrated that it is feasible and necessary to distinguish children with ISS from those with GHD; these previous studies have focused on the 3D volume of the pituitary gland. Han *et al* (6) studied the pituitary glands of 23 children with ISS, 32 children with GHD and 75 normal children. Compared with the normal children, 65.6% of the children with GHD and 34.8% of the children with ISS had smaller pituitary gland volumes. Kessler *et al* (7) reported that the average pituitary volume of children with GHD was $230.8 \pm 89.6 \text{ mm}^3$, for children with ISS it was $286.8 \pm 108.2 \text{ mm}^3$ and in the healthy group it was $343.7 \pm 145.9 \text{ mm}^3$ ($P < 0.001$). However, to the best of our knowledge, no previous study has focused on the MRI textures of the pituitary gland. The present study aimed to develop predictive models using clinical parameters and MRI texture features, in order to distinguish between children with GHD and ISS.

Correspondence to: Professor Zhenzhou Hou, Department of Computed Tomography and Magnetic Resonance, Children's Hospital of Hebei Province, 133 Jianhua South Street, Shijiazhuang, Hebei 050031, P.R. China
E-mail: houzhenzhou@yeah.net

Key words: growth hormone deficiency, idiopathic short stature, magnetic resonance imaging, prediction model, texture

Materials and methods

Children selection. The present retrospective analysis was approved by the ethical review board of Children's Hospital of Hebei Province (Shijiazhuang, China) and the need for informed consent was waived. Patients were retrospectively selected from hospital records between December 2018 and February 2019. Inclusion criteria were as follows: i) Children whose height was <2 standard deviations for the same age and sex; ii) children with subnormal growth velocity for ≤ 6 months; iii) children that underwent GH stimulation test; iv) children whose pituitary gland MRI was normal. The exclusion criteria included: i) Children that were or are using GH; ii) children that had intracranial abnormalities, such as malignant tumors, hypothalamic hamartoma or Arnold Chiari malformation; ii) children that had chronic liver and kidney diseases, skeletal diseases, congenital heart disease, thyroid hormone axis abnormalities and chromosomal abnormalities; iv) children whose MRI scans had artifacts or could not be outlined in the sagittal plane.

GHD was defined as peak serum GH levels <10 ng/ml in the GH stimulation tests, and ISS was defined as serum GH levels ≥ 10 ng/ml in the GH stimulation tests.

GH stimulation tests. For GH stimulation tests, the children fasted after 8 p.m. and blood was collected from the median cubital vein in the arm for basic GH tests in bed the next morning. The dosage of arginine hydrochloride was 0.5 g/kg, which was diluted with water for injection to a total of 200 ml and infused intravenously within 30 min. At 30, 60, 90 and 120 min after administration, blood was collected from the other arm to measure serum GH which was tested by the laboratory physician in Department of Laboratory Medicine, Children's Hospital of Hebei Province.

The recruitment pathway for the 362 children included in the present study is shown in Fig. 1. All of the recruited patients were children with short stature. The 362 children included 235 male patients and 127 female patients, with an average age of 8.94 years (range: 5-13 years). There were 190 children with GHD and 172 children with ISS.

MRI. MRI scans were performed on a Philips 1.5 Tesla MRI system (Achieva 1.5 T; Philips Medical Systems B.V.) using an 8-channel head coil. T1WI sagittal (TR/TE, 550/15 msec; FOV, 140x140 mm²; reconstruction matrix, 256; slice thickness, 3 mm; slice gap, 0 mm) images of the brain were obtained.

Segmentation and feature extraction. The region of interest (ROI) from the pituitary gland was segmented by a radiologist with 6 years of experience using ITK-SNAP software (version 3.6.0; www.itksnap.org). The ROI included the largest level of the sagittal T1 non-enhanced sequence, which displayed the pituitary stalk, as shown in Fig. 2. A total of 57 MRI textures (12 histogram features, nine form factor features and 36 grey level co-occurrence matrix features) were extracted from the pituitary gland ROI using Matlab R2014a software (MathWorks) and C++ language was used to write features.

To assess the intra- and interobserver reproducibility of MRI texture features, A total of 50 children were randomly

selected from the 362 children for pituitary gland segmentation 1 month later by the same radiologist and another radiologist with 10 years of experience in diagnosis. They both had no information or knowledge of the clinical and laboratory data. Subsequently, an intragroup consistency analysis was performed on the 50 images drawn by the same radiologist, and inter-group consistency analysis was performed on the 50 images drawn by the two radiologists.

The quantitative features available for analysis in this study were 57 MRI textures, age, IGF-1, IGF binding protein (IGFBP)-3, serum calcium and serum alanine aminotransferase (ALT). All the blood examinations were routine procedures performed by the laboratory in the Children's Hospital of Hebei Province. The qualitative feature was sex.

Statistical analysis. Continuous variables are presented as mean \pm SD, whereas categorical variables are presented as counts. The differences in continuous variables were analyzed through unpaired Student's t-test, including 57 MRI textures, age, IGF-1, IGFBP-3, serum calcium and serum ALT. Differences in categorical variables were analyzed with the χ^2 test, including sex. $P < 0.05$ was considered to indicate a statistically significant difference. Intraobserver and interobserver consistency was evaluated by intraclass correlation coefficient (ICC) using the absolute agreement method. The property of the model was determined by applying the receiver operating characteristic (ROC) curve analysis. The accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were calculated at a cutoff value (the maximum value of Youden index) to evaluate the efficiency of the predictive model. Statistical analysis was performed using SPSS software version 21.0 (IBM Corp.). The flow chart of the analysis is shown in Fig. 3.

Results

Characteristics of the children recruited. The characteristics of all of the children recruited are detailed in Table I. There were 362 children with short stature recruited in the present study.

Feature selection and performance in the predictive model. Two clinical features, IGF-1 and IGFBP-3, were chosen to establish the clinical model since the P-values of these two clinical features were statistically different between children with ISS and GHD ($P < 0.05$; Table I). The ROC curve was as shown in Fig. 4. The area under the curve (AUC) value of the clinical model in differentiating ISS from GHD was 0.607 (95% CI, 0.546-0.667), with an accuracy of 0.645, a sensitivity of 0.576, a specificity of 0.721, a PPV of 0.279 and an NPV of 0.424.

Three MRI texture features: Fractal dimension, normalized entropy and fourth-order moment (kurtosis), were chosen to establish the texture model since the P-values of the three features were statistically different ($P < 0.05$; Table II). The ROC curve was shown in Fig. 5. The AUC value of the MRI texture model in differentiating ISS from GHD was 0.852 (95% CI, 0.813-0.891), with an accuracy of 0.804, a sensitivity of 0.936, a specificity of 0.658, a PPV of 0.751 and an NPV of 0.903. Fractal dimension mainly describes the most important

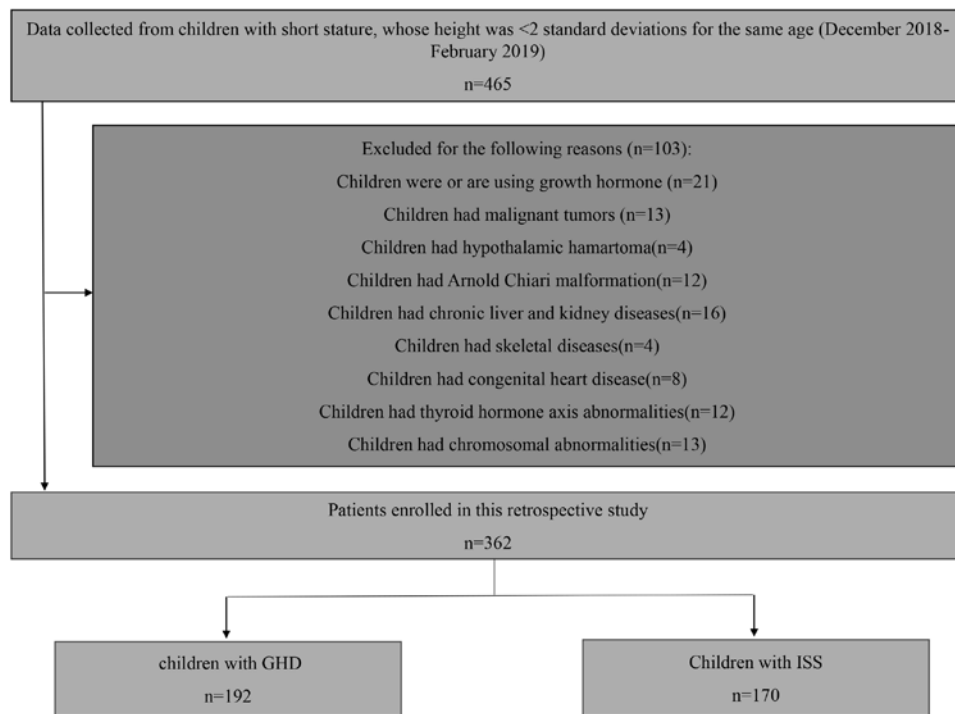


Figure 1. Recruitment pathway for patients in the present study. GHD, growth hormone deficiency; ISS, idiopathic short stature.



Figure 2. Midline sagittal images of the pituitary gland in a child with growth hormone deficiency. The red area represents the region of interest.

fractal parameters, and the similarity of textures or structures obtained by random processes. Normalized entropy is the value of entropy after normalization. Entropy value expresses the asymmetry of image histogram distribution. The higher the entropy value is, the more asymmetrical the histogram (lesion) distribution is; the smaller the value is, the more symmetrical the histogram distribution is. The fourth-order moment (kurtosis) feature describes the shape of the flatness or peak of the probability density function. The peak value of the peak distribution is greater than that of the plane distribution.

Formula results and interpretation. The clinical model prediction probability of ISS can be calculated using the following formula: $\text{Ln}[P/(1-P)] = -1.194 + 0.001 \times \text{IGF-1} + 0.213 \times \text{IGFBP3}$. The texture model prediction probability of ISS can be calculated using the following formula: $\text{Ln}[P/(1-P)] = -11.352 + 8.853 \times \text{fractal dimension} - 5.189 \times \text{normalized entropy} + 3.595 \times \text{fourth-order moment (kurtosis)}$. Firstly, the P-value was selected,

then the features for which the P-value was <0.05 were selected. Secondly, binary logistic regression analysis was used to calculate the coefficient of each feature. Finally, the formula $\text{Ln}[P/(1-P)] = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$ was used; α represents the constant, β_n represents the B value and X_n represents each feature.

The pituitary gland image was inputted into Matlab software for one new child with short stature for whom it was not known whether they had GHD or ISS. The fractal dimension, normalized entropy and kurtosis of the image were extracted. The IGF-1 and IGFBP3 of the child were inputted into the clinical model formula, and the fractal dimension, normalized entropy and kurtosis of the child were inputted into the texture model formula; the P-value was obtained through the aforementioned calculation. When the P-value was >0.5 , the child was considered more likely to have ISS, and when the P-value was <0.5 , the child was considered more likely to have GHD. SPSS software defined the threshold as 0.5.

Consistency test. The intraobserver and interobserver ICCs among the 57 MRI texture features ranged between 0.809 and 0.967 (data not shown). These findings indicated that the intra- and interobserver consistency was reliable.

Discussion

To the best of our knowledge, the present study is the first to use predictive models to distinguish GHD from ISS in children. This is particularly noteworthy, because the present study did not assess children with GHD and ISS according to pituitary volume or height, as numerous previous studies have done (1,6-8). In the present study, a predictive model was established that used MRI texture features to distinguish

Table I. Clinical features of children with GHD and ISS.

Variable	GHD	ISS	P-value
Age, years	9.07±2.59	8.61±2.76	0.101
Sex			0.246
Male	138	134	
Female	52	38	
ALT, IU	16.80±12.17	15.30±8.77	0.177
Ca, mmol/l	1.60±0.08	1.61±0.08	0.85
IGF-1, ng/ml	186.33±81.94	206.94±106.35	0.041
IGFBP-3, μ g/ml	3.89±0.84	4.11±0.94	0.018

Differences in continuous variables were analyzed through the unpaired Student's t-test, including age, IGF-1, IGFBP-3, Ca and ALT. Differences in categorical variables were analyzed using χ^2 test, including sex. GHD, growth hormone deficiency; ISS, idiopathic short stature; IGF-1, insulin growth factor 1; IGFBP-3, IGF binding protein 3; ALT, alanine aminotransferase.

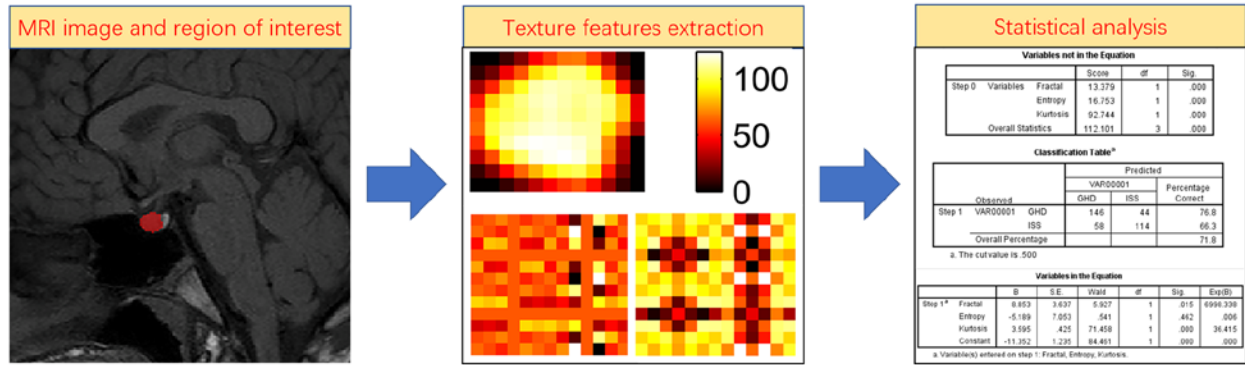


Figure 3. Flow chart of the present study. Firstly, the radiologist drew the ROI of the pituitary gland (labeled area). Secondly, 57 MRI textures (12 histogram features, nine form factor features and 36 grey level co-occurrence matrix features) were extracted from the pituitary gland ROI using Matlab software, and C++ language was used to write features. Thirdly, the P-value was calculated and the features for which the P-value was <0.05 were selected. Fourthly, binary logistic regression analysis was used to calculate the coefficient of each feature. Finally, the formula $\ln[P/(1-P)] = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$ was used; α represents the constant, β_n represents the B value and X_n represents each feature. ROI, region of interest.

between children with GHD and those with ISS. This MRI texture prediction model was superior to the clinical prediction model for distinguishing between children with GHD and ISS. The ROC analysis indicated that the predictive AUC values had good predictive capacity. In the present study, the children with ISS had higher IGF-1 and IGFBP-3 levels than the children with GHD, and the difference was statistically significant. This result was concordant with previous studies (9,10). IGF-1 is mainly synthesized and secreted by the liver, and its expression varies under the influence of various factors, including GH. GH has a direct effect on target cells, stimulating the synthesis and secretion of IGF-1, and inhibiting the secretion of GH through negative feedback (11,12).

Fractal analysis is a type of mathematical structure, which can be used to evaluate and quantitatively analyze the texture or heterogeneity of tumors; therefore, it has great application value in brain tumors (13-16). Entropy refers to the uniformity of pixel value distribution in the image histogram. The higher the entropy is, the more evenly the image pixel value is distributed. Kurtosis refers to the approximate state when the pixel value of the image is close to the mean value. The smaller the peak, the more concentrated it is. Based on this, the present

study aimed to assess whether the features of the normal pituitary gland were related to its hormone secretion level. In the present study, the accuracy of using the three features to distinguish GHD from ISS was 80.4%. MRI texture analysis is a quantitative tool (17); through MRI texture analysis, more detailed and quantitative data of the ROI can be obtained compared with that obtained from the naked eye. Texture analysis is applied to the pituitary gland for several reasons. Pituitary and surrounding slope structures make the pituitary gland an ideal object to evaluate. There is a magnetic resonance signal difference between the pituitary gland and the slope, so it is helpful for the physician to sketch the pituitary gland. There is a difference in the pituitary gland which is hard to distinguish between GHD and ISS with the naked eye. The present study assessed children with short stature and normal pituitary glands; therefore, GHD and ISS could not be distinguished with the naked eye. Therefore, texture analysis was used to identify GHD and ISS. Another study, similar to the present study, has previously been published (18). This previous study reported that the visual image texture of EGFR mutations in patients with lung cancer could not be distinguished by the naked eye alone (18). Similarly, the present

Table II. One-way ANOVA analysis for differentiating GHD from ISS.

Variable	GHD	ISS	P-value
Fractal dimension	2.970±0.278	3.324±0.328	<0.001
Normalized entropy	0.058±0.015	0.066±0.023	<0.001
Fourth-order moment	0.018±0.023	0.070±0.193	0.001

GHD, growth hormone deficiency; ISS, idiopathic short stature.

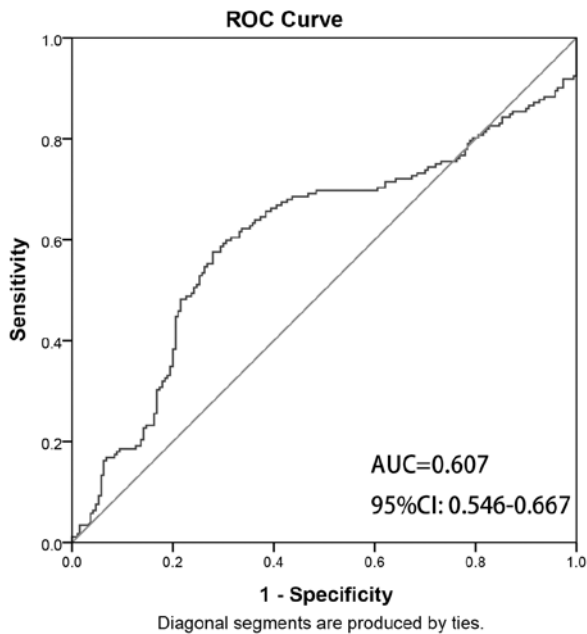


Figure 4. ROC curve of the clinical model. ROC, receiver operating characteristic; AUC, area under the curve.

study was unable to distinguish the visual images between children with GHD and ISS by the naked eye.

Children with short stature are often diagnosed by pediatric endocrinologists as having GHD or ISS. Cohen *et al* (19) reported that ISS may represent a partial GH-insensitive state that manifests during treatment with higher doses of GH. To distinguish between the two conditions, endocrinologists rely on the GH stimulation test; however, this test requires frequent blood sampling of children. More likely, GH production is continuous, and although some patients classified as ISS can show normal GH production in GH stimulation trials, they may not be able to consistently produce enough GH to support optimal linear growth under physiological conditions. Therefore, it is particularly important to distinguish GHD from ISS.

The present study has some limitations. Firstly, this is a retrospective study, and the number of children with short stature was small. In future studies, we aim to include more children with short stature. Secondly, this study only selected the maximum level of the sagittal position of the T1WI adenohypophysis, which may result in the loss of part of the MRI information. Finally, this procedure is not suitable for children who could not undergo the GH stimulation test; this situation could not be overcome. Because there were a series of problems in the pituitary gland of the child (such as pituitary stalk

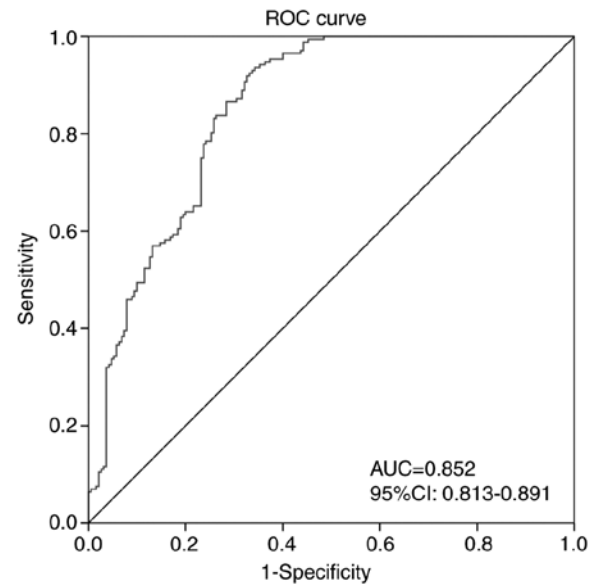


Figure 5. ROC curve of the MRI texture predictive model. ROC, receiver operating characteristic; AUC, area under the curve.

interruption syndrome and pituitary adenoma), which were mentioned in the inclusion and exclusion criteria, and the GH stimulation test could not be performed, it was impossible to distinguish whether the child had GHD or ISS.

In conclusion, the present study indicated that a predictive model using MRI texture features may have potential for differentiating between GHD and ISS in children.

Acknowledgements

Not applicable.

Funding

No funding was received.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

MC collected clinical data on short stature, MRI and outlined the ROI, and was a major contributor in writing the manuscript.

SQ used the software to extract the texture features of the ROI. RL made substantial contributions to design and conception of the study. HS outlined ROI of the pituitary gland. LC collected the clinical and imaging data. ZH made substantial contributions to conception and design of the study, and the acquisition, analysis and interpretation of data. MC and SQ confirm the authenticity of all the raw data. All authors read and approved the final manuscript.

Ethics approval and consent to participate

This retrospective analysis was approved by the ethical review board of our hospital, and the need for informed consent was waived.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Pedicelli S, Peschiaroli E, Violi E and Cianfarani S: Controversies in the definition and treatment of idiopathic short stature (ISS). *J Clin Res Pediatr Endocrinol* 1: 105-115, 2009.
2. Wilson TA, Rose SR, Cohen P, Rogol AD, Backeljauw P, Brown R, Hardin DS, Kemp SF, Lawson M, Radovick S, *et al*: Update of guidelines for the use of growth hormone in children: The lawson wilkins pediatric endocrinology society drug and therapeutics committee. *J Pediatr* 143: 415-421, 2003.
3. Albrecht A, Penger T, Marx M, Hirsch K and Dörr HG: Short-term adverse effects of testosterone used for priming in prepubertal boys before growth hormone stimulation test. *J Pediatr Endocrinol Metab* 31: 21-24, 2018.
4. Grimberg A, DiVall S, Polychronakos C, Allen DB, Cohen LE, Quintos JB, Rossi WC, Feudtner C and Murad MH; Drug and Therapeutics Committee and Ethics Committee of the Pediatric Endocrine Society: Guidelines for growth hormone and insulin-like growth factor-i treatment in children and adolescents: Growth hormone deficiency, idiopathic short stature, and primary insulin-like growth factor-i deficiency. *Horm Res Paediatr* 86: 361-397, 2016.
5. Maghnie M, Lindberg A, Koltowska-Häggström M and Ranke MB: Magnetic resonance imaging of CNS in 15,043 children with GH deficiency in KIGS (Pfizer International Growth Database). *Eur J Endocrinol* 168: 211-217, 2013.
6. Han X, Xiu J, Huang Z, Zhang J, Zhang Z, Dong Y, Yuan X and Liu Q: Three-dimensional magnetic resonance volumetry of the pituitary gland is effective in detecting short stature in children. *Exp Ther Med* 8: 551-556, 2014.
7. Kessler M, Tenner M, Frey M and Noto R: Pituitary volume in children with growth hormone deficiency, idiopathic short stature and controls. *J Pediatr Endocrinol Metab* 29: 1195-1200, 2016.
8. Fink AM, Vidmar S, Kumbla S, Pedreira CC, Kanumakala S, Williams C, Carlin JB and Cameron FJ: Age-related pituitary volumes in prepubertal children with normal endocrine function: Volumetric magnetic resonance data. *J Clin Endocrinol Metab* 90: 3274-3278, 2005.
9. Ren S, Nie Y and Wang A: Effects of recombinant human growth hormone in the treatment of dwarfism and relationship between IGF-1, IGFBP-3 and thyroid hormone. *Exp Ther Med* 12: 3579-3582, 2016.
10. Wang Y, Zhang H, Cao M, Kong L and Ge X: Analysis of the value and correlation of IGF-1 with GH and IGFBP-3 in the diagnosis of dwarfism. *Exp Ther Med* 17: 3689-3693, 2019.
11. Clayton PE, Cuneo RC, Juul A, Monson JP, Shalet SM and Tauber M; European Society of Paediatric Endocrinology: Consensus statement on the management of the GH-treated adolescent in the transition to adult care. *Eur J Endocrinol* 152: 165-170, 2005.
12. Takeshita H, Fujihara J, Soejima M, Koda Y, Kimura-Kataoka K, Ono R, Yuasa I, Iida R, Ueki M, Nagao M and Yasuda T: Confirmation that SNPs in the high mobility group-A2 gene (HMGA2) are associated with adult height in the Japanese population; Wide-ranging population survey of height-related SNPs in HMGA2. *Electrophoresis* 32: 1844-1851, 2011.
13. Liu S, Fan X, Zhang C, Wang Z, Li S, Wang Y, Qiu X and Jiang T: MR imaging based fractal analysis for differentiating primary CNS lymphoma and glioblastoma. *Eur Radiol* 29: 1348-1354, 2019.
14. Czyz M, Radwan H, Li JY, Filippi CG, Tykocki T and Schuller M: Fractal analysis may improve the preoperative identification of atypical meningiomas. *Neurosurgery* 80: 300-308, 2017.
15. Marusina MY, Mochalina AP, Frolova EP, Satikov VI, Barchuk AA, Kuznetsov VI, Gaidukov VS and Tarakanov SA: MRI image processing based on fractal analysis. *Asian Pac J Cancer Prev* 18: 51-55, 2017.
16. Smitha KA, Gupta AK and Jayasree RS: Fractal analysis: Fractal dimension and lacunarity from MR images for differentiating the grades of glioma. *Phys Med Biol* 60: 6937-6947, 2015.
17. Herlidou-Même S, Constans JM, Carsin B, Olivie D, Eliat PA, Nadal-Desbarats L, Gondry C, Le Rumeur E, Idy-Peretti I and de Certaines JD: MRI texture analysis on texture test objects, normal brain and intracranial tumors. *Magn Reson Imaging* 21: 989-993, 2003.
18. Jia TY, Xiong JF, Li XY, Yu W, Xu ZY, Cai XW, Ma JC, Ren YC, Larsson R, Zhang J, *et al*: Identifying EGFR mutations in lung adenocarcinoma by noninvasive imaging using radiomics features and random forest modeling. *Eur Radiol* 29: 4742-4750, 2019.
19. Cohen P, Germak J, Rogol AD, Weng W, Kappelgaard AM and Rosenfeld RG; American Norditropin Study Group: Variable degree of growth hormone (GH) and insulin-like growth factor (IGF) sensitivity in children with idiopathic short stature compared with GH-deficient patients: Evidence from an IGF-based dosing study of short children. *J Clin Endocrinol Metab* 95: 2089-2098, 2010.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.