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

Chiara Guzzetti¹, Anastasia Ibba¹, Sabrina Pilia¹, Nadia Beltrami²,
Natascia Di Iorgi³, Alessandra Rollo⁴, Nadia Fratangeli³, Giorgio Radetti²,
Stefano Zucchini⁴, Mohamad Maghnie³, Marco Cappa⁵ and Sandro Loche¹

¹SSD Endocrinologia Pediatrica, Ospedale Pediatrico Microcitemico "A. Cao", Cagliari, Italy, ²Ospedale Generale Regionale, Bolzano, Italy, ³Clinica Pediatrica, Istituto Giannina Gaslini, Università di Genova, Genova, Italy, ⁴Università di Bologna, Ospedale S.Orsola-Malpighi, Bologna, Italy, and ⁵UOC di Endocrinologia e Diabetologia, Ospedale Pediatrico Bambino Gesù IRCCS, Roma, Italy

Correspondence
should be addressed
to S Loche
Email
sandro.loche@aob.it

Abstract

Objective: The diagnosis of GH deficiency (GHD) in children and adolescents is established when GH concentrations fail to reach an arbitrary cut-off level after at least two provocative tests. The objective of the study was to define the optimal GH cut-offs to provocative tests in children and adolescents.

Design: Retrospective study in 372 subjects who underwent evaluation of GH secretion. GH and IGF-I were measured by chemiluminescence assay in all samples. Receiver operating characteristic (ROC) analysis was used to evaluate the optimal GH cut-offs and the diagnostic accuracy of provocative tests.

Methods: Seventy four patients with organic GHD (GH peak <10 µg/L after two provocative tests) and 298 control subjects (GH response >10 µg/L to at least one test) were included in the study. The provocative tests used were arginine, insulin tolerance test (ITT) and clonidine. Diagnostic criteria based on cut-offs identified by ROC analysis (best pair of values for sensitivity and specificity) were evaluated for each test individually and for each test combined with IGF-I SDS. **Results:** The optimal GH cut-off for arginine resulted 6.5 µg/L, 5.1 µg/L for ITT and 6.8 µg/L for clonidine. IGF-I SDS has low accuracy in diagnosing GHD (AUC=0.85). The combination of the results of provocative tests with IGF-I concentrations increased the specificity.

Conclusions: The results of the ROC analysis showed that the cut-off limits which discriminate between normal and GHD are lower than those commonly employed. IGF-I is characterized by low diagnostic accuracy.

European Journal of
Endocrinology
(2016) **175**, 41–47

Introduction

The diagnosis of growth hormone (GH) deficiency (GHD) during childhood is classically based on clinical assessment, combined with laboratory tests of the GH-insulin-like growth factor (IGF) axis and radiological evaluation (1, 2, 3). Stimulation tests are commonly used to assess pituitary GH secretory capacity. GHD is traditionally confirmed when peak GH concentration does not reach

the established cut-off level after two different stimulation tests (1, 2, 3). A large number of stimulation tests for GH secretion have been proposed in the last 50 years (4). The cut-off level was initially set at 5–7 µg/L, and subsequently was arbitrarily increased to 10 µg/L (5, 6, 7, 8, 9, 10). The choice of this single fixed cut-off level consider neither the type of stimulus nor the variation in response

due to gender, age, puberty, BMI and other factors. The interpretation of the results of the provocative tests has proven to be problematic for many other reasons including nonphysiological test procedures, type of stimulation test and type of assay. In fact, several GH assays are available, making it difficult to compare results from different laboratories (11, 12). Based on the foregoing, it has been made clear that provocative tests have low reproducibility, specificity and sensitivity (13).

IGF-I, which mediates most of the growth-promoting action of GH, is also commonly used in the diagnostic workup of GHD. IGF-I shows little diurnal variation and has good specificity. However, it is markedly influenced by nutritional status, and there is still the need for an acceptable IGF-I standard reference preparation. It has been shown that IGF-I concentration does not correlate with the GH response to stimulation tests (14, 15).

Based on these premises, we designed a retrospective study in order to define the most appropriate cut-off limits for GH peak after different stimulation tests. We used receiver operating characteristic (ROC) curve analysis to generate the best pair of values for sensitivity and specificity (16) in a large group of children and adolescents with an unequivocal diagnosis of GHD and in control children who underwent GH stimulation tests with insulin hypoglycemia (insulin tolerance test (ITT)), arginine and clonidine. Serum GH was measured in all samples using the same immunoassay.

Subjects and methods

Subjects

Data were collected from 372 children and adolescents (237 boys and 135 girls, 205 prepubertal and 167 pubertal, median age 10.5 (6.5; 13.5), H-SDS -2.4 (-2.8 ; -1.9). These children underwent diagnostic procedures for suspected GHD according to the consensus guidelines (1), including GH secretion studies after exclusion of other causes for their shortness. All children included in the study were seen in five Pediatric Endocrine Units in Italy (Bologna, Bolzano, Cagliari, Genova and Roma) between 2005 and 2013.

The subjects were subdivided into patient group (cases) and control group (controls). Their main clinical characteristics are summarized in Table 1.

The cases consisted of 74 patients with an established diagnosis of GHD of organic origin (55 central nervous system tumors including craniopharyngioma ($n=32$), germinoma ($n=17$), astrocytoma ($n=2$), pituitary adenoma

Table 1 Clinical characteristics of the subjects studied Data are presented as median (IQR).

	Cases ($n=74$)	Controls ($n=298$)
Sex (M/F)	47/27	190/108
Prepubertal/pubertal	37/36	168/130
Age (years)	12.6 (8.6; 14)	10 (5.8; 13.2)
Peak GH to arginine ($\mu\text{g/L}$)	1.3 (0.3; 2.6)	12.3 (9.3; 16.6)
Peak GH to ITT ($\mu\text{g/L}$)	1.2 (0.4; 3.5)	12.6 (9.3; 16.5)
Peak GH to clonidine ($\mu\text{g/L}$)	1.8 (1.1; 4.2)	13.1 (10.1; 18.9)
H-SDS	-2 (-2.7 ; -0.8)	-2.5 (-2.8 ; -2)
BMI-SDS	0.7 (-0.2 ; 1.5)	-0.6 (-1.4 ; 0.1)
HV pretest	-2 (-3.5 ; -1)	-1.3 (-2.2 ; -0.5)
HV posttest	2.2 (0.3 ; 4.2)	-0.2 (-1 ; 0.6)
IGF-I SDS	-2.9 (-4 ; -1.9)	-0.9 (-1.8 ; -0.03)

ITT, insulin tolerance test; H-SDS, height-SDS; BMI-SDS, body mass index-SDS; HV, height velocity; IGF-I SDS, insulin-like growth factor-I SDS.

($n=1$), glioma ($n=3$); 18 malformations including empty sella ($n=3$), ectopic neurohypophysis and/or pituitary hypoplasia ($n=14$), ethmoidal-sellar cephalocele ($n=1$); 1 PROP1 mutation) (Table 1). At the time of diagnosis, all patients had presented reduced height-velocity (HV)-SDS (median 2 (-3.5 ; -1)). In 35 patients (47.3%), the diagnosis of GHD was established before recognition of its cause (tumor/malformation). Forty three patients had additional pituitary hormone deficiencies (15 TSH, 1 ACTH/cortisol, 27 both TSH and ACTH). Patients with additional hormone deficiency received appropriate replacement therapy with hydrocortisone and/or L-thyroxine.

The controls consisted of 298 children who underwent GH stimulation tests for their short stature (Table 1). Their HV at the time of evaluation was normal and were ultimately found to have a GH response $>10\mu\text{g/L}$ to at least one test.

Methods

The three stimulation tests used in this study were arginine, ITT and clonidine. The tests were performed in a random order. All tests were performed between 08:00 and 09:00 after overnight fasting. Arginine test was performed in 61 cases and 151 controls. Arginine was administered intravenously (0.5 g/kg , max 30 g) during 30min and blood samples for GH determination were collected at times -30 , 0 , 15 , 30 , 45 , 60 , 90 and 120 min . ITT was performed in 46 cases and 60 controls. Insulin was administered intravenously (0.05 – 0.1 U/kg) and blood samples for GH and glucose determinations were collected at times 0 , 30 , 60 , 90 and 120 min . A nadir glucose value during ITT below 40 mg/dL (2.2 mmol/L) was recorded in all subjects at time 30 min . Clonidine test

was performed in 26 cases and 192 controls. Clonidine was administered orally (0.15 mg/m²) and blood samples for GH determination were collected at times 0, 30, 60, 90 and 120 min. Stimulation tests were performed on separate days (at least 2 days apart). IGF-I was also determined in 334 children at baseline (58 cases and 276 controls). Steroid priming was not used in any of the subjects. Brain MRI was performed in all children with a GH response <10 µg/L after two stimulation tests. Ten subjects had a peak GH <10 after two stimulation tests and a normal MRI. In these patients, a third test was performed few months apart which resulted normal.

Assays

Serum GH and IGF-I were measured by chemiluminescent immunometric assay (Immulite 2000; Diagnostic Products Corp, Los Angeles, CA, USA). GH assay was calibrated against the recommended IS 98/574. IGF-I assay was calibrated against the WHO IRR 87/518. The sensitivity of the method was 0.01 µg/L for GH and 2.6 nmol/L for IGF-I. The GH intra- and interassay coefficients of variation were, respectively, 4.2–6.6 and 2.9–4.6% at GH levels of 2.6–17 µg/L. The intra- and interassay coefficients of variation for IGF-I were 3.4 and 7.1% respectively. Consistency of assay performance was assessed by regular use of internal controls.

Statistical analysis

Distribution of the data was evaluated using Kolmogorov–Smirnov test. GH peak after ITT in patients and age, H-SDS, and GH peak after clonidine in controls were the only not normally distributed variables. Comparisons between two groups were performed using the Student *t*-test for normally distributed variables and the Mann–Whitney *U*-test for not normally distributed variables.

The possibly confounding effect of age, gender, puberty and BMI-SDS on peak GH and IGF-I was assessed by multiple regression analysis performed in cases and controls separately.

ROC analysis was applied to evaluate the statistically best cut-off value in order to provide the best pair of values for sensitivity and specificity. To compare the diagnostic accuracy of the tests the area under the curve (AUC) with 95% confidence interval (CI) was calculated. Likelihood ratios for a positive test result (LR+) and efficiency (the number of correct results divided by the total number of tests) were provided as a measure of diagnostic accuracy.

Diagnostic criteria based on cut-offs identified by ROC analysis were evaluated for each test individually and for each test combined with IGF-I SDS (considering as positive a subject who was positive on each test). The association between test results and the presence of disease was evaluated by Pearson χ^2 test.

H-SDS and BMI-SDS were derived from the Italian reference data (17). HV-SDS was derived from Tanner's charts (18). IGF-I-SDS was calculated using the normative data for the method (19). All values are reported as median and interquartile range (IQR) (continuous variables) or as counts and percentages (categorical variables). Sensitivity, specificity, positive and negative predictive values (PPV, NPV), and diagnostic accuracy were expressed as percentage. $P < 0.05$ (two sided) was considered significant. All statistical calculations were performed using STATISTICA version 9.1 software (StatSoft Inc, Tulsa, OK, USA).

Results

Subjects' characteristics

All clinical and biochemical characteristics of the subjects are summarized in Table 1. The median age of the patients was higher than in controls. However, age had a significant effect of the GH peak neither in the patients nor in the control group. Likewise, multiple regression analysis excluded a significant influence of gender, and puberty on GH responses, except for pubertal status and age that significantly influenced the GH peak after clonidine in controls ($B = -0.45$, $P = 0.041$; $B = 3.14$, $P = 0.049$ respectively). However, the median peak GH after clonidine in prepubertal and pubertal controls was similar (13.4 (10.3–18.8) µg/L; 12.6 (10–19.5) µg/L, respectively; $P = 0.88$). Therefore, they have been considered together in the statistical analysis.

We found a significant correlation between BMI-SDS and the GH peak after arginine and clonidine, and IGF-I SDS in controls ($B = -1$, $P = 0.034$; $B = -1.65$, $P = 0.0003$; $B = 0.14$, $P = 0.019$ respectively). However, the number of patients with pathological BMI was small and did not consent to stratify the population according to their BMI.

Mean peak GH to all three stimulation tests was significantly lower in the cases than in the controls (data not shown). Likewise, mean baseline IGF-I SDS concentrations were significantly lower in the cases than in the controls (data not shown).

Fifty, 39 and 10 patients had a GH peak <10 after the arginine, clonidine and ITT tests performed as the first test, respectively, and a normal response to a second test.

Figure 1

ROC curves for peak GH serum responses to stimulation tests.

Age, H-SDS and BMI-SDS were significantly higher in the GHD patients than in controls. Many of the GHD patients had craniopharyngioma, and it is well known that growth retardation in these patients is often associated with weight gain (20). After 1–2 years, their HV-SDS significantly increased (2.2 (0.3; 4.2, $P < 0.0001$).

Cut-off identification

Results of ROC analysis are shown in Figs 1, 2 and Table 2. The AUC (measure of discriminative ability of the test) for arginine was 0.96 (95% CI: 0.93–0.99), ITT was 0.97 (95% CI: 0.95–0.99), clonidine was 0.95 (95% CI: 0.9–1) and IGF-I SDS was 0.85 (95% CI: 0.79–0.9). For arginine, the best pair of values for sensitivity and specificity was found at a peak GH of 6.5 µg/L (LR+ 11.76, efficiency 92.5%). For ITT, the optimal peak GH cut-off point was 5.1 µg/L (LR+ 10.96, efficiency 91.5%), and for clonidine, it was 6.8 µg/L (LR+ 33.97, efficiency 96.3%). ROC analysis of IGF-I SDS showed the best diagnostic accuracy at a cut-off of –1.8 SDS (LR+ 3.17, efficiency 75.5%).

Cut-off application and test combination

Then we applied the cut-off limits identified (Table 3). Arginine alone exhibited a sensitivity of 94%, specificity of 92.7%, LR+ of 12.84 and efficiency of 93%. When arginine was combined with IGF-I SDS, the sensitivity was 77.1%, specificity 95.6%, LR+ 17.6 and efficiency 90.8%.

ITT alone exhibited a sensitivity of 94.4%, specificity of 89.6%, LR+ of 9.07 and efficiency of 91.7%. ITT combined with IGF-I SDS exhibited a sensitivity of 80.6%, specificity of 95.8%, LR+ of 19.33 and efficiency of 89.3%.

Clonidine alone exhibited a sensitivity of 88.5%, specificity of 97.3%, LR+ of 33.26 and efficiency of 96.3%. Clonidine combined with IGF-I SDS exhibited a

sensitivity of 65.4%, specificity of 98.4%, LR+ of 40.97 and efficiency of 94.4%.

Discussion

In the 1960s, some studies showed that GH serum concentrations rose after various stimulation tests in children with idiopathic short stature but not in children with GHD (5, 6, 7, 8, 9, 10). These studies were performed in small numbers of short children and led to the definition of a cut-off value for GH peak between 5 and 7 µg/L. When biosynthetic GH became available in 1985, the GH limit was arbitrarily increased to 10 µg/L (1), with the same cut-off applied to children of all ages and both sexes (13). Although biochemical tests for GH secretion clearly distinguish children with severe GHD of pituitary origin, recognition of more subtle forms of GH insufficiency still presents a diagnostic dilemma. Furthermore, more than 70% (up to 100% in some studies) of isolated idiopathic

Figure 2

ROC curve for IGF-I SDS.

Table 2 Optimal cut-off points of peak GH responses to arginine, ITT, and clonidine, and optimal cut-off point of IGF-I SDS evaluated with ROC analysis. Cut-off points for sensitivity and specificity at 95% are also reported.

	Arginine	ITT	Clonidine	IGF-I SDS
Cut-off point	6.5 µg/L	5.1 µg/L	6.8 µg/L	−1.8 SDS
Sensitivity (%)	93.4	91.3	88.5	79.3
Specificity (%)	92.1	91.7	97.4	75
Efficiency (%)	92.5	91.5	96.3	75.5
LR+	11.76	10.96	33.97	3.17
Cut-off point for sensitivity at 95%	7.7 µg/L	6.8 µg/L	12.1 µg/L	−0.5 SDS
Cut-off point for specificity at 95%	3.6 µg/L	4.2 µg/L	6.8 µg/L	−2.9 SDS

ITT, insulin tolerance test; IGF-I SDS, insulin-like growth factor-I SDS; ROC, receiver operating characteristic; LR+, positive likelihood ratio.

GHD patients diagnosed in childhood have normal GH responses to stimulation when retested at the end of statural growth or even a few months apart (21, 22, 23). Spontaneous resolution of transient GHD in adolescence has been proposed. However, false-positive results to GH provocative tests may account for the discrepancy in GH responses (22). Many factors may be involved in this variability in the GH response to the stimulation test including age, sex, nutrition and BMI (24).

In our cohort, ten children failed the first two test, but since they had a normal MRI and no evidence of other disease, they were retested after 3–6 months and all had a peak GH >10 µg/L. We recognize that some patients with IGHD may well have a permanent deficit which will not normalize at retesting either shortly or at the end of statural growth. However, no patient with such characteristics was present in our cohort.

IGF-I measurement may help in the diagnosis of GHD but presents some limitations. The regulation of IGF-I secretion and action is complex, and many factors, such as age, endogenous GH secretion, BMI, physical

fitness, glucocorticoids, sex steroids and increased IGFBP-3 binding capacity, are potential determinants of IGF-I serum concentration (2, 12, 13). We and others (25, 26, 27) have shown that in patients with severe unequivocal GHD, IGF-I concentrations are invariably reduced, while as many as 30% of children with idiopathic GHD show normal IGF-I concentrations (28). A considerable overlap has been shown between children with GHD and children with idiopathic short stature, making it impossible to discriminate between the two groups (2, 14, 29).

Among the numerous GH provocative tests, arginine, ITT and clonidine are routinely used in most centers to evaluate GH secretion in children and are among the preferred tests in the last Consensus Guidelines (1). In our study, we applied ROC analysis for these three different stimulation tests.

Arginine stimulates GH secretion via inhibition of endogenous somatostatin release, and enhances the GH response to exogenous GHRH (29, 30). In our study, ROC curve analysis showed an AUC of 0.96, which classifies the arginine test as highly predictive in the diagnosis of GHD (16). Moreover, ROC curve analysis indicated that a peak GH of 6.5 µg/L provides the best pair of sensitivity and specificity. This GH cut-off value allows to correctly identify 92.5% of patients. Binder *et al.* (31) identified the best diagnostic accuracy at a GH peak of 6.6 µg/L (sensitivity 84.3%, specificity 82.4%, AUC 0.83). Ghigo *et al.* (29) found that in 12.6 and 32.9% of normally growing children, the peak GH responses were lower than 7 and 10 µg/L respectively.

Insulin-induced hypoglycemia has been used for decades to test GH secretion and is considered the gold standard (32, 33). Also, hypoglycemia lowers hypothalamic somatostatin tone, thereby increasing GH release. Our results showed that ITT has a high predictive value in the diagnosis of GHD (16), with AUC of 0.97 in ROC analysis. The cut-off value with the best pair of sensitivity and specificity was 5.1 µg/L. This cut-off allows to correctly

Table 3 Diagnostic accuracy of GH stimulation tests alone and associated with IGF-I SDS (positivity to each test to be affected). Only patients with available IGF-I measurement included.

	n	Sensitivity (%)	Specificity (%)	PPV	NPV	LR+	Efficiency (%)	χ^2	P
Arginine	185	93.8	92.7	81.8	97.7	12.84	93	127.2	<0.0001
Arginine + IGF-I SDS		77.1	95.6	86.1	92.3	17.6	90.8	105.3	<0.0001
ITT	84	94.4	89.6	87.2	95.6	9.07	91.7	58.4	<0.0001
ITT + IGF-I SDS		80.6	95.8	93.6	86.8	19.33	89.3	51.5	<0.0001
Clonidine	214	88.5	97.3	82.1	98.4	33.26	96.3	147.9	<0.0001
Clonidine + IGF-I SDS		65.4	98.4	85	95.4	40.97	94.4	109.7	<0.0001

ITT, insulin tolerance test; IGF-I SDS, insulin-like growth factor-I SDS; PPV, positive predictive value; NPV, negative predictive value, LR+, positive likelihood ratio.

Clinical Study	C Guzzetti and others	GH response to stimulation test	175:1	46
----------------	-----------------------	---------------------------------	-------	----

identify 91.5% of patients. In normally growing children (29), the mean GH response to ITT was $13.2 \pm 1.2 \mu\text{g/L}$ (range 2.7–46.4). In this population (29), using a cut-off of $3 \mu\text{g/L}$, the prediction of GHD is 100% and lowers to 85% when a cut-off of $10 \mu\text{g/L}$ is considered.

Clonidine is an alpha-receptor agonist that stimulates GH secretion by a central mechanism, probably by inducing endogenous GHRH release (29, 30). In our study, the ROC analysis showed that the clonidine test has a high predictive value with an AUC of 0.95 (16). The best diagnostic accuracy is at peak GH of $6.8 \mu\text{g/L}$. This cut-off value allows to correctly identify 96.3% of patients. Silva *et al.* (34) determined the normal GH peak values after clonidine stimulation in normal prepubertal children using an IFMA and suggested the use of a cut-off value of $3.3 \mu\text{g/L}$, which yields 100% sensitivity with 93% specificity to exclude the diagnosis of GHD. In normally growing children (29), the mean GH response to clonidine stimulation was $13.1 \pm 1.8 \mu\text{g/L}$ (range 3.8–86). In 10.1 and 23.2% subjects, the peak GH responses were lower than 7 and $10 \mu\text{g/L}$ respectively.

Recently, Wagner *et al.* (35), re-evaluated the cut-off value in diagnosis of GHD, using ROC analysis and Youden index, in 44 GHD children and 52 controls tested with arginine, ITT and glucagon. They identified $7.09 \mu\text{g/L}$ as the cut-off point where the rate of correct classification was at its maximum. The cut-off they obtained for the chemiluminescent immunometric assay (Immulite 2000) after correction was $7.77 \mu\text{g/L}$, which is higher than in our results.

Despite the above-discussed limitations, IGF-I may be helpful in the diagnosis of GHD. Binder *et al.* (31) defined new limits for IGF-I SDS for screening purposes, and in their population, 100% sensitivity was provided by an IGF-I cut-off value of -1.4 SDS. This value, however, shows very low specificity and an AUC of 0.8, indicating the poor accuracy of the test. In our study, IGF-I SDS alone shows moderate prediction power in the diagnosis of GHD with an AUC value of 0.85 (16). The most accurate cut-off level for IGF-I SDS was -1.8 SDS, but still with moderate low sensitivity and specificity (79.3 and 75% respectively). Despite the low efficiency of IGF-I SDS (75.5%) alone in diagnosing GHD, IGF-I SDS combined with GH peak after provocative test increases the specificity of the test results, reaching 95.6% for arginine, 95.8% for ITT and 98.4% for clonidine. Thus, combining the results of the stimulation tests with IGF-I measurement reduces the risk of false-negative results.

Despite the well-known poor diagnostic accuracy (13), GH stimulation testing is a routine procedure in most pediatric endocrinology centers. Although current guidelines

still indicate a cut-off of $10 \mu\text{g/L}$, such value is not supported by robust clinical data (1) and, most likely, produces a high number of false-positive results. These observations, coupled with the results of this study, suggest that this limit should be reconsidered. Our study has the advantage of being conducted on a well-selected, large population of patients with an unequivocal diagnosis of GHD of organic origin. Furthermore, GH and IGF-I were measured by the same chemiluminescence assay in all samples. The study has the limitation of being retrospective, despite these testing procedures and the assays used were homogeneous among the five centers involved in the study.

In conclusion, this study shows that the cut-off limits of peak GH after provocative tests, which better discriminate between normal children and children with GHD, are lower than those reported in the current guidelines (1), and differ according to the stimulation test. IGF-I alone is characterized by low diagnostic accuracy but is useful in combination with results of GH stimulation testing. Thus, based on these results, new normal values for the GH peak after stimulation test should be established.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This research did not receive any specific grant from any funding agency in the public, commercial or not-for-profit sector.

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. GH Research Society. *Journal of Clinical Endocrinology and Metabolism* 2000 **85** 3990–3993. (doi:10.1210/jcem.85.11.6984)
- 2 Shalet SM, Toogood A, Rahim A & Brennan BM. The diagnosis of growth hormone deficiency in children and adults. *Endocrine Reviews* 1998 **19** 203–223. (doi:10.1210/edrv.19.2.0329)
- 3 Frasier SD. A preview of growth hormone stimulation tests in children. *Pediatrics* 1974 **53** 929–937.
- 4 Saggese G, Ranke MB, Saenger P, Rosenfeld RG, Tanaka T, Chaussain JL & Savage MO. Diagnosis and treatment of growth hormone deficiency in children and adolescents: towards a consensus. Ten years after the Availability of Recombinant Human Growth Hormone Workshop held in Pisa, Italy, 27–28 March 1998. *Hormone Research* 1998 **50** 320–340. (doi:10.1159/000023298)
- 5 Kaplan SL, Abrams CAL, Bell JJ, Conte FA & Grumbach MM. Serum growth hormone response to insulin-induced hypoglycemia in disorders of growth. *Journal of Pediatrics* 1965 **67** 956–959. (doi:10.1016/S0022-3476(65)81883-2)
- 6 Frohman LA, Aceto T & MacGillivray MH. Studies of growth hormone secretion in children: normal, hypopituitary and constitutionally delayed. *Journal of Clinical Endocrinology and Metabolism* 1967 **27** 1409–1417. (doi:10.1210/jcem-27-10-1409)

- 7 Raiti S, Davis WT & Blizzard RM. A comparison of the effects of insulin hypoglycaemia and arginine infusion on release of human growth hormone. *Lancet* 1967 **2** 1182–1183. (doi:10.1016/S0140-6736(67)91896-X)
- 8 Root AW, Saenz-Rodriguez C, Bongiovanni AM & Eberlein WR. The effect of arginine infusion on plasma growth hormone and insulin in children. *Journal of Pediatrics* 1969 **74** 187–197. (doi:10.1016/S0022-3476(69)80066-1)
- 9 Root AW, Rosenfield RL, Bongiovanni AM & Eberlein WR. The plasma growth hormone response to insulin-induced hypoglycemia in children with retardation of growth. *Pediatrics* 1967 **39** 844–852.
- 10 Frasier SD. The serum growth-hormone response to hypoglycemia in dwarfism. *Journal of Pediatrics* 1967 **71** 625–638. (doi:10.1016/S0022-3476(67)80197-5)
- 11 Trainer PJ, Barth J, Sturgeon C & Wieringaon G. Consensus statement on the standardisation of GH assays. *European Journal of Endocrinology* 2006 **155** 1–2. (doi:10.1530/eje.1.02186)
- 12 Bidlingmaier M. Problems with GH assays and strategies toward standardization. *European Journal of Endocrinology* 2008 **159** S41–S44. (doi:10.1530/EJE-08-0284)
- 13 Rosenfeld RG. Is growth hormone deficiency a viable diagnosis? *Journal of Clinical Endocrinology and Metabolism* 1997 **82** 349–351. (doi:10.1210/jcem.82.2.3841)
- 14 Mitchell H, Dattani MT, Nanduri V, Hindmarsh PC, Preece MA & Brook CG. Failure of IGF-I and IGFBP-3 to diagnose growth hormone insufficiency. *Archives of Disease in Childhood* 1999 **80** 443–447. (doi:10.1136/adc.80.5.443)
- 15 Moore DC, Ruvalcaba RHA, Smith EK & Kelley VC. Plasma somatomedin-C as a screening test for growth hormone deficiency in children and adolescents. *Hormone Research* 1982 **16** 49–55. (doi:10.1159/000179484)
- 16 Zweig MH & Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry* 1993 **39** 561–577.
- 17 Cacciari E, Milani S, Balsamo A, Spada E, Bona G, Cavallo L, Cerutti F, Gargantini L, Greggio N, Tonini G *et al.* Italian cross-sectional growth charts for height, weight and BMI (2 to 20 yr). *Journal of Endocrinological Investigation* 2006 **29** 581–593. (doi:10.1007/BF03344156)
- 18 Tanner JM & Whitehouse RH. Clinical longitudinal standards for height, weight, height velocity, weight velocity, and stages of puberty. *Archives of Disease in Childhood* 1976 **51** 170–179. (doi:10.1136/adc.51.3.170)
- 19 Elmlinger MW, Kühnel W, Weber MM & Ranke MB. Reference ranges for two automated chemiluminescent assays for serum insulin-like growth factor I (IGF-I) and IGF-binding protein 3 (IGFBP-3). *Clinical Chemistry and Laboratory Medicine* 2004 **42** 654–664. (doi:10.1515/CCLM.2004.112)
- 20 Müller HL. Childhood craniopharyngioma. *Pituitary* 2013 **16** 56–67. (doi:10.1007/s11102-012-0401-0)
- 21 Maghnie M, Strigazzi C, Tinelli C, Autelli M, Cisternino M, Loche S & Severi F. Growth hormone (GH) deficiency (GHD) of childhood onset: reassessment of GH status and evaluation of the predictive criteria for permanent GHD in young adults. *Journal of Clinical Endocrinology and Metabolism* 1999 **84** 1324–1328. (doi:10.1210/jcem.84.4.5614)
- 22 Loche S, Bizzarri C, Maghnie M, Faedda A, Tzialla C, Autelli M, Casini MR & Cappa M. Results of early reevaluation of growth hormone secretion in short children with apparent growth hormone deficiency. *Journal of Pediatrics* 2002 **140** 445–449. (doi:10.1067/mpd.2002.122729)
- 23 Cacciari E, Tassoni P, Parisi G, Pirazzoli P, Zucchini S, Mandini M, Cicognani A & Balsamo A. Pitfalls in diagnosing impaired growth hormone (GH) secretion: retesting after replacement therapy of 63 patients defined as GH deficient. *Journal of Clinical Endocrinology and Metabolism* 1992 **74** 1284–1289. (doi:10.1210/jcem.74.6.1592872)
- 24 Vottero A, Guzzetti C & Loche S. New aspects of the physiology of the GH-IGF-1 axis. *Endocrine Development* 2013 **24** 96–105. (doi:10.1159/000342573)
- 25 Federico G & Cianfarani S. Usefulness of serum insulin-like growth factor I assessment in the diagnosis of childhood-onset growth hormone deficiency. *Hormone Research* 2010 **74** 145–148. (doi:10.1159/000314895)
- 26 Cianfarani S, Tondinelli T, Spadoni GL, Scirè G, Boemi S & Boscherini B. Height velocity and IGF-I assessment in the diagnosis of childhood-onset GH insufficiency: do we still need a second GH stimulation test? *Clinical Endocrinology* 2002 **57** 161–167. (doi:10.1046/j.1365-2265.2002.01591.x)
- 27 Federico G, Street ME, Maghnie M, Caruso-Nicoletti M, Loche S, Bertelloni S, Cianfarani S & Study Group on Physiopathology of Growth Processes; Council of ISPED. Assessment of serum IGF-I concentrations in the diagnosis of isolated childhood-onset GH deficiency: a proposal of the Italian Society for Pediatric Endocrinology and Diabetes (SIEDP/ISPED). *Journal of Endocrinological Investigation* 2006 **29** 732–737. (doi:10.1007/BF03344184)
- 28 Cianfarani S, Liguori A, Boemi S, Maghnie M, Iughetti L, Wasniewska M, Street ME, Zucchini S, Aimaretti G & Germani D. Inaccuracy of insulin-like growth factor (IGF) binding protein (IGFBP)-3 assessment in the diagnosis of growth hormone (GH) deficiency from childhood to young adulthood: association to low GH dependency of IGF-II and presence of circulating IGFBP-3 18-kilodalton fragment. *Journal of Clinical Endocrinology and Metabolism* 2005 **90** 6028–6034. (doi:10.1210/jc.2005-0721)
- 29 Ghigo E, Bellone J, Aimaretti G, Bellone S, Loche S, Cappa M, Bartolotta E, Damacco F & Camanni F. Reliability of provocative tests to assess growth hormone secretory status. Study in 472 normally growing children. *Journal of Clinical Endocrinology and Metabolism* 1996 **81** 3323–3327. (doi:10.1210/jcem.81.9.8784091)
- 30 Rochiccioli P, Enjaume C, Tauber MT & Pienkowski C. Statistical study of 5473 results of nine pharmacological stimulation tests: a proposed weighting index. *Acta Paediatrica* 1993 **82** 245–248. (doi:10.1111/apa.1993.82.issue-3)
- 31 Binder G, Huller E, Blumenstock G & Schweizer R. Auxology-based cut-off values for biochemical testing of GH secretion in childhood. *Growth Hormone & IGF Research* 2011 **21** 212–218. (doi:10.1016/j.ghir.2011.05.007)
- 32 Erturk E, Jaffe CA & Barkan AL. Evaluation of the integrity of the hypothalamic-pituitary-adrenal axis by insulin hypoglycemia test. *Journal of Clinical Endocrinology and Metabolism* 1998 **83** 2350–2354. (doi:10.1210/jcem.83.7.4980)
- 33 Lange M, Svendsen OL, Skakkebaek NE, Müller J, Juul A, Schmiegelow M & Feldt-Rasmussen U. An audit of the insulin-tolerance test in 255 patients with pituitary disease. *European Journal of Endocrinology* 2002 **147** 41–47. (doi:10.1530/eje.0.1470041)
- 34 Silva EGP, Shessarenko N, Arnhold IJP, Batista MC, Estefan V, Osorio MG, Marui S & Mendonça BB. GH values after clonidine stimulation measured by immunofluorometric assay in normal prepubertal children and GH-deficient patients. *Hormone Research* 2003 **59** 229–233. (doi:10.1159/000070222)
- 35 Wagner IV, Paetzold C, Gausche R, Vogel M, Koerner A, Thiery J, Arsene CG, Henrion A, Guettler B, Keller E *et al.* Clinical evidence-based cutoff limits for GH stimulation tests in children with a backup of results with reference to mass spectrometry. *European Journal of Endocrinology* 2014 **171** 389–397. (doi:10.1530/EJE-14-0165)

Received 5 February 2016

Revised version received 28 April 2016

Accepted 4 May 2016