Original Article

Pragmatic Evaluation of Growth Hormone Stimulation Tests in Short Stature

Rahul Gupta*, Aashima Dabas[1]*, Shweta Kohli, Rekha Ramot, Viveka P. Jyotsna, Alpesh Goyal, Yashdeep Gupta, Rajesh Khadgawat

Department of Endocrinology and Metabolism, All India Institute of Medical Sciences, New Delhi, [1]Department of Pediatrics, Maulana Azad Medical College and Lok Nayak Hospital, New Delhi, India

*Both contributed equally

Abstract

Introduction: To assess the performance of growth hormone stimulation tests (GHSTs) in the evaluation of short stature. **Methods:** It was a single-centre retrospective study carried out in children evaluated for short stature between January 2005 to March 2020. The clonidine stimulation test (CST) and glucagon stimulation test (GST) were used to assess growth hormone (GH) reserve (GST was performed only when peak GH levels were between 5 to ≤10 ng/mL on CST). A GH level of <5 ng/mL on CST or ≤10 ng/ml on both was used to corroborate GH deficiency. **Results:** A total of 556 children were eligible for this study. The mean (SD) age was 12.9 (3.5) years, and 66.3% were male. The peak GH level [median (IQR)] was 5.50 ng/ml (1.90 – 7.50) on CST (at 60 minutes) and 7.45 ng/ml (2.15 – 10.77) on GST (at 120 minutes). On restricting sampling to two time points, the false positive rate was 13.6% on CST (60, 90 minutes) and 11.5% on GST (120, 150 minutes). Similarly, restricting to three time points was associated with a false positive rate of 8.5% on CST (60, 90, 120 minutes) and 3.8% on GST (90, 120, 150 minutes). Using the treating clinician-determined diagnosis of GHD as a reference standard, the optimal cut-off of peak GH on CST was 7.79 ng/ml (sensitivity: 83.8%; specificity: 89.4%). **Conclusion:** Restricting the GH sampling to fewer time points is associated with an increase in the false positivity rate (FPR).

Keywords: Diagnosis, growth, growth hormone stimulation test, short stature

Introduction

Short stature encompasses clinical phenotypes such as growth hormone deficiency (GHD), growth plate abnormalities, idiopathic short stature, familial short stature, and syndromic short stature.[1] Among these, GHD remains a challenge to diagnose and treat. Growth hormone stimulation tests (GHST) have been a standard diagnostic tool to detect pituitary growth hormone (GH) reserve.[2] Pharmacological provocative stimuli such as insulin-induced hypoglycaemia, glucagon, arginine, clonidine, L-dopa, and GH-releasing peptide-2 have shown variable performance in detecting GH reserve but are recommended to improve the objectivity of diagnosis.[3] A diagnosis of GHD requires at least two failed provocative tests or a single failed test with high clinical suspicion.[4]

Among the agents available for GHST, oral clonidine and glucagon remain the most feasible and simple to administer.[3,5] The performance of these two agents has been evaluated in a limited number of studies.[6-8] The performance of GHST also varies in different clinical indications that carry diagnostic connotations. Significantly low GH levels suggest severe GHD and a good response to recombinant GH is likely, while apparently normal growing children may also test below acceptable limits.

Given the existing controversies in the interpretation of GHST, there is a need to evaluate the performance of these tests in different clinical settings. The present study aimed to evaluate (1) the feasibility of restricting GH sampling to lesser time points in GHST, (2) diagnostic performance in detecting

Address for correspondence: Dr. Rajesh Khadgawat, Room No. 303, 3rd Floor, Biotechnology Block, All India Institute of Medical Sciences, New Delhi, India.
E-mail: rajeshkhadgawat@hotmail.com

Submitted: 28-Jul-2023 **Revised:** 09-Sep-2023
Accepted: 24-Sep-2023 **Published:** 29-Apr-2024

Supplementary material available online

Access this article online

Quick Response Code:

Website: https://journals.lww.com/indjem/

DOI: 10.4103/ijem.ijem_326_23

How to cite this article: Gupta R, Dabas A, Kohli S, Ramot R, Jyotsna VP, Goyal A, *et al*. Pragmatic evaluation of growth hormone stimulation tests in short stature. Indian J Endocr Metab 2024;28:137-44.

GH deficiency, and (3) the effect of body mass index (BMI), pubertal status, and gender on GHST.

Materials and Methods

This was a single-centre, retrospective observational study carried out at the Paediatric and Adolescent Clinic of a tertiary care hospital. Children of chronological age between 1 to 18 years who underwent GHST between January 2005 and March 2020 were eligible. In case of duplication, the record of only the first GHST was included.

Management of participants

Children with short stature were systematically evaluated at the Paediatric and Adolescent Clinic. The evaluation included history, physical examination and anthropometry, and investigations. The anthropometric measurements were made by trained staff with children dressed in minimal light clothing and without footwear. Height was measured to the nearest 0.1 cm using Holtain's stadiometer (Holtain Inc., Crymych, Pembs., UK) and weight was measured to the nearest 0.1 kg using an electronic scale. The stadiometer and scale were calibrated using standard height and weight, respectively. Mid-parental height was calculated in cm as (father's height + mother's height + 13)/2 for boys and (father's height + mother's height –13)/2 for girls. BMI was calculated as weight (kg)/height (m) ^{2}. Height, weight, and BMI were expressed as standard deviation scores (SDSs) using country- and gender-specific reference charts for boys and girls, KN Agarwal charts from 2000–2020, and revised IAP growth charts thereafter.[9,10] The pubertal stage was estimated according to the Tanner staging. A testicular volume of ≥4 ml in boys and the larche (≥B2) in girls was considered as the onset of puberty.

Investigations were done using a three-tiered approach. The first-tier investigations included complete blood counts, renal and liver function tests, fasting plasma glucose, serum calcium, phosphate, IgA tissue transglutaminase (tTg), thyroid function tests, FSH for all girls, urine routine microscopy and pH, and bone age (BA) estimation (anteroposterior radiograph of the left hand and wrist, interpreted as per Greulich and Pyle atlas).[11] The second-tier investigation included the measurement of serum IGF-1 and GHST. Additional investigations under second-tier included anterior pituitary hormone (serum cortisol, ACTH, dehydroepiandrosterone, prolactin, estradiol) and whole blood karyotype (based on clinical suspicion).

GHSTs were performed in a fasting state after achieving an euthyroid and eucortisolemic state. A clonidine stimulation test (CST) was performed using tablet clonidine, orally in a dose of 0.15 mg/m^2 followed by blood samples for GH at 0, 30, 60, 90, and 120 minutes.[2] According to the departmental protocol, a peak serum GH level <5 ng/mL on CST (with supporting anthropometrics and investigations) was considered adequate to corroborate the presence of GHD. In patients with peak serum GH levels of 5 to ≤10 ng/mL, a second GHST using intramuscular glucagon (GST) in a dose of 0.03 mg/kg (maximum dose 1 mg) was conducted after an interval of 48 hours, with samples for GH collected at 0, 60, 90, 120, 150 and 180 minutes. A peak serum GH level ≤10 ng/mL on GST was considered to corroborate the presence of GHD.[2,3] Both tests were considered congruent when the peak GH level was ≤10 ng/mL in both and incongruent when paired GHST suggested peak GH level ≤10 ng/mL on CST and >10 ng/mL on GST [Figure 1]. Sex-steroid priming was not performed for peripubertal children as per departmental protocol. The children underwent further workup (third-tier investigations) as per the clinical indication that included

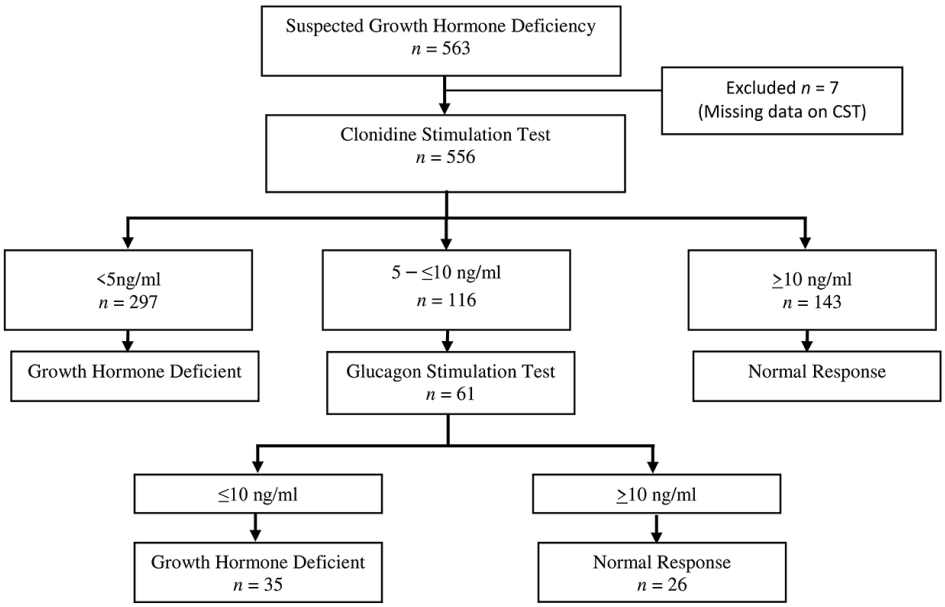

Figure 1: Flow of participants

magnetic resonance imaging of the sella. The final aetiology was determined by the treating clinician based on the findings from history, auxology and examination, biochemical investigations, and imaging.[4]

Serum GH and IGF-1 estimation was done by CLIA using the Diasorin Liaison auto-analyzer (Diasorin Inc., Stillwater, MN, USA). The recombinant 22 kDa GH standard (WHO 98/574) was adopted as a reference standard for the GH assay. The functional and analytical sensitivity of the GH assay was 0.05 ng/ml and 0.1 ng/ml, respectively. Serum thyroid stimulating hormone (TSH), luteinizing hormone (LH), follicle stimulating hormone (FSH), prolactin, testosterone, estradiol and plasma adrenocorticotropic hormone (ACTH) estimation were conducted by ECLIA assay using Cobas e-411 auto-analyzer (Roche Diagnostics, Mannheim, Germany).

Statistical analysis

Statistical analysis was carried out using SPSS for Windows (SPSS 21.0, SPSS Inc., Chicago, IL, USA). Continuous variables were presented as mean (±SD) or median (IQR) as appropriate and categorical variables as proportion. Independent Student's *t*-test, Wilcoxon rank-sum test, and Mann-Whitney were used for the comparison of continuous variables as appropriate. The Chi-square test was used to compare proportions. To estimate the optimal GH cut-off to detect a GH-sufficient state, receiver operating curves were created. In analysis involving the proportion of individuals attaining their peak GH levels at a given time point, those with all GH values of <0.1 ng/ml and where the difference between the GH levels on GHSTs was <0.05 ng/ml were excluded. A *P* value of < 0.05 was considered statistically significant in a two-tailed test

Ethical Aspect

The study was approved by the Institutional ethics committee (IEC-184/04.03.2022). The need for informed written consent was waived off as the study procedures did not reveal the patient's identity and its retrospective study design. The study was conducted in accordance with the principles outlined in Declaration of Helinski(2013).

Results

The flow of participants is shown in Figure 1. A total of 563 individuals were subjected to GHST and after excluding those with missing data (n = 7), 556 (66.3% males) were eligible. The mean (SD) age was 12.9 (3.5) years and 33.3% were prepubertal. The treating clinician determined diagnosis of idiopathic growth hormone deficiency (IGHD), multiple pituitary hormone deficiency (MPHD), idiopathic short stature (ISS), constitutional delay in growth and puberty (CDGP), and, others were present in 285 (51.3%), 91 (16.4%), 23 (4.1%), 88 (15.8%), and 68 (12.2%) children, respectively.

Of the 556 children who underwent GHST with clonidine, 297 (53.4%), 116 (20.8%), and 143 (25.7%) had peak GH levels of <5 ng/ml, 5 to ≤10 ng/ml, and >10 ng/ml, respectively. Of the 116 children with peak GH levels between 5 to ≤10 ng/ml on CST, 61 underwent GST. The GH levels across different time points on CST and GST are shown in Table 1. As compared to CST where peak levels were achieved at 60 minutes, the peak levels on GST were higher [7.45 (2.15 – 10.77) vs. 5.50 (1.90 – 7.50); *P* = <0.001] and attained at 120 minutes. The proportion of children attaining their peak GH levels at 0, 30, 60, 90, and 120 minutes on CST were 10.9%, 10.1%, 39.3%, 26.7%, and 13.0%, respectively. Using GST, the proportion of children attaining their peak GH levels at 0, 60, 90, 120, 150, and 180 minutes were 4.9%, 6.6%, 26.2%, 39.3%, 18.0%, and 4.9%, respectively.

Using peak GH >10 ng/ml as a reference standard, the false positivity rate (FPR) on restricting the GH sampling to fewer time points is shown in Table 2. If sampling was restricted to two, the time points associated with the lowest FPRs were 60 and 90 minutes for CST (FPR: 13.6%) and 120 and 150 minutes for GST (FPR: 11.5%). The best time points if the number of samples were restricted to three were 60, 90, and 120 minutes for CST (FPR: 7.2%) and 90, 120, and 150 minutes for GST (FPR: 3.8%%). Sampling at 60, 90, 120, and 150 minutes on GST performed similarly to 0, 60, 90, 120, 150, and 180 minutes; while CST sampling at 30, 60, 90, and 120 minutes gave the lowest FPR of 4.3% amongst all time-restricted sampling strategies. In children with severe short stature (Height [Ht] SDS <–3.0), restricting

Table 1: Growth hormone levels after clonidine and glucagon stimulation test at various time points

Time points	Clonidine stimulation test (n=556)	Clonidine stimulation test* (n=116)	Glucagon stimulation test (n=61)	$P^{\#}$
0 min	0.41 (0.11–1.48)	0.74 (0.19–2.30)	0.32 (0.14–1.45)	0.029
30 min	0.57 (0.13–2.50)	1.21 (0.30–3.60)	0.43 (0.14–2.12)	0.014
60 min	1.70 (0.33–6.8)	5.50 (1.90–7.50)	0.44 (0.15–1.05)	<0.001
90 min	1.95 (0.49–6.7)	4.91 (2.53–6.47)	1.40 (0.37–6.15)	0.117
120 min	1.30 (0.34–3.97)	2.60 (1.50–5.00)	7.45 (2.15–10.77)	<0.001
150 min	-	-	4.00 (1.10–8.70)	
180 min	-	-	2.75 (120–4.87)	

GH values are presented as median (IQR) *Individuals with peak GH levels between 5–≤10 ng/ml on clonidine stimulation test. $^{\#}P$ value between individuals with paired data available on CST and GST

Table 2: False positivity rate with various time-restricted GH sampling strategies with clonidine and glucagon stimulation test

Clonidine Stimulation Test (n=556)			Glucagon Stimulation Test (n=61)		
Sampling Time	**False Positive* (%)**	**True Negative# (%)**	**Sampling Time**	**False Positive* (%)**	**True Negative# (%)**
0, 30, 60, 90, and 120 min	0	100	0, 60, 90, 120, 150, and 180 min	0	100
30, 60, 90, and 120 min	4.3	95.7	60, 90, 120, and 150 min	0	100
0, 60, 90, and 120 min	5.0	95.0	90, 120, 150, and 180 min	3.8	96.2
0, 30, 60, and 90 min	6.4	93.6	60, 120, 150, and 180 min	4.0	96.0
60, 90, and 120 min	7.2	92.8	90, 120, and 150 min	3.8	96.2
30, 60, and 90 min	10.0	90.0	60, 120, and 150 min	7.7	92.3
0, 60, and 90 min	10.0	90.0	120, 150, and 180 min	8.0	92.0
			0, 120, and 150 min	11.5	88.5
60 and 90 min	13.6	86.4%	120 and 150 min	11.5	88.5

*False positive: wrongly labelled as having growth hormone deficiency; #True negative: correctly labelled as not having growth hormone deficiency

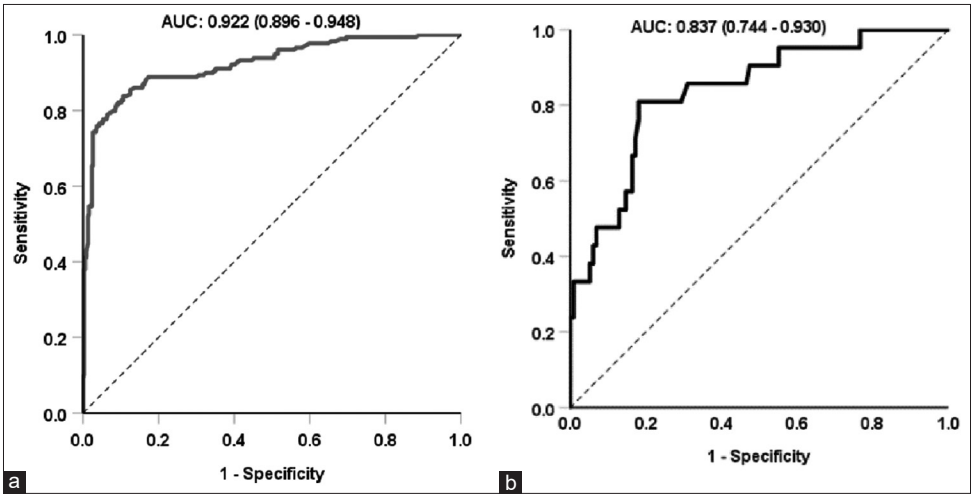

Figure 2: (a) ROC curve of peak growth hormone value on clonidine stimulation for detecting growth hormone sufficiency; (b) ROC curve of peak growth hormone value on glucagon stimulation for diagnosing growth hormone sufficiency

the samples in CST to two, three, and four time points gave a similar FPR of 13.0% (60, 90 minutes), 9.1% (60, 90, 120 minutes), and 5.2% (30, 60, 90, 120 minutes). However, restricting the samples in GST to two (120,150 minutes), three (90,120,150 minutes), and four time points (60, 90, 120, 150 minutes) yielded no false positive cases.

Using treating clinician-determined diagnosis as a reference standard, the Receiver-operating characteristic (ROC) curve of peak GH levels on CST and GST in detecting GH-sufficient states is shown in Figure 2. The area under curve (AUC) for CST and GST were 0.922 (0.896 –0.948) and 0.837 (0.744 –0.930), respectively. The optimal cut-off of peak GH for CST and GST were >7.79 ng/ml (sensitivity: 83.8% and specificity: 89.4%) and >10.25 ng/ml (sensitivity: 81.0% and specificity: 81.9%), respectively.

The effect of BMI, puberty, gender, and aetiology of GHD on GH levels and its timing is shown in Figure 3. On CST, GH levels were higher in patients with BMI SDS <–1.0 [3.20 (0.80 – 5.74) ng/ml vs. 1.66 (0.40 – 5.00) ng/ml; P = <0.001], in peripubertal children [5.80 (1.80 – 10.90) ng/ml vs. 2.70 (0.52 – 9.52) ng/ml; P = 0.002], and those having IGHD as compared to MPHD [3.35 (1.08 – 5.80) ng/ml vs. 0.41 (0.14 – 1.40) ng/ml; P = <0.001]. The GH levels were similar in males and females (P = 0.311) [Supplementary Table 1]. In addition, the time to attain peak GH was similar across BMI (BMI SDS <–1.0 vs. >–1.0; P = 0.974), pubertal status (prepubertal vs. peripubertal; P = 0.144), gender (male vs. female; P = 0.116), and aetiology (IGHD vs. MPHD; P = 0.550) [Supplementary Table 2].

Similarly, on GST GH levels were higher in patients with BMI SDS <–1.0 [7.00 (3.20 – 11.85) ng/ml vs. 2.90 (0.60 – 7.60) ng/ml; P = 0.018], peripubertal children [10.90 (5.20 – 20.40) ng/ml vs. 7.10 (5.10 – 10.20) ng/ml; P = 0.026], and in IGHD in comparison to MPHD [5.74 (2.05 – 9.11) ng/ml vs. 0.51 (0.22 – 0.97) ng/ml; P = <0.001] [Supplementary Table 1]. However, it was similar in males and females (P = 0.145). In addition, the time to attain peak GH was similar across BMI (BMI SDS <–1.0 vs. >–1.0; P = 0.479), pubertal

status (prepubertal vs. peripubertal; $P = 0.866$), gender (male vs. female; $P = 0.467$), and aetiology (IGHD vs. MPHD; $P = 0.774$) [Supplementary Table 2].

The comparison of children with peak GH levels of <5 ng/ml, 5–10 ng/ml, and >10 ng/ml on CST is shown in Table 3. As compared to those with GH levels of >10 ng/ml and 5 ng/ml to ≤10 ng/ml, those with peak GH levels <5 ng/ml were shorter, had higher BMI, substantially shorter for their genetic potential, had more delay in bone age, and had lower serum IGF-1, respectively. In comparison, children with peak GH levels between 5 ng/ml to ≤10 ng/ml and >10 ng/ml were similar to each other, except for being more obese and having a lesser delay in bone age, respectively. Of the 116 children with peak GH levels between 5–10 ng/ml on CST, the results on GST were available for 61 (52.6%). Amongst these, a concordant response (peak GH ≤ 10 ng/ml on GST) was seen in 57.4%. In children with peak GH levels of 5–7 ng/ml on CST, a discordant response on GST was seen in 48.0%. The comparison of children subjected to paired GHST (CST and GST) is shown in Table 3.

The comparison of various aetiologies of short stature is shown in Table 4. As compared to MPHD, those with IGHD had similar height and weight but had lower BMI, lesser bone age delay, and higher serum IGF-1 levels and peak GH values on CST and GST. As compared to ISS, those with IGHD presented at a younger age were shorter, were shorter in relation to MPHD, and had lower IGF-1 levels and peak GH levels on CST and GST. As compared to CDGP, those with IGHD were shorter, weighed more, were shorter in relation to MPH, and had lower serum IGF-1 levels and peak GH levels on CST and GST.

Discussion

The study has evaluated the performance of clonidine and glucagon as provocative agents for assessing GH reserve in

Table 3: Comparison of children with short stature based on peak GH levels on clonidine stimulation and glucagon stimulation test

Peak GH Levels on Clonidine Stimulation				Peak GH Levels on Glucagon Stimulation Test			
	<5 ng/ml (*n*=297)	5-10 ng/ml (*n*=116)	>10 ng/ml (*n*=143)	<10 ng/ml (*n*=35)	>10 ng/ml (*n*=26)	*P*
Age	13.2 (10.2, 15.1)	13.0 (11.0,15.3)	13.1 (9.8,15.2)	Age	12.4 (9.0,15.3)	13.9 (11.7, 15.7)	0.135
Height SDS[ac]	-3.52 (–2.59, –4.86)	–3.10 (–2.54, –3.81)	–3.07 (–2.53, –3.76)	Height SDS	–3.29 (–2.39, –3.97)	–2.84 (–2.31, –3.44)	0.159
Weight SDS[b]	–2.48 (–1.02, –3.71)	–2.34 (–1.51, –3.15)	–2.78 (–1.92, –3.34)	Weight SDS	–2.33 (–1.46, –3.16)	–2.58 (–1.56, –3.35)	0.498
BMI SDS[abc]	–0.47 (0.56, –1.55)	–0.79 (–0.05, –1.64)	–1.49 (–0.65, –2.05)	BMI SDS	–0.58 (0.25, –1.44)	–1.36 (–0.35, –2.11)	0.021
MPH SDS[a]	–0.92 (–0.47, –1.53)	–1.14 (–0.63, –1.56)	–1.01 (–0.49, –1.58)	MPH SDS	–1.09 (–0.48, –1.56)	–1.08 (–0.73, –1.74)	0.748
MPH SDS – Height SDS[ac]	2.56 (1.66,3.68)	1.85 (1.30,2.77)	2.01 (1.57,2.73)	MPH SDS – Height SDS	2.09 (1.38, 3.03)	1.65 (0.89, 2.44)	0.133
BA-CA[ab]	3.6 (2.1, 4.9)	2.7 (1.2, 3.8)	3.5 (2.0, 4.4)	BA-CA	2.2 (1.1, 3.8)	3.0 (2.0, 3.8)	0.268
Serum IGF-1[ac], ng/mL	49.6 (18.9,129.5)	128.0 (69.7,195.5)	163 (82.6, 290.9)	Serum IGF-1, ng/mL	96.3 (27.7, 183.8)	115.8 (52.5, 288.5)	0.731

Data is presented as Median (IQR) $P<0.05$ between: [a] patient with GH <5 and 5–10 ng/mL; [b] between 5–10 ng/mL and >10 ng/mL; [c]between <5 ng/mL and >10 ng/mL for CST; BA-CA: Difference between bone age and chronological age; BMI: Body mass index; MPH: mid-parental height; SDS: standard deviation score

Table 4: Comparison of various aetiologies of short stature

	IGHD	MPHD	ISS	CDGP
Age	13.1[b] (10.5–15.1)	13.1[d] (10.1–15.3)	15.0[b,d,f] (14.3–15.7)	13.1[f] (10.1–15.2)
Height SDS	–3.40[b,c] (–2.60 – –4.75)	–3.40[d,e] (–2.52 – –4.64)	–2.57[b,d,f] (–2.20 – –3.02)	–3.05[c,e,f] (–2.54 – –3.68)
Weight SDS	–2.49 (–1.27 – –3.63)	–2.21[e] (–1.26 – –3.42)	–2.18[f] (–1.42 – –2.98)	–2.93[e,f] (–2.16 – –3.35)
BMI SDS	–0.72[a,c] (0.34 – –1.22)	–0.37[a,d,e] (0.32 – –1.22)	–1.36[d] (0.20 – –2.10)	–1.64[c,e] (–0.76 – –2.06)
MPH SDS–Height SDS	2.29[b,c] (1.53–3.44)	2.87[d,e] (1.84–3.69)	1.61[b,d] (1.34–2.22)	1.93[c,e] (1.41–2.55)
BA-CA	3.1[a] (1.8–4.5)	4.5[a,d,e] (2.9–6.3)	3.5[d] (1.7–4.6)	3.4 [e] (2.1–4.3)
Serum IGF-1, ng/mL	86.1[a,b,c] (27.1–167.9)	24.9[a,d,e] (13.1–64.2)	291.0[b,d] (160.0–398)	163.6[c,e] (81.1–215.0)
Peak GH value on CST (ng/ml)	3.35[a,b,c] (1.06–5.92)	0.41[a,d,e] (0.14–1.40)	10.10[b,d,f] (7.80–19.50)	17.05[c,e,f] (12.07–25.95)
Peak GH value on GST (ng/ml)	5.70[a,b,c] (2.00–8.95)	0.51[a,d,e] (0.22–0.97)	20.00[b,d] (11.65–37.45)	31.60[c,e] (9.10–39.75)

[a]$P<0.05$ between GHD and MPH; [b]$P<0.05$ between GHD and ISS; [c]$P<0.05$ between GHD and CDGP; [d]$P<0.05$ between MPHD and ISS; [e]$P<0.05$ between MPHD and CDGP; [f]P between ISS and CDGP

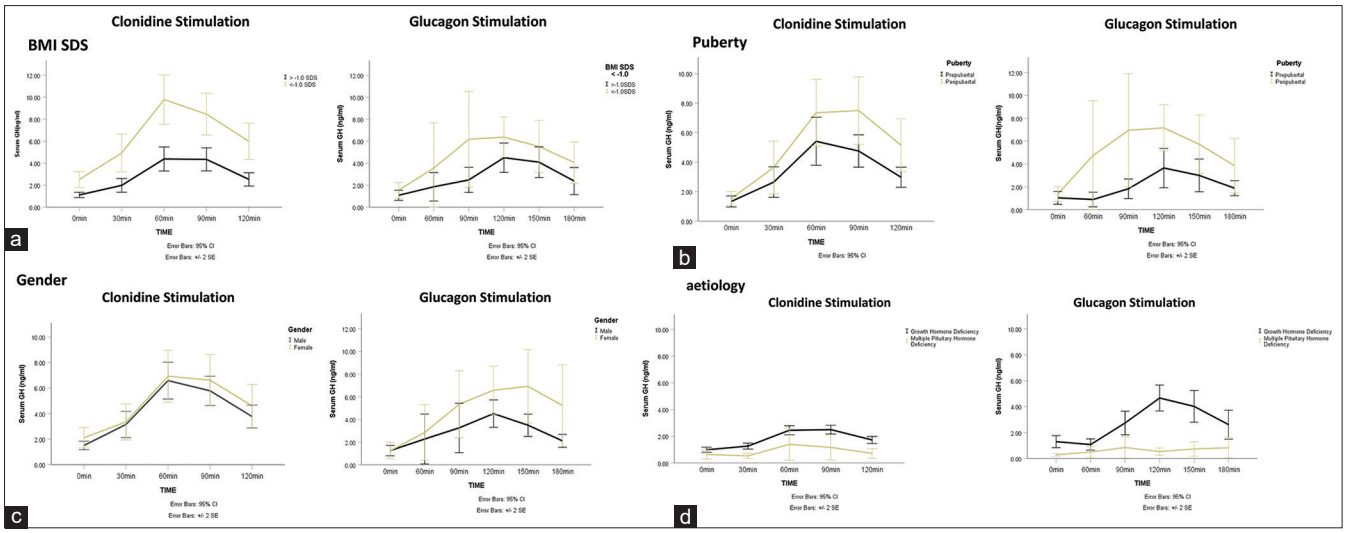

Figure 3: The effect of (a) BMI, (b) puberty, (c) gender, and (d) aetiology of growth hormone deficiency, on peak growth hormone levels on clonidine and glucagon stimulation

563 children with short stature. The key findings from the study are decreasing the number of samples in GHSTs is associated with an increase in false positive rates and BMI influences the peak GH levels.

GHSTs are used in the evaluation of short stature to substantiate the diagnosis of GH deficiency.[4] While many agents are available for stimulating GH secretion, the two most commonly used are clonidine and glucagon, as they are readily available, have lower cost, a better safety profile, and are easy to administer. For these reasons, these were used in the current study. Performing GHST entails multiple blood samples for GH estimation for 2 to 3 hours, which can be challenging in a short child.[12] Furthermore, robust evidence supporting the timing of GH sampling is lacking. In this regard, the findings from the current study suggest that a progressive decrease in the number of samples for GH estimation is associated with an increasing FPR (overdiagnosis). Interestingly, a time-restricted strategy of sampling at 60, 90, 120, and 150 minutes on GST gave the same diagnostic yield as that of conventional sampling at 0, 60, 90, 120, 150, and 180 minutes. However, given the small number of children who were subjected to GST, this finding needs to be evaluated in a larger number. Lastly, decreasing the number of samples only in children with severe short stature (Ht SDS <–3.0) also yielded a similar FPR on CST. The study from Georeli, *et al*.[13] (n = 56) and Thakur, *et al*.[14] (n = 79) using CST estimated a false positive rate of 20% if peak GH levels were estimated at the abovementioned time points only. In studies by Galluzzi, *et al*.[15] (n = 291), Christoforidis, *et al*.[6] (n = 258), and Morris, *et al*.[16] (n = 30), using CST with GH sampling at 0, 30, 60, and 90 minutes the false positive rates were 7.9%, 4.43%, and 3.3%, respectively. In the study by Christoforidis, *et al*.,[6] which also evaluated GST, a similar approach of skipping GH estimation at 180 minutes gave a false positive rate of 5.75%. Overall, these findings suggest that every attempt should be made to sample for GH up to 120 minutes in CST (0, 30, 60, 90, 120 minutes) and 180 minutes in GST (0, 60, 120, 150, 180 minutes) to reduce the false positive rates to avoid overtreatment.

The second issue pertaining to the GHST is the cut-off for peak GH levels to diagnose GHD. The cut-offs are arbitrary and have decreased over time despite improvement in assays analytical sensitivity. In the recent guidelines by the Growth Hormone Research Society, a 7 ng/ml threshold was suggested by participating delegates.[4] The findings from the current study were in agreement, where a cut-off of 7 ng/ml on CST was estimated to provide a sensitivity and specificity of 89.4% and 83.8%, respectively. The validity of this cut-off was further supported by the finding of children with peak GH level <5 ng/ml representing an auxologically distinct subgroup with them being shorter, shorter in relation to mid-parental height, having more delayed bone age, a higher BMI, and low serum IGF-1 levels. However, the cut-off of 7 ng/ml needs to be interpreted cautiously. The first issue pertains to the use of reference standards to define GHD. In this regard, the guideline recommends that the diagnosis of GHD is a clinical one and is based on the findings of auxology, laboratory investigations, and imaging. In addition, it goes on to mention that GHD should not be diagnosed solely based on GHST.[4] To reflect these recommendations, 'treating clinician diagnosis' was used as a reference standard in the current study. Such an approach is likely to reflect the holistic assessment of a child by an endocrinologist with years of experience in managing short stature. This was further substantiated by the findings on the comparison of short stature due to various aetiologies [Table 4], where children with IGHD were found to have substantial differences in auxology and investigations in comparison to other aetiologies. However, given the retrospective nature of the study, a uniform method of labelling a short child as GHD cannot be assured due to evolving knowledge, changes in practices, and available investigations.

The second issue pertains to GST, as it was only performed in selected children the applicability of findings will be only to those who have intermediate peak GH levels (5–10 ng/ml) on CST. Lastly, the test being evaluated (GHST) was also a component of a reference standard.

The third issue pertaining to the GHST is the number and sequence to diagnose GHD. The guidelines have recommended two separate provocative tests with different agents.[4] This has been recommended to reduce the FPR due to the limitations of GHST. A few of these are their non-physiological nature, they assess GH reserve and not integrated GH secretion, have poor intraindividual reproducibility, and inter-assay differences amongst GH assays. In the current study, two GHSTs (CST and GST) were used and the use of a second GHST was confined to those having peak GH levels between 5 ng/ml to ≤10 ng/ml. This approach was used to reduce the number of GHST in children with high probability GHD; as suggested by auxology and peak GH levels of <5 ng/ml on CST. This was in agreement with current guidelines that have recommended using a single GHST for diagnosis of GHD in children with high-probability GHD as suggested by auxology and serum IGF-1 levels. In addition, the study from Dori, *et al*.[7] suggested that the probability of a second positive test is increased if peak GH levels are <5 ng/ml on the first GHST. While the concordance rate between CST and GST at peak GH level of <5 ng/ml on CST cannot be ascertained from the current study due to non-availability of GST in these children, it was 57.3% when peak GH levels were 5 ng/ml to ≤10 ng/ml on CST. A low concordance rate when peak GH levels are in the intermediate range (5 to ≤10 ng/ml) supports the use of two GHSTs. In regard to the sequence of GHST, there is a dearth of data supporting the superiority of one over another. In the current study, clonidine was used as the first provocative agent and glucagon as the second. This preference was primarily based on the low cost, easy availability, and ease of administration (oral vs. IM) with clonidine. The sequence of CST followed by GST is also supported by the findings from Yackobovitch-Gavan, *et al*.,[8] who have reported a higher occurrence of divergent results in GST first than CST first approach. A similar concordant response of 56.5% has been reported by Dori, *et al*. (n = 200).[7] Overall, these findings suggest that the use of two GHSTs is likely to continue due to the FPR of individual GHSTs. In short children with a high pretest probability of GHD, extremely low levels (<5 ng/ml) on the first GHST can obviate the need for the second. Almost three-fourths of the children can be appropriately identified using clonidine as the first provocative agent. In the remaining one-fourth, almost half will have a concordant response using glucagon as the second provocative agent.

The last issue pertains to the effect of BMI on peak GH levels. The current study showed a negative effect of BMI on peak GH levels on both CST and GST. A negative correlation between peak GH levels and BMI in GHD is documented earlier as well and is postulated to impair GH response in children with GHD.[17,18]As per a meta-analysis, a per unit increase in BMI decreases the peak GH level by 11.6%.[17] These findings suggest that there is a need for BMI SDS-specific peak GH cut-off on GHST. Till the time they are available, one should remain cautious of over-diagnosing GHD in obese children.

The study is one of the largest to evaluate the performance of GHST in short children. However, there are certain limitations: (1) retrospective study design; (2) missing data on GST and serum IGF-1 levels; (3) use of only clonidine and glucagon as provocative agents; (4) use of treating clinician-determined diagnosis as a reference standard for GHD; and (5) availability of data on GST in selected children (peak GH levels between 5 ng/ml to ≤10 ng/ml on CST).

Conclusion

The peak GH levels are seen between 60 to 90 minutes on CST and 150 to 180 minutes on GST. Restricting the GH sampling to a lesser time point is associated with an increase in false positivity. BMI negatively affects the peak GH levels, and caution should be used while interpreting GHST in overweight or obese children.

Acknowledgment

None.

Authors' contribution

RG was responsible for patient management, data analysis, manuscript drafting and revisions. AD handled patient management, data collection and analysis, manuscript drafting and revisions. SK and RR contributed to data collection. VPJ, AG, YG were involved in patient management and manuscript revisions. RK managed patients, collected and analyzed data, and drafted and revised the manuscript.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Wit JM, Oostdijk W, Losekoot M, van Duyvenvoorde HA, Ruivenkamp CA, Kant SG. Mechanisms in endocrinology: Novel genetic causes of short stature. Eur J Endocrinol 2016;174:R145-73.
2. Kamoun C, Hawkes CP, Grimberg A. Provocative growth hormone testing in children: How did we get here and where do we go now? J Pediatr Endocrinol Metab 2021;34:679-96.
3. Grimberg A, DiVall SA, Polychronakos C, Allen DB, Cohen LE, Quintos JB, *et al.* 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 2016;86:361-97.
4. Collett-Solberg PF, Ambler G, Backeljauw PF, Bidlingmaier M, Biller BMK, Boguszewski MCS, *et al.* Diagnosis, genetics, and therapy of short stature in children: A growth hormone research society international perspective. Horm Res Paediatr 2019;92:1-14. doi: 10.1159/000502231.
5. Ibba A, Guzzetti C, Casula L, Salerno M, Di Iorgi N, Allegri AME, *et al*. Reliability of clonidine testing for the diagnosis of growth hormone deficiency in children and adolescents. Clin Endocrinol (Oxf) 2018;89:765-70.

6. Christoforidis A, Triantafyllou P, Slavakis A, Katzos G. Clonidine and glucagon stimulation for testing growth hormone secretion in children and adolescents: Can we make it with fewer samples? J Endocrinol Invest 2012;36:1046-50.
7. Dori EB, Avnon Ziv C, Auerbach A, Greenberg Y, Zaken H, Levy-Khademi F. The inter-Test variability of growth hormone stimulation tests and factors affecting this variability. Growth Horm IGF Res 2020;55:101361. doi: 10.1016/j.ghir. 2020.101361.
8. Yackobovitch-Gavan M, Lazar L, Diamant R, Phillip M, Oron T. Diagnosis of growth hormone deficiency in children: The efficacy of glucagon versus clonidine stimulation test. Horm Res Paediatr 2020;93:470-6.
9. Khadilkar VV, Khadilkar AV, Choudhury P, Agarwal KN, Ugra D, Shah NK. IAP growth monitoring guidelines for children from birth to 18 years. Indian Pediatr 2007;44:187-97.
10. Indian Academy of Pediatrics Growth Charts Committee, Khadilkar V, Yadav S, Agrawal KK, Tamboli S, Banerjee M, *et al.* Revised IAP growth charts for height, weight and body mass index for 5- to 18-year-old Indian children. Indian Pediatr 2015;52:47-55.
11. Greulich WW, Pyle SI, editor. Radiographic Atlas of Skeletal Development of the Hand and Wrist. 2nd ed. Repr. Stanford, Calif: Stanford Univ. Press; 2011.
12. Bright GM, Julius JR, Lima J, Blethen SL. Growth hormone stimulation test results as predictors of recombinant human growth hormone treatment outcomes: Preliminary analysis of the national cooperative growth study database. Pediatrics 1999;104:1028-31.
13. Georeli I, Triantafyllou P, Dimitriadou M, Slavakis A, Christoforidis A. Timing of GH peak in both glucagon and clonidine tests is of major clinical importance. Endocr Pract 2019;25:800-8.
14. Thakur DS, Bhagwat NM, Bhide MM, Yerawar CG, Ghanekar GA, Sonawane AB, *et al.* Clonidine stimulation test: Is single best time point, convenient yet efficacious? Indian J Endocrinol Metab 2018;22:511-4.
15. Galluzzi F, Stagi S, Parpagnoli M, Losi S, Pagnini I, Favelli F, *et al.* Oral clonidine provocative test in the diagnosis of growth hormone deficiency in childhood: Should we make the timing uniform? Horm Res 2006;66:285-8.
16. Morris AH, Harrington MH, Churchill DL, Olshan JS. Growth hormone stimulation testing with oral clonidine: 90 minutes is the preferred duration for the assessment of growth hormone reserve. J Pediatr Endocrinol Metab 2001;;14:1657-60.
17. Abawi O, Augustijn D, Hoeks SE, de Rijke YB, van den Akker ELT. Impact of body mass index on growth hormone stimulation tests in children and adolescents: A systematic review and meta-analysis. Crit Rev Clin Lab Sci 2021;58:576-95.
18. Lee NY, Kim SE, Kim S, Ahn MB, Kim SH, Cho WK, *et al.* Effect of body mass index on peak growth hormone level after growth hormone stimulation test in children with short stature. Ann Pediatr Endocrinol Metab 2021;26:192-8.

Supplementary Table 1: Peak GH levels on CST and GST at various time points stratified by BMI, Age, Gender, and etiology

	BMI SDS		Puberty		Gender		Aetiology	
	>-1.0SDS	<-1.0SDS	Prepubertal	Peripubertal	Male	Female	IGHD	MPHD
Clonidine Stimulation Test								
	1.66 (0.40-5.0)	3.20 (0.80-5.74)	2.70 (0.52-9.52	5.80 (1.80-10.90)	1.75 (0.53-5.26)	3.00 (0.47-5.40)	3.35 (1.08-5.80)	0.41 (0.14-1.40)
P	<0.001		0.02		0.311		<0.001	
Glucagon Stimulation Test								
	2.90 (0.60-7.60)	7.00 (3.20-11.85)	7.10 (5.10-10.20)	10.90 (5.20-20.40)	4.30 (0.80-8.10)	5.80 (1.07-13.65)	5.74 (2.05-9.15)	0.51 (0.22-0.97)
P	0.018		0.026		0.145		<0.001	

Supplementary Table 2: Proportion of patients attaining their peak GH levels on CST at various time points stratified by BMI, pubertal status, gender, and etiology

Time Points	BMI SDS		Pubertal Status		Gender		Aetiology	
	>-1.0SDS	<-1.0SDS	Prepubertal	Peripubertal	Male	Female	IGHD	MPHD
Clonidine Stimulation Test								
0 min	10.3% (28)	10.6%(25)	6.9%(12)	13.6%(24)	8.4%(29)	15.2%(27)	11.5% (31)	14.7%(11)
30 min	9.6% (26)	11.4%(27)	14.4%(25)	9.1%(16)	11.9%(41)	7.9%(14)	11.2%(30)	14.7%(11)
60 min	39.9% (108)	39.0%(92)	35.6%(62)	39.2%(69)	39.1%(135)	39.9%(71)	34.9%(94)	28.0%(21)
90 min	27.3% (74)	26.3%(62)	28.7%(50)	26.1%(46)	27.0%(93)	25.8%(46)	29.4%(79)	25.3%(19)
120 min	12.9% (35)	12.7%(30)	14.4%(25)	11.9%(21)	13.6%(47)	11.2%(20)	13.0%(39)	17.3%(13)
P	0.974		0.144		0.116		0.550	
Glucagon Stimulation Test								
0 min	3.4%(1)	6.9%(2)	5.3%(1)	3.7%(1)	7.0%(3)	0	5.9%(3)	0
60 min	10.3%(3)	3.4%(1)	10.5%(2)	7.4%(2)	7.0%(3)	5.6%(1)	5.9%(3)	0
90 min	17.2%(5)	31.0%(9)	36.8%(7)	29.6%(8)	30.2%(13)	16.7%(3)	27.5%(14)	1 (100%)
120 min	44.8%(13)	34.5%(10)	31.6%(6)	37.0%(10)	37.2%(16)	44.4%(8)	41.2%(21)	0
150 min	13.8%(4)	24.1%(7)	15.8%(3)	14.8%(4)	16.3%(7)	22.2%(4)	13.7%(5)	0
180 min	10.3%(3)	0%(0)	0	7.4%(2)	2.3%(1)	11.1%(2)	5.9%(3)	0
P	0.479		0.866		0.467		0.774	