

Diagnostic accuracy of IGF-1 for screening growth hormone deficiency: a prospective, single-center, diagnostic accuracy study

Hideyuki Iwayama (✉ iwahide1976@gmail.com)

Aichi Medical University: Aichi Ika Daigaku <https://orcid.org/0000-0002-5218-5815>

Sachiko Kitagawa

Daiyukai Medical Corp: Shakai Iryo Hojin Daiyukai

Jyun Sada

Aichi Medical University: Aichi Ika Daigaku

Ryosuke Miyamoto

Aichi Medical University: Aichi Ika Daigaku

Tomohito Hayakawa

Aichi Medical University: Aichi Ika Daigaku

Yoshiyuki Kuroyanagi

Aichi Medical University: Aichi Ika Daigaku

Taichiro Muto

Aichi Medical University: Aichi Ika Daigaku

Hirokazu Kurahashi

Aichi Medical University: Aichi Ika Daigaku

Wataru Ohashi

Aichi Medical University: Aichi Ika Daigaku

Junko Takagi

Aichi Medical University: Aichi Ika Daigaku

Akihisa Okumura

Aichi Medical University: Aichi Ika Daigaku

Research Article

Keywords: Insulin-Like Growth Factor 1, Growth Hormone, ROC Curve, Area Under Curve, Sensitivity and Specificity

Posted Date: April 20th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-430014/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Purpose

We evaluated the diagnostic accuracy of insulin-like growth factor-1 (IGF-1) for screening growth hormone deficiency (GHD) to determine the usefulness of IGF-1 as a screening test.

Methods

On 298 consecutive children who had short stature or decreased height velocity, we measured IGF-1 levels and performed growth hormone (GH) secretion test using clonidine, arginine, and, in cases with different results of the two tests, L-dopa. Patients with congenital abnormalities were excluded. GHD was defined as peak GH \leq 6.0 ng/mL in the two tests.

Results

We identified 60 and 238 patients with and without GHD, respectively. The mean IGF-1 (SD) was not significantly different between the GHD and non-GHD groups ($p = 0.23$). Receiver operating characteristic curve analysis demonstrated the best diagnostic accuracy at an IGF-1 cutoff of -1.493 SD, with sensitivity of 0.685, specificity of 0.417, positive predictive value of 0.25, negative predictive value of 0.823, and area under the curve of 0.517. Spearman's rank correlation coefficient showed that IGF-1 (SD) was weakly correlated with age, bone age, height velocity before examination, weight (SD), and BMI (SD) and very weakly correlated with height (SD), target height (SD), and maximum GH peak.

Conclusion

IGF-1 level had poor diagnostic accuracy as a screening test for GHD. Correlation analysis revealed that none of the items increased the diagnostic power of IGF-1. Therefore, IGF-1 should not be used alone in the screening of GHD. A predictive biomarker for GHD should be developed in the future.

Introduction

Growth hormone (GH) secretion needs to be assessed for the diagnosis of GH deficiency (GHD) by stimulation tests. However, there are several challenges associated with the GH secretion test [1]. Pharmacological stimuli are not physiological, and their accuracy is poor. It is well known that normally growing children may have falsely low GH responses. Moreover, the diagnostic criteria for GHD are not uniform worldwide [2]. Furthermore, the GH secretion test may be influenced by factors such as obesity, undernutrition, sex, age, puberty, and presence of chronic diseases. It also has potential adverse reactions and may sometimes result in hospitalization. Therefore, a predictive biomarker for GHD is desired to avoid unnecessary GH secretion test.

Insulin-like growth factor-1 (IGF-1) is a small polypeptide hormone secreted by the liver when stimulated by GH. As serum levels of IGF-1 show little circadian variation, IGF-1 has been considered as a predictive biomarker for GHD [2]. The utility of IGF-1 for the screening of GHD was reported in some studies [3-7] but not in others [1, 8]. As the study settings in these reports were different, it is difficult to compare the diagnostic accuracy of IGF-1. For example, the inclusion criteria for GH secretion test comprise not only short stature but also bone age [2], target height [3,6], and catch-up growth [3]. Furthermore, the studies used different GH cutoff levels [1, 3-6, 8, 9]. Therefore, a prospective cohort study was required to determine the diagnostic accuracy of IGF-1. We prospectively analyzed a cohort of children with short stature to evaluate the diagnostic accuracy of IGF-1 for the diagnosis of GHD.

Materials And Methods

Patients

This was a prospective cross-sectional study on children with short stature or decreased growth velocity who were examined at Aichi Medical University Hospital between April 2015 and March 2020. We used the following inclusion criteria: (a) referred to Aichi Medical University for the evaluation of short stature or decreased growth velocity; (b) short stature of ≤ -2 SD or height velocity of ≤ -1.5 SD in >2 years below the mean for sex and age [10]; and (c) >1 year of age and before the completion of puberty, according to Tanner stages. The exclusion criteria were the presence of recognized congenital abnormalities, such as hypothyroidism; small for gestational age; Turner's syndrome; and trisomy 21.

General biochemical tests, thyroid function test, bone age, and IGF-1 were examined before GH secretion test in consecutive patients who met the inclusion criteria. The radius, ulna, and short bone method was used for evaluating bone age [11]. The patients were divided into GH and non-GH groups according to the response to the GH secretion test. In Japan, GHD is diagnosed when the peak GH is ≤ 6.0 ng/mL in two GH secretion tests [9]. A cutoff of 6 ng/mL was determined by the Japanese National Health Insurance program. Stimulation tests using clonidine, arginine, and L-dopa were performed in that order, using the algorithm shown in Figure 1. GHD was diagnosed if the GH peak levels were < 6 ng/mL in the two stimulation tests. If the GH peak was above the cutoff level in the clonidine stimulation test, the next stimulation test was not performed. If the GH peak of the arginine stimulation test was 6–8 ng/mL, the third L-dopa stimulation test was performed. If the GH peak of the arginine stimulation test was > 8 ng/mL, the third test was not performed as GHD was unlikely to be present. Glucagon and insulin were not used in this study because glucagon requires a long examination time of 180 min and insulin results in an adverse effect of severe hypoglycemia.

After overnight fasting, the stimulation test was started at 6:30 for children <6 years old and at 9:00 for those >6 years old because of fasting tolerance. Sampling was done at 0, 30, 60, 90, and 120 minutes. Clonidine (5 μ g/kg), arginine (10 mg/kg), and L-dopa (10 mg/kg) were administered as the stimuli for the GH secretion test. After the diagnosis of GHD, head MRI was performed before starting GH replacement therapy.

Hormone assays

Serum IGF-1 was measured by electrochemiluminescence immunoassay (Elecsys IGF-1; Roche Diagnostics, Tokyo, Japan), which was calibrated against the WHO International Standard 02/254. The values of serum IGF-1 were transformed into SDs, according to the established reference ranges of the assay for sex and calendar age [12]. GH was measured by immunoenzymometric assay (E Test TOSOH II HGH; Tosoh Co., Ltd., Tokyo, Japan), which was standardized against the WHO International Standard 98/574.

Statistical analysis

We calculated point estimates for IGF-1 (SD) sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), diagnostic efficiency (DE), positive likelihood ratio (PLR), and negative likelihood ratio (NLR) for predicting the presence of GHD. Data were shown as median (interquartile range) for chronological and bone age, and as mean \pm SD for the other numerical variables. Based on the f-test, Student's t-test was performed in the case of homoscedasticity and the Mann–Whitney U test was performed in the case of unequal variances to compare the IGF-1 level and other variables between the two groups. Spearman's rank correlation coefficient test was performed to investigate the relationship of IGF-1 (SD) with age, bone age, height (SD), target height (SD), height velocity before examination (SD), weight (SD), body mass index (BMI) (SD), and maximum peak GH (ng/mL). Correlation was defined as very weak if <0.2 , weak if ≥ 0.2 and <0.4 , moderate if ≥ 0.4 and <0.6 , strong if ≥ 0.6 and <0.8 , and very strong if ≥ 0.8 . Receiver operating characteristic (ROC) analysis with the Youden index was used to compare the discriminatory performances of IGF-1 in the diagnosis of GHD. Based on the area under the ROC curve (AUC), performance was considered as acceptable if >0.7 and ≤ 0.8 and excellent if >0.8 .

All statistical analyses were performed using EZR (Saitama Medical Center, Jichi Medical University, Saitama, Japan) [13], which is a graphical user interface for R (The R Foundation for Statistical Computing, Vienna, Austria). More precisely, it is a modified version of R commander designed to add statistical functions frequently used in biostatistics.

Results

The patients included in this study had a median age of 4.98 years (interquartile range, 3.21–9.38 years). We identified 60 children with GHD and 238 children without GHD (non-GHD) (Figure 1), with male preponderance of 53.4%. Four patients were diagnosed with organic GHD [lymphocytic hypophysitis ($n = 2$), craniopharyngioma ($n = 1$), and cerebral myeloma ($n = 1$)]. The background of these patients are described in Table 1. Height velocity after examination, body weight (SD), and BMI (SD) were significantly higher in the GHD group than in the non-GHD group. Maximum peak GH was significantly lower in the GHD group than in the non-GHD group.

ROC analysis showed the best diagnostic accuracy at an IGF-1 cutoff of -1.493 SD (sensitivity, 0.417; specificity, 0.689; PPV, 0.25; NPV, 0.823; DE, 0.631; PLR, 1.322; NLR, 0.852; AUC = 0.517) (Figure 2). Thus,

using the IGF-1 (SD) cutoff of -1.493 , a correct diagnosis was possible in 26 patients with GHD and 161 subjects without GHD.

IGF-1 (SD) across the GHD and non-GHD groups showed normality in the Kolmogorov–Smirnov test but not in the Shapiro–Wilk test, and outliers were in the IGF-1 (SD) phase. Because of the possible significant impact on the number of relationships, we decided to analyze the IGF-1 (SD) using Spearman's rank correlation coefficient. The correlation of IGF-1 (SD) was weak with age ($r = -0.264$, $p < 0.001$), bone age ($r = -0.26$, $p < 0.001$), height velocity before examination (SD) ($r = 0.22$, $p < 0.001$), weight (SD) ($r = 0.219$, $p < 0.001$), and BMI (SD) ($r = 0.241$, $p < 0.001$) and very weak with height (SD) ($r = 0.0815$, $p = 0.16$), target height (SD) ($r = -0.05$, $p = 0.393$), and maximum GH peak ($r = 0.129$, $p = 0.0257$).

To clarify the relationship between pre-treatment IGF-1 and response to GH, we compared the height velocity (SD) between groups with IGF-1 above ($n = 34$) and below ($n = 26$) the cutoff value (-1.493 SD). Pre- and post-treatment height velocity (SD) were similar between the groups (pre-treatment, -1.28 ± 2.25 vs. -1.52 ± 2.51 , $p = 0.696$; post-treatment, 2.48 ± 3.15 vs. 2.74 ± 3.09 , $p = 0.755$).

To assess the efficacy of the third stimulation test, patients diagnosed with GHD on the second ($n = 39$) and third tests ($n = 21$) were compared. Pre- and post-treatment growth velocities (SD) were similar between these groups (pre-treatment, -1.66 ± 2.36 vs. -0.89 ± 2.31 , $p = 0.232$; post-treatment, 2.44 ± 3.00 vs. 2.87 ± 3.33 , $p = 0.618$).

Discussion

We found that IGF-1 had poor accuracy as demonstrated by low AUC, and poor sensitivity, specificity, and DE for the best cutoff of -1.493 SD. The correlation analysis revealed that none of the items increased the diagnostic power of IGF-1 for GHD screening.

IGF-1 has been reported to be useful in the screening of GHD in some studies [3-7] but not in others [1, 8]. The reason for these contradictory results is that the patient groups and GHD cutoff values differed between these studies. First, the inclusion criteria may create differences in patient backgrounds. In previous studies on the usefulness of IGF-1 for GHD screening, patients were selected according to bone age [2], target height [3, 6], or catch-up growth [3] in addition to short stature and/or height velocity. These variations in inclusion criteria might superficially improve the sensitivity and specificity of IGF-1. Second, different GH cutoff levels for GHD were selected: ≤ 5 ng/mL [5, 6], ≤ 6 ng/mL [9], ≤ 7 ng/mL [1], ≤ 8 ng/mL [3, 8], and ≤ 10 ng/mL [4]. In the case of ≤ 8 or ≤ 10 ng/mL [3, 8], the prevalence of GHD in patients with short stature was $>30\%$, which was higher than that in our study (20.1%). Since disease prevalence affects sensitivity, specificity, PPV, and NPV, IGF-1 is not a useful screening test in a patient population with low prevalence of GHD. In the cohort of this study, the prevalence of GHD was decreased to 12.8% when the GH cutoff level of 5 ng/mL was selected. Therefore, when evaluating the efficacy of IGF-1, comparisons should be made at the same GH cutoff levels.

Bone age, target height, and height velocity should be taken into consideration before selecting patients for the GH secretion test [2]. In our study, bone age, target height, and height velocity before the examination were similar between the GHD and non-GHD groups. Even after combining these conditions with IGF-1, the diagnostic power of IGF-1 for GHD screening did not increase. Therefore, it would be difficult to distinguish patients with GHD from those without GHD using those parameters.

To clarify the relationship between pre-treatment IGF-1 and response to GH, height velocity (SD) between groups with IGF-1 above ($n = 34$) and below ($n = 26$) the cutoff value (-1.493 SD) was compared. Pre- and post-treatment height velocity (SD) were similar between the groups. IGF-1 was reported to be weakly correlated with the clinical endpoints of GH treatment [14]. Therefore, it would be difficult to predict the degree of improvement prior to GH treatment using pre-treatment IGF-1.

We performed the third stimulation test when the result of GH secretion in the first and second tests were different. Although a sufficient GH response in one stimulation test rules out GHD in most cases [2], the utilization and interpretation of the drugs used in the stimulation test depends on the facility [1, 15]. In this study, pre- and post-treatment growth velocities were similar between the patients diagnosed with GHD on the second and third tests. This indicates that, regardless of the number of stimulation tests, patients with GHD have similar response to growth hormone. Therefore, the third stimulation test may be effective in diagnosing patients with GHD.

This study had several limitations. First, immunoassay for IGF-1 analysis is not the most sensitive assay. The variations in immunoassays used in different studies may result in variations in the reported efficacy of IGF-1. More accurate assays, such as LC-MS, may reveal the actual usefulness of IGF-1 for GHD screening. Second, it is unclear whether the third stimulation should be performed or not. In this study, the third stimulation test was performed when the second stimulation test provided a result of 6–8 ng/mL. However, when the first stimulation test provided a result of 6–8 ng/mL, the next test was not performed. Depending on the order of the tests, the diagnosis of GHD may vary. Therefore, it needs to be clarified whether the third stimulation test should be performed.

In conclusion, IGF-1 level had poor diagnostic accuracy as a screening test for GHD. Correlation analysis revealed that none of the items increased the diagnostic power of IGF-1. Therefore, IGF-1 should not be used alone for the screening of GHD. A predictive biomarker for GHD should be developed in the future.

Declarations

Funding

The authors declare that they received no funding for this research.

Conflicts of interest

The authors declare that there is no conflict of interest that could be perceived as a prejudice to the impartiality of this study.

Availability of data and material

The data that support the findings of this study are available upon request from the corresponding author. The data are not publicly available because of privacy and ethical restrictions.

Authors' contributions

HI contributed to study design and data interpretation and was a major contributor in manuscript writing. SK, JS, RM, TH, YK, TM, and HK contributed to data acquisition and analysis. WO, JT, and AO critically revised the manuscript for important intellectual content. HI contributed to the final approval of the version to be published. All authors read and approved the final manuscript.

Ethics approval

The study was approved by the ethics committee of Aichi Medical University (originally 2015-H359 but revised to 2020-H041 as the study period was updated).

Consent to participate and for publication

The parents of the study subjects provided consent to participate and for publication after full explanation of the purpose and nature of all the procedures used in this study.

References

1. Ibba, A., Corrias, F., Guzzetti, C., Casula, L., Salerno, M., di Iorgi, N., Tornese, G., Patti, G., Radetti, G., Maghnie, M., Cappa, M., Loche, S.: IGF1 for the diagnosis of growth hormone deficiency in children and adolescents: a reappraisal. *Endocr. Connect.* 9(11), 1095–1102 (2020). doi:10.1530/ec-20-0347
2. Collett-Solberg, P.F., Jorge, A.A.L., Boguszewski, M.C.S., Miller, B.S., Choong, C.S.Y., Cohen, P., Hoffman, A.R., Luo, X., Radovick, S., Saenger, P.: Growth hormone therapy in children; research and practice - a review. *Growth Horm. IGF Res.* 44, 20–32 (2019). doi:10.1016/j.ghir.2018.12.004
3. Binder, G., Huller, E., Blumenstock, G., Schweizer, R.: Auxology-based cut-off values for biochemical testing of GH secretion in childhood. *Growth Horm. IGF Res.* 21(4), 212–218 (2011). doi:10.1016/j.ghir.2011.05.007
4. Guzzetti, C., Ibba, A., Pilia, S., Beltrami, N., Di Iorgi, N., Rollo, A., Fratangeli, N., Radetti, G., Zucchini, S., Maghnie, M., Cappa, M., Loche, S.: Cut-off limits of the peak GH response to stimulation tests for the diagnosis of GH deficiency in children and adolescents: study in patients with organic GHD. *Eur. J. Endocrinol.* 175(1), 41–47 (2016). doi:10.1530/eje-16-0105
5. Inoue-Lima, T.H., Vasques, G.A., Scalco, R.C., Nakaguma, M., Mendonca, B.B., Arnhold, I.J.P., Jorge, A.A.L.: IGF-1 assessed by pubertal status has the best positive predictive power for GH deficiency diagnosis in peripubertal children. *J. Pediatr. Endocrinol. Metab.* 32(2), 173–179 (2019). doi:10.1515/jpem-2018-0435

6. Inoue-Lima, T.H., Vasques, G.A., Nakaguma, M., Brito, L.P., Mendonça, B.B., Arnhold, I.J.P., Jorge, A.A.L.: A bayesian approach to diagnose growth hormone deficiency in children: insulin-like growth factor type 1 is valuable for screening and IGF-binding protein type 3 for confirmation. *Horm. Res. Paediatr.* 93(3), 197–205 (2020). doi:10.1159/000509840
7. Shen, Y., Zhang, J., Zhao, Y., Yan, Y., Liu, Y., Cai, J.: Diagnostic value of serum IGF-1 and IGFBP-3 in growth hormone deficiency: a systematic review with meta-analysis. *Eur. J. Pediatr.* 174(4), 419–427 (2015). doi:10.1007/s00431-014-2406-3
8. Galluzzi, F., Quaranta, M.R., Salti, R., Saieva, C., Nanni, L., Seminara, S.: Are IGF-I and IGF-BP3 useful for diagnosing growth hormone deficiency in children of short stature? *J. Pediatr. Endocrinol. Metab.* 23(12), 1273–1279 (2010). doi:10.1515/jpem.2010.201
9. Katsumata, N., Shimatsu, A., Tachibana, K., Hizuka, N., Horikawa, R., Yokoya, S., Tatsumi, K.I., Mochizuki, T., Anzo, M., Tanaka, T.: Continuing efforts to standardize measured serum growth hormone values in Japan. *Endocr. J.* 63(10), 933–936 (2016). doi:10.1507/endocrj.EJ16-0198
10. Isojima, T., Kato, N., Ito, Y., Kanzaki, S., Murata, M.: Growth standard charts for Japanese children with mean and standard deviation (SD) values based on the year 2000 national survey. *Clin. Pediatr. Endocrinol.* 25(2), 71–76 (2016). doi:10.1297/cpe.25.71
11. Murata, M.: Japanesespecific bone age standard on the TW2. *Clin. Pediatr. Endocrinol.* 2(Supple3), 35–41 (1993). doi:10.1297/cpe.2.Supple3_35
12. Isojima, T., Shimatsu, A., Yokoya, S., Chihara, K., Tanaka, T., Hizuka, N., Teramoto, A., Tatsumi, K.I., Tachibana, K., Katsumata, N., Horikawa, R.: Standardized centile curves and reference intervals of serum insulin-like growth factor-I (IGF-I) levels in a normal Japanese population using the LMS method. *Endocr. J.* 59(9), 771–780 (2012). doi:10.1507/endocrj.ej12-0110
13. Kanda, Y.: Investigation of the freely available easy-to-use software 'EZR' for medical statistics. *Bone Marrow Transplant.* 48(3), 452–458 (2013). doi:10.1038/bmt.2012.244
14. Johannsson, G., Bidlingmaier, M., Biller, B.M.K., Boguszewski, M., Casanueva, F.F., Chanson, P., Clayton, P.E., Choong, C.S., Clemmons, D., Dattani, M., Frystyk, J., Ho, K., Hoffman, A.R., Horikawa, R., Juul, A., Kopchick, J.J., Luo, X., Neggers, S., Netchine, I., Olsson, D.S., Radovick, S., Rosenfeld, R., Ross, R.J., Schilbach, K., Solberg, P., Strasburger, C., Trainer, P., Yuen, K.C.J., Wickstrom, K., Jorgensen, J.O.L.: Growth Hormone Research Society perspective on biomarkers of GH action in children and adults. *Endocr. Connect.* 7(3), R126–R134 (2018). doi:10.1530/ec-18-0047
15. Chesover, A.D., Dattani, M.T.: Evaluation of growth hormone stimulation testing in children. *Clin. Endocrinol. (Oxf.)* 84(5), 708–714 (2016). doi:10.1111/cen.13035

Tables

Table 1. Background characteristics of the GHD and non-GHD groups

	GHD group (n = 60)	Non-GHD group (n = 238)	p value
Male:female	34:26	125:113	0.664
Age (year)	5.10 (3.37–10.03)	4.94 (3.19–9.17)	0.565
Bone age (year)	3.21 (2.42–7.90)	3.21 (2.42–7.5)	0.778
Prepubertal:pubertal	58:2	224:14	0.748
Height (SD)	-2.32 ± 0.76	-2.37 ± 0.82	0.645
Target height (SD)	-0.49 ± 0.69	0.47 ± 0.71	0.869
Height velocity before examination (SD)	-1.39 ± 2.35	-0.97 ± 1.93	0.157
Height velocity after examination (SD)	2.60 ± 3.10	-0.12 ± 2.96	<0.001
Weight (SD)	-1.61 ± 0.96	-2.00 ± 1.08	0.0144
BMI (SD)	-0.11 ± 1.11	-0.50 ± 1.12	0.0207
Max. Peak GH (ng/mL)	5.94 ± 2.63	11.17 ± 3.97	<0.001
IGF-1 (SD)	-1.43 ± 1.14	-1.12 ± 0.88	0.23

The values were expressed as median (IQR) or mean ± SD

BMI, body mass index; GH, growth hormone; GHD, growth hormone deficiency; IGF-1, insulin-like growth factor; IQR, interquartile range;

Max., maximum

Figures

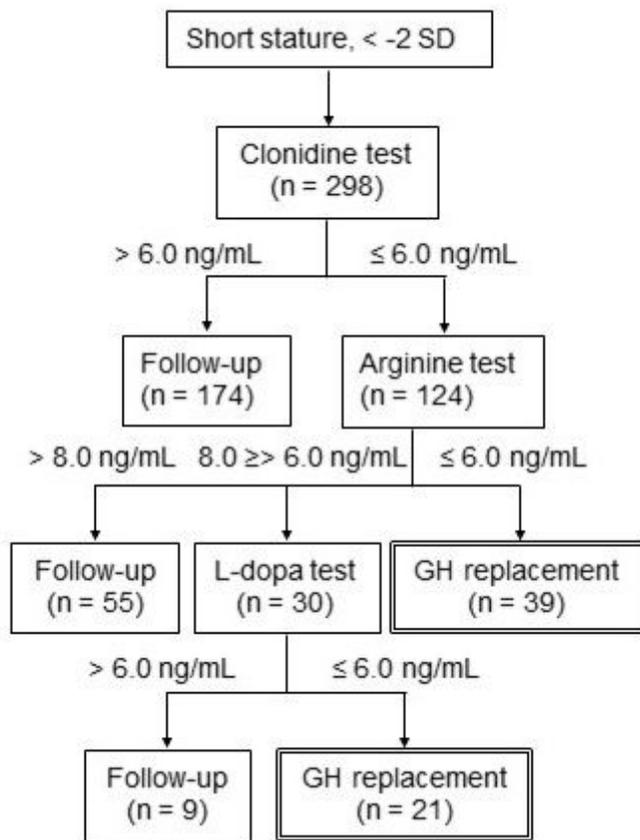


Figure 1

Algorithm of the stimulation tests using clonidine, arginine, and L-dopa. If peak GH (growth hormone) in both clonidine and arginine tests was $\leq 6.0\text{ ng/mL}$, GH replacement therapy was initiated. When the peak GH in arginine test was $6.0\text{--}8.0\text{ ng/mL}$, the third stimulation test using L-dopa was performed.

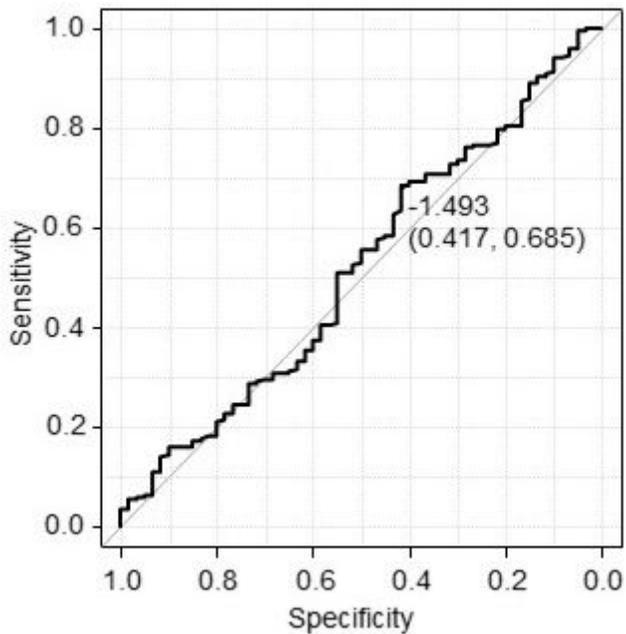


Figure 2

Receiver operating characteristic (ROC) curve of IGF-1 (SD) for the diagnosis of growth hormone deficiency (GHD) ROC analysis showed the best diagnostic accuracy at an IGF-1 cutoff of -1.493 SD (sensitivity, 0.685; specificity, 0.417; area under the ROC curve = 0.517).