GSC Advanced Research and Reviews

eISSN: 2582-4597 CODEN (USA): GARRC2

Cross Ref DOI: 10.30574/gscarr

Journal homepage: https://gsconlinepress.com/journals/gscarr/

(REVIEW ARTICLE)

Check for updates

Growth Hormone Secretion Patterns and Diagnostic Evaluation of Pediatric Growth Hormone Deficiency: A Comprehensive 25-Year Narrative Review (2000–2025)

Ashraf T. Soliman [1, *], Fawzia Alyafei [1], Nada Alaaraj [1], Noor Hamed [1], Shayma Ahmed [1], Dina Fawzy [2], Ahmed Elawwa [2] and Ahmed Khalil [3]

[1] *Department of Pediatrics, Hamad General Hospital, Hamad Medical Corporation, Doha, Qatar.*
[2] *Department of Pediatrics, Alexandria University, Alexandria, Egypt.*
[3] *Department of Clinical Pharmacy, Hamad General Hospital, Hamad Medical Corporation, Doha, Qatar.*

GSC Advanced Research and Reviews, 2025, 25(03), 048-067

Publication history: Received on 29 October 2025; revised on 03 December 2025; accepted on 06 December 2025

Article DOI: https://doi.org/10.30574/gscarr.2025.25.3.0373

Abstract

Background: Growth hormone (GH) plays a central role in childhood growth and metabolic regulation, yet its diagnostic evaluation remains challenging because of its pulsatile secretion and sensitivity to factors such as puberty, sleep, nutrition, and adiposity. Over the past 25 years, advances in assay technology and a deeper understanding of GH–IGF physiology have highlighted major limitations of traditional stimulation tests and IGF-1–based screening, underscoring the need for an updated, integrated diagnostic framework for pediatric GH deficiency (GHD).

Objectives: To synthesize current evidence (2000–2025) on GH secretion physiology, diagnostic performance of GH stimulation tests, IGF-1, and IGFBP-3, and the modifying effects of puberty, BMI, and assay generation; and to propose a modern stepwise diagnostic strategy for accurate early identification of GHD.

Methods: A narrative literature review was performed using PubMed, Scopus, and Google Scholar (2000–2025), including pediatric studies assessing GH secretion, GH stimulation tests, IGF-1/IGFBP-3 performance, and pubertal or BMI influences. Mechanistic and classical physiologic studies were included for context. Findings were synthesized into thematic domains and structured comparative tables.

Results: Across the literature, GH secretion increased two- to three-fold from Tanner I to Tanner III–IV, driven mainly by increased pulse amplitude rather than frequency. Parallel rises in IGF-1 and IGFBP-3 peaked in mid-puberty and declined in late adolescence, creating clearly stage-dependent physiological reference points. Failure of GH or IGF-1 to rise appropriately during mid-puberty strongly differentiated true GHD from constitutional or obesity-related patterns.

GH stimulation tests showed moderate-to-good sensitivity but only modest specificity, with substantial variability across stimuli and assays. Newer chemiluminescent assays produced lower GH peaks, indicating that historical 10 ng/mL cutoffs are no longer valid for many laboratories. Diagnostic accuracy improved when two different stimulation tests were used and when GH values were interpreted alongside Tanner stage and BMI rather than as isolated results.

IGF-1 and IGFBP-3 alone lacked sufficient sensitivity but demonstrated good specificity; combining both biomarkers or using IGF-1/IGFBP-3 ratios improved classification. Obesity consistently suppressed spontaneous GH pulsatility and reduced stimulated GH peaks, while also lowering IGF-1 specificity, confirming the need for BMI-adjusted cutoffs. Tanner III–IV represented the most physiologically informative stage for distinguishing delayed puberty from true GHD. Nocturnal GH sampling and IGF-1 generation tests had limited routine value but remained useful in selecting diagnostic dilemmas.

* Corresponding author: Ashraf T. Soliman

Conclusion: Accurate diagnosis of pediatric GHD requires an integrated, context-dependent approach that combines auxologic parameters, pubertal staging, BMI, IGF-1/IGFBP-3 interpretation, assay-specific GH cutoffs, and pituitary MRI. Such a strategy reduces misclassification and supports timely, appropriate therapy for children with true GHD.

Keywords: Growth Hormone Deficiency; GH Stimulation Tests; IGF-1/IGFBP-3; Pubertal Physiology; Pediatric Endocrinology

1. Introduction

Growth hormone (GH) plays a central role in linear growth, skeletal maturation, metabolic regulation, and body composition through its direct actions and through stimulation of hepatic insulin-like growth factor-1 (IGF-1), making the GH–IGF-1 axis a fundamental determinant of normal childhood growth (1,2).

GH secretion is pulsatile, ultradian, and influenced by circadian rhythm, sleep architecture, and hypothalamic regulation by GHRH, somatostatin, and ghrelin. This complex physiology renders random GH measurements clinically meaningless and underpins the need for dynamic or surrogate assessments in suspected GH deficiency (GHD) (3,4).

IGF-1 and its major binding protein IGFBP-3 are widely used as integrated biomarkers of GH action, reflecting overall GH exposure over days to weeks. However, their diagnostic accuracy is limited by variability due to age, puberty, undernutrition, chronic illness, and obesity, reducing sensitivity in many clinical scenarios (5,6).

Since the early 2000s, GH stimulation (provocation) tests—using clonidine, glucagon, arginine, or insulin-induced hypoglycemia—have remained the cornerstone of GHD diagnosis. Their diagnostic performance, however, varies markedly with assay methodology, peak GH cut-points, stimulus type, sex-steroid priming, and physiologic modulators such as BMI and puberty (7,8).

Diagnostic cut-points for GH stimulation tests have been inconsistent across decades, ranging from 5 to 10 ng/mL depending on assay standardization and clinical guidelines. More recent chemiluminescent and mass-spectrometry–aligned assays have produced lower GH values, necessitating recalibration of diagnostic thresholds and creating discrepancies between historical and modern data (9).

Obesity is now recognized as one of the most important confounders in GH axis evaluation. Excess adiposity suppresses spontaneous and stimulated GH secretion through increased somatostatinergic tone, hyperinsulinemia, elevated free fatty acids, and altered leptin signalling. As a result, obese children frequently show blunted GH peaks and reduced IGF-1, increasing the risk of false-negative diagnoses (10,11).

Sex and pubertal maturation strongly influence GH dynamics. Rising estradiol concentrations during puberty—regardless of biological sex—amplify GH pulse amplitude and stimulate IGF-1 production. Boys typically exhibit higher IGF-1 peaks in mid-puberty, whereas girls may have slightly lower stimulated GH peaks at comparable maturational stages, affecting test interpretation (12).

Other proposed diagnostic modalities, including nocturnal GH sampling and IGFBP-3 assessment, have shown limited incremental value due to low specificity, wide inter-individual variability, and strong modulation by non-GH factors such as nutrition, sleep disturbance, systemic inflammation, and comorbidities (13).

Imaging of the hypothalamic–pituitary region using MRI remains essential for identifying congenital or acquired structural abnormalities, such as pituitary hypoplasia, ectopic posterior pituitary, stalk defects, tumors, or post-surgical changes. MRI findings not only support the diagnosis of organic GHD but also guide prognosis and long-term management (14).

Despite 25 years of scientific progress, no universal gold standard for diagnosing pediatric GHD has emerged. Heterogeneity in GH assays, variable cut-points, lack of BMI-adjusted norms, inconsistent use of sex-steroid priming, and physiologic influences of obesity and puberty continue to challenge diagnostic accuracy. A comprehensive updated review is therefore urgently needed to synthesize current evidence and improve early diagnosis and monitoring of short children—with or without true GH deficiency—and to prevent both over- and under-diagnosis (15).

Objectives

- To synthesize and critically evaluate 25 years of evidence (2000–2025) on GH secretion physiology and the diagnostic performance of available tests for pediatric growth hormone deficiency.
- To assess how key modifiers—BMI/obesity, sex, age, and pubertal status—affect GH stimulation responses, IGF-1/IGFBP-3 levels, and diagnostic accuracy.
- To develop an updated, evidence-based diagnostic framework that improves early identification, risk stratification, and monitoring of short children with and without true GH deficiency.

2. Materials and methods

This review was conducted as a comprehensive narrative synthesis of literature examining growth hormone (GH) secretion patterns and diagnostic approaches to pediatric growth hormone deficiency (GHD) over the past 25 years. A structured search strategy was applied to ensure broad and representative coverage of research published between January 2000 and January 2025. PubMed, Scopus, and Google Scholar were used as primary databases, and search terms were combined using Boolean operators to maximize sensitivity and relevance. The core search concepts included "growth hormone deficiency," "pediatric" or "children," "GH stimulation test," "IGF-1," "IGFBP-3," "diagnosis," "sensitivity," "specificity," "puberty," and "obesity." Reference lists of the retrieved articles, clinical guidelines, and systematic reviews were also screened manually to identify any additional studies that were not captured by database searches but were relevant to GH physiology or diagnostic practices.

Studies were considered eligible for inclusion if they evaluated pediatric subjects (aged 0–18 years) undergoing diagnostic assessment for suspected GHD and if they provided data on GH secretion patterns, GH stimulation tests, GH assay characteristics, IGF-1 or IGFBP-3 performance, or related diagnostic measures. Eligible study designs included observational cohorts, cross-sectional diagnostic studies, randomized trials, and systematic reviews or meta-analyses. Only peer-reviewed articles published in English, with full-text availability, were included. Although the primary time window was 2000–2025, certain landmark studies predating this period were incorporated when necessary to provide essential physiological context.

Exclusion criteria were applied to enhance the methodological robustness of the review. Studies were excluded if they reported insufficient diagnostic data, focused exclusively on adult GHD, were non–peer reviewed, or were available only as conference abstracts without full datasets. Case reports and small case series with fewer than 10 participants were also omitted, as studies centered solely on syndromic short stature unrelated to GH secretory physiology, unless they explicitly evaluated GHD. Animals, in vitro, and purely mechanistic molecular studies were excluded to maintain clinical relevance.

Data extraction was performed systematically for all included studies. The extracted variables included publication year, study design, patient demographics, sample size, BMI or obesity classification, pubertal staging methods, GH stimulation protocols, laboratory assays used, peak GH cut-points, IGF-1 and IGFBP-3 z-score thresholds and reported diagnostic accuracy indices such as sensitivity and specificity. When available, ancillary findings such as pituitary MRI abnormalities were also recorded. All extracted information was cross-checked for accuracy and consistency.

Quality assessment of the selected studies was undertaken using established evaluation tools appropriate to each study type. Diagnostic accuracy studies were appraised using the QUADAS-2 instrument, with particular attention to patient selection, reference standards, and test interpretation. Observational cohort or case–control studies were evaluated using the Newcastle–Ottawa Scale, focusing on representativeness, comparability, and outcome assessment. Systematic reviews identified during the search were assessed using the AMSTAR-2 framework. Only studies judged to have moderate or high methodological quality in the key domains were included in the final synthesis.

Because of the anticipated heterogeneity across studies—stemming from differences in GH assays, stimulation test protocols, cut-offs, age distribution, BMI stratification, and IGF-1 measurement techniques—formal meta-analysis was not attempted. Instead, findings were narratively synthesized and organized into thematic domains covering GH physiology, GH stimulation testing, IGF-1 and IGFBP-3 diagnostic performance, the influence of obesity and sex, and the role of MRI in diagnostic confirmation. Four comprehensive summary tables were prepared to consolidate the key findings from the included literature.

As this investigation synthesizes published data without involving human subjects, institutional ethics approval was not required.

3. Results

The results of this narrative review synthesize 25 years of evidence on GH secretion physiology, diagnostic test performance, and the modifying effects of puberty, BMI, and assay evolution, integrating findings from multiple comparative tables to provide a unified interpretation of the GH–IGF axis in pediatric evaluation.

Table 1 GH pulsatile secretion across childhood and puberty: changes in amplitude, frequency, and 24-hour GH output

Study (Year)	Population & Tanner Stages	Sampling Method	Findings on GH Pulsatility	Influence of Sex, BMI, Puberty	Ref
Miller, 1982	38 healthy children (4–16 y), both sexes; Tanner I–V	GH every 20 min × 8 h; pulse analysis	GH pulsatility present prepubertally; pulse amplitude increases markedly with advancing puberty; frequency relatively stable	Pubertal rise driven mainly by pulse amplitude, not frequency	(16)
Martha, 1992	33 normal boys; prepubertal → late puberty	24-h frequent sampling; deconvolution	2–3 fold increase in 24-h GH production from prepubertal to late puberty; GH half-life unchanged	GH output inversely related to BMI, even in normal boys	(17)
Mauras, 1987	12 boys, early–late puberty	24-h GH sampling; deconvolution	Puberty induces marked GH burst mass increase with minimal change in frequency; correlates strongly with rising estradiol	Pulse amplitude is sex-steroid driven	(18)
Van Cauter, 2000	Children → adolescents → adults	24-h GH profiles with sleep staging	Early-night slow-wave sleep (SWS) drives major GH pulses; GH output declines with age parallel to SWS decline	Pubertal GH rise tightly linked to sleep architecture	(20)
Roemmich, 2002	Boys & girls; pre-, mid-, late puberty	Nocturnal GH sampling + body composition	Mean nocturnal GH increases from pre- to mid-puberty; remains high in late puberty; tracks with IGF-1	Pubertal GH rise modulated by sex steroids and body composition	(22)
Roemmich, 2005	57 lean vs overweight youth	12-h nocturnal sampling + deconvolution	Overweight youth show marked reduction in GH burst mass, integrated GH, and GH half-life; pulse frequency preserved	Adiposity blunts GH pulsatility, masking expected pubertal rise	(21)
Veldhuis, 2008	Review of pediatric & adult datasets	Deconvolution synthesis	Puberty increases GH burst mass > frequency; obesity suppresses amplitude	GHRH drives amplitude; somatostatin drives frequency	(23)
Kasa-Vubu, 2006	38 postpubertal girls (lean vs overweight)	24-h GH + leptin sampling	Overweight girls have lower GH pulsatility despite similar pubertal stage; GH inversely correlated with leptin	Adiposity + leptin physiology modulate GH after puberty	(24)

Denny-Brown, 2012	108 adolescents/adults; subset peri-pubertal	GH stimulation + overnight sampling	Vitamin C and nutrient intake correlate positively with spontaneous GH burst mass and total GH output	Nutritional quality modulates GH secretion beyond puberty/BMI	(25)
Calvert, 2022	20 pubertal children (Tanner II–IV)	Overnight sampling with sleep disruption	Acute sleep disruption reduces nocturnal GH peaks but daytime compensation preserves 24-h GH	Timing of pulses changes more than total output	(26)

Table 1 : Across classical and recent studies, GH pulsatile secretion is present throughout childhood but increases substantially during puberty due to a quantitative rise in burst mass and amplitude, rather than major changes in frequency (16–18,20,22,23). Slow-wave sleep remains the dominant driver of large nocturnal pulses (20,26). Adiposity strongly suppresses GH pulsatility—including burst mass, integrated GH output, and GH half-life—across both sexes and all pubertal stages (17,21,24). Newer research shows additional modulators, such as micronutrient intake and acute sleep disruption, influencing GH output (25,26). These combined findings highlight that interpretation of GH testing must account for pubertal stage, BMI, sex steroid milieu, sleep context, and nutritional profile to avoid diagnostic misclassification during adolescence.

Table 2 Diagnostic performance of GH provocation tests for pediatric GHD (2000–2025)

Study (Year)	**Population & Setting**	**Test(s) & GH Cut-off**	**Reported Diagnostic Performance**	**Key Notes / Limitations**	**Ref**
Maghnie, 2005	150 children with short stature; Italian multicenter GHD workup	ITT, arginine, clonidine; GH cut-off 10 ng/mL	Overall sensitivity of single test ~80–90%; specificity ~60–75%; reproducibility between tests modest (27)	Concluded that **no single stimulation test is fully reliable**; discordant results common; recommended combining auxology, IGF-1 and ≥1 test rather than relying on a single GH cut-off (27).	(27)
Zimmermann, 2011	209 short children; GH stimulation vs MRI & auxology	ITT vs glucagon (GTT); GH cut-off 7 vs 10 ng/mL	Lower cut-off (7 ng/mL) improved specificity but reduced sensitivity; higher cut-off (10 ng/mL) increased sensitivity but more false positives; ITT and GTT showed similar AUCs (28)	Highlighted that **cut-off choice shifts sensitivity–specificity balance**; suggested assay-specific, age-adjusted thresholds rather than universal 10 ng/mL (28).	(28)
Ibba, 2014	94 children evaluated for GHD; two stimulation tests performed	Clonidine + arginine; GH cut-off 8 ng/mL	When **both tests were concordant**, sensitivity ~90% and specificity ~92%; single-test performance was significantly lower (29)	Showed that **using two different stimuli** improves diagnostic precision and reduces misclassification versus one test alone (29).	(29)
Elbornsson, 2015	120 individuals with childhood-onset GHD retested in transition	ITT vs GHRH–arginine; adult cut-offs applied	ITT sensitivity ~90% for persistent GHD; GHRH–arginine high sensitivity but lower specificity in obese subjects (30)	Demonstrated that **obesity reduces specificity** of GHRH–arginine; underlines BMI-adjusted interpretation of provocation tests (30).	(30)
Felício, 2019	116 children with suspected	Various stimuli (clonidine,	AUCs for GH stimulation tests ranged 0.78–0.87;	Emphasized impact of **assay generation** on GH	(31)

	GHD; Italian cohort	arginine, ITT); ROC analysis by assay	optimal cut-offs often **below 10 ng/mL** with modern assays; combining tests and IGF-1 improved discrimination (31)	peaks and advocated **redefining cut-offs** according to the calibrated assay used (31).	
Fatani, 2023	165 children undergoing stimulation; multicenter	Two GH stimulation tests (mostly clonidine, glucagon); IGF-1 integrated	Concordant double-fail on two tests strongly predicted GHD; normal IGF-1 with borderline GH peaks reduced post-test probability; IGF-1 ≥ 0 SDS allowed **omission of second test** in some models (32)	Suggested that **IGF-1 combined with a single failed test** may obviate a second test in selected children, improving practicality while maintaining acceptable sensitivity (32).	(32)

Table 2: Across pediatric cohorts, GH stimulation tests show moderate-to-good sensitivity but only modest specificity when used in isolation, with performance strongly dependent on the test type, assay generation, and GH cut-off chosen (27–31). Insulin-induced hypoglycemia, clonidine, arginine, and glucagon all provide comparable areas under the ROC curve in many series, but no individual stimulus emerges as a perfect gold standard, and reproducibility between tests in the same child is limited (27–29,31). Modern chemiluminescent and standardized assays tend to yield lower GH peak values, meaning that historical cut-offs of 10 ng/mL often overcall deficiency and more appropriate thresholds may lie in the 6–8 ng/mL range for some methods (28,31). Recent work supports combined strategies, in which concordant failure of two different tests or integration with low IGF-1 substantially improves diagnostic confidence, while normal IGF-1 with borderline peaks may allow more conservative approaches or omission of a second test in selected children (29,32). These data collectively reinforce the concept that GH provocation tests should be interpreted within a multimodal framework rather than used as binary stand-alone criteria.

Table 3 Diagnostic performance of IGF-1 and IGFBP-3 as screening markers in pediatric growth hormone deficiency (GHD)

Study (ref)	**Marker / cut-off definition**	**Sensitivity**	**Specificity**	**Key diagnostic message**
Shen et al., 2015 (systematic review & meta-analysis) (33)	Pooled analysis of serum IGF-1 and IGFBP-3 (various cut-offs, mostly <–1 to –2 SDS) across pediatric GHD cohorts	IGF-1: pooled sensitivity ≈ 60–70%; IGFBP-3: slightly lower	IGF-1: pooled specificity ≈ 65–75%; IGFBP-3: often higher (≈ 75–85%)	Meta-analysis showed moderate sensitivity and specificity for both IGF-1 and IGFBP-3 and confirmed that neither marker alone can replace GH stimulation tests, but low values increase post-test probability of GHD (33).
Inoue-Lima et al., 2020 (Bayesian approach) (34)	IGF-1-SDS and IGFBP-3-SDS in short children (GHD if 2 failed GH tests)	IGF-1-SDS: 92% sensitivity	IGF-1-SDS: 69% specificity; IGFBP-3-SDS: 45.8% sensitivity, 93.8% specificity	IGF-1-SDS is a high-sensitivity screening test, whereas IGFBP-3-SDS has high specificity but low sensitivity; combining auxology with IGF-1 for screening and using IGFBP-3 for confirmation improved post-test probabilities in a Bayesian model (34).
Haj-Ahmad et al., 2023 (IGF-1/IGFBP-3 ratio) (35)	IGF-1, IGFBP-3, and IGF-1/IGFBP-3 molar ratio in 235 short children; GHD defined as peak GH <6.25 ng/mL in 2 tests	Low IGF-1/IGFBP-3 ratio: sensitivity 87.5%	Low IGF-1/IGFBP-3 ratio: specificity 83.0%; combination of low IGF-1, IGFBP-3, and ratio: specificity 97.7%	The IGF-1/IGFBP-3 molar ratio outperformed either analyte alone, providing high sensitivity and good specificity; combining all three markers yielded near-rule-in specificity and normal values for all three effectively ruled out GHD (35).

Iwayama et al., 2021 (prospective cohort) (36)	IGF-1-SDS in 298 short/decelerating children; best ROC cut-off −1.493 SDS	68.5% sensitivity	41.7% specificity; AUC 0.517	IGF-1-SDS had poor overall diagnostic accuracy with low AUC and modest sensitivity/specificity, leading authors to conclude that IGF-1 alone should not be used as a stand-alone screening test for GHD (36).
Galluzzi et al., 2010 (short-stature clinic cohort) (37)	IGF-1 and IGFBP-3 (age- and sex-adjusted SDS) in children with short stature referred for GHD evaluation	IGF-1: low-to-moderate sensitivity (≈ 30–60% depending on cut-off)	IGF-1 and IGFBP-3: moderate-to-high specificity (≈ 70–90%)	In this real-world cohort, both markers had limited sensitivity but reasonable specificity; neither test alone could replace GH stimulation, but low IGF-1/IGFBP-3 increased suspicion of GHD and normal values reduced it (37).
Bereket et al., 2009 (narrative evidence synthesis) (38)	Review of IGF-1 and IGFBP-3 studies and cut-offs in prepubertal/early-pubertal short children	Reported sensitivities for IGF-1 generally in the 50–70% range; IGFBP-3 often lower	Specificities frequently ≥70–80% for low IGF-1 or IGFBP-3; higher specificity at more stringent (≤−2 SDS) cut-offs	Summarized data indicate that low IGF-1 and/or IGFBP-3 are supportive but not diagnostic of GHD; best use is as screening/triage tests before GH stimulation, interpreted with auxology, bone age, and comorbidities (38).
Kota et al., 2012 (South-Asian cohort) (39)	IGF-1 and IGFBP-3 levels in children with pituitary dwarfism vs non-GHD short stature	Sensitivity of IGF-1 and IGFBP-3 individually was modest; higher when both low	Specificity improved when both markers were low	Authors concluded that IGF-1 and IGFBP-3 are useful as initial screening tools, but must always be followed by GH provocative testing for definitive diagnosis; combined low values substantially increase likelihood of GHD (39).
Je et al., 2014 (Korean short-stature cohort) (40)	IGF-1 and IGFBP-3 levels in 207 short children undergoing GH testing	Sensitivity of either marker alone was limited; negative predictive value high when both normal	Specificity increased when stricter SDS cut-offs applied and when both markers low	This lab-based study showed that normal IGF-1 and IGFBP-3 together make GHD unlikely, whereas low levels of both strongly suggest GHD; however, due to imperfect sensitivity, GH stimulation tests remained indispensable (40).

Table 3: Throughout studies, IGF-1 and IGFBP-3 function as supportive, not definitive, tools for diagnosing pediatric GHD. Meta-analytic data show only moderate pooled sensitivity and specificity, with IGF-1 generally more sensitive and IGFBP-3 more specific, especially at stringent SDS cut-offs (33,38). Single-center cohorts and Bayesian modeling indicate that IGF-1-SDS is useful for high-sensitivity screening, while IGFBP-3-SDS contributes high specificity, particularly in younger children, when used as a confirmatory marker (34,37). Newer work suggests that composite indices such as the IGF-1/IGFBP-3 molar ratio and multi-marker combinations can substantially improve diagnostic discrimination, achieving sensitivities around 85–90% and specificities over 80–95% in selected populations (35). At the same time, prospective data with broad inclusion criteria demonstrate that IGF-1 alone may perform poorly when pre-test probability of GHD is modest, with low AUC and limited specificity (36). Overall, the evidence supports using IGF-1 and IGFBP-3 as integrated components of a diagnostic algorithm—interpreted alongside auxology, puberty, BMI, comorbidities, and GH stimulation tests—rather than as stand-alone replacements for provocative testing in short children being evaluated for GHD.

Table 4 Influence of obesity / BMI on GH and IGF-1 test performance in children and adolescents

Study (ref)	Population / BMI profile	Test or marker evaluated	Effect of BMI / obesity on GH or IGF-1	Diagnostic impact
Cornier et al. (10)	Children, adolescents, and adults with varying adiposity and features of metabolic syndrome	GH physiology and metabolic markers (narrative synthesis)	Obesity associated with suppressed spontaneous and stimulated GH secretion, hyperinsulinemia, and increased free fatty acids; IGF-1 often normal or low-normal despite reduced GH pulses	Mechanistic basis for GH axis suppression in obesity, implying that short obese children may show blunted GH responses and misleadingly low-normal IGF-1, increasing risk of false-negative GHD diagnoses
Bizzarri et al. (11)	Children undergoing GH stimulation tests, stratified by BMI SDS	Standard GH stimulation tests (various stimuli; peak GH vs BMI SDS)	Peak GH inversely correlated with BMI SDS; overweight/obese children showed significantly lower GH peaks than lean peers under the same stimulus	Demonstrated that BMI is a major determinant of GH test peak, so using a fixed cut-off across BMI groups can over-diagnose GHD in lean and under-diagnose in obese children
Martha et al. (17)	Normal boys across puberty with a range of BMIs	24-h GH secretion with deconvolution analysis	Higher BMI associated with lower 24-h GH production and reduced secretory burst mass, despite otherwise normal growth and puberty	Indicates that adiposity alone can mimic features of GHD physiology, suggesting that reduced GH peaks in heavier children do not automatically imply true deficiency
Roemmich et al. (21)	57 lean vs overweight youth at different pubertal stages	Nocturnal GH profiles (pulsatility; integrated GH)	Overweight youth had markedly reduced GH burst mass, integrated GH, and GH half-life, with preserved pulse frequency, across all pubertal stages	Shows that obesity blunts expected pubertal increase in GH pulsatility, so obese adolescents may fail to reach "normal pubertal" GH peaks during provocation testing
Kasa-Vubu et al. (24)	Postpubertal adolescent girls: lean vs overweight	24-h GH and leptin secretion	Overweight girls showed significantly lower GH pulsatility, inversely related to leptin and adiposity, despite similar pubertal status	Suggests that in postpubertal girls, leptin-mediated adiposity effects further suppress GH, complicating interpretation of both GH tests and IGF-1 levels in heavier adolescents
Elbornsson et al. (30)	Individuals with childhood-onset GHD retested in transition; BMI-stratified	ITT and GHRH–arginine stimulation	In overweight/obese subjects, GHRH–arginine showed reduced specificity, with more false-positive "GHD" results at standard cut-offs; ITT less affected but still influenced by BMI	Indicates that obesity reduces specificity of certain stimuli, particularly GHRH–arginine, reinforcing the need for BMI-adjusted GH cut-offs and

				multimodal confirmation
Felício et al. (31)	116 children with suspected GHD; modern GH assays, BMI recorded	Multiple GH stimuli + IGF-1; ROC analysis	Study noted that BMI negatively correlated with GH peak; optimal GH cut-offs with modern assays tended to be lower than historical 10 ng/mL, especially in heavier children	Supports the concept that fixed historical cut-offs are inappropriate in the modern era and that BMI and assay type must be considered when defining GH "normality"
Fatani et al. (32)	165 children undergoing two GH stimulation tests; BMI considered in models	Two GH tests plus IGF-1 SDS	Obesity was associated with lower GH peaks but not uniformly low IGF-1, leading to discordant profiles (borderline GH but non-low IGF-1)	Suggested that a normal or near-normal IGF-1 in obese short children with modestly reduced GH peaks lowers the probability of true GHD, and that integrated algorithms can reduce unnecessary second tests
Shen et al. (33)	Meta-analysis of pediatric IGF-1/IGFBP-3 diagnostic performance; heterogeneous BMI	IGF-1 and IGFBP-3 diagnostic indices across cohorts	Considerable heterogeneity between studies partly explained by differences in BMI and nutritional status; cohorts with more obesity tended to show lower IGF-1 for non-GHD controls, reducing specificity	Indicates that obesity can lower IGF-1 in non-GHD children, thereby reducing the specificity of low IGF-1 as a screening marker for GHD in modern populations
Iwayama et al. (36)	298 short/decelerating children with mixed BMI; IGF-1 vs GH tests	IGF-1 SDS as a diagnostic indicator	IGF-1 SDS showed poor AUC (≈0.52) with modest sensitivity and low specificity; authors noted that comorbidities and BMI distribution likely contributed to poor discrimination	Demonstrates that in real-world mixed-BMI cohorts, IGF-1 alone performs poorly as a discriminator of GHD, and its interpretation must consider BMI, comorbidities, and pretest probability

The studies summarized in Table 4 show consistently that higher BMI and obesity blunt GH secretion and distort the diagnostic performance of both GH stimulation tests and IGF-1–based screening. In children and adolescents, adiposity reduces GH secretory burst mass and peak responses, so that short obese patients may fail to reach conventional GH cut-offs despite not having true structural or genetic GHD (11,17,21,24,30). At the same time, obesity and associated metabolic changes can also lower IGF-1 or keep it only mildly reduced, which decreases the specificity of low IGF-1 for GHD in contemporary, more overweight pediatric populations (10,32,33,36). Together, these data indicate that diagnostic algorithms that ignore BMI are prone to both false-negative (true GHD masked by obesity-related suppression) and false-positive (obesity misclassified as GHD when using rigid cut-offs or certain stimuli) errors. Therefore, BMI-adjusted interpretation of GH peaks, cautious use of IGF-1, and reliance on integrated clinical, auxologic, and imaging data are essential when evaluating short children in the context of the global obesity epidemic.

Table 5 Comparative advantages and limitations of key diagnostic methods for evaluating the GH–IGF-1 axis in short children

Method	Main information provided	Advantages in clinical practice	Limitations / pitfalls	Recommended clinical role	Ref
GH stimulation tests (clonidine, glucagon, ITT, arginine, etc.)	Pharmacologic maximal GH secretory capacity; helps define biochemical GHD vs adequate reserve (7,27–32)	Long-standing standard; directly tests pituitary GH release; high sensitivity when two different stimuli are used; widely available; supported by guidelines (7,27,29,31)	Limited reproducibility; influenced by assay type, BMI, puberty, sex-steroid priming, and comorbidities; risk of hypoglycemia with ITT; specificity only moderate; historical 10 ng/mL cut-off often too high with modern assays (8,11,27–32)	Remains the core confirmatory test for GHD, ideally using two different stimuli plus integrated auxology and IGF-1; cut-offs should be assay- and BMI-aware rather than fixed (7,27–32)	(7,8,11,27–32)
IGF-1 (age/sex/BMI-adjusted SDS)	Integrated marker of chronic GH exposure and hepatic GH action; correlates with growth velocity (5,6,33–38)	Non-invasive, single blood sample; relatively stable across the day; good specificity at low SDS cut-offs; helpful to estimate pre-test probability and prioritize testing (33,34,38)	Sensitivity only moderate; affected by age, puberty, BMI, undernutrition, chronic disease, liver function and assay variability; normal IGF-1 does not exclude GHD; performance can be poor in unselected mixed-BMI cohorts (5,6,33,36,38)	Best used as a screening and triage marker in combination with auxology and clinical context to decide who needs GH stimulation; very low IGF-1 increases suspicion, normal IGF-1 lowers it but does not rule out GHD (33–38)	(5,6,33–38)
IGFBP-3 (SDS)	Marker of GH-dependent binding protein production; less age-dependent than IGF-1 in younger children (33,34,37,38)	Often more specific than IGF-1 in younger/prepubertal children; less influenced by short-term nutritional variation; useful when IGF-1 is equivocal (34,37,38)	Sensitivity frequently lower than IGF-1; still influenced by nutrition, chronic illness, and liver function; diagnostic performance alone insufficient to replace GH testing (33,38)	Acts as a complementary marker to IGF-1; low IGF-1 *and* low IGFBP-3 together substantially increase post-test probability of GHD, especially in younger children (34,37–39)	(33,34,37–39)
IGF-1/IGFBP-3 molar ratio /	Composite index reflecting	In several cohorts, the ratio and combined	Assay standardization	Promising second-line	(35,38)

combined markers	bioavailable IGF-1 and GH sensitivity (35,38)	low IGF-1 + IGFBP-3 achieve higher sensitivity and specificity than either marker alone; near "rule-in" specificity when all are low (35)	for both analytes required; reference ranges less widely established; data still limited to specialized centers and research cohorts (35,38)	laboratory tool in expert centers to refine pre-test probability before or after GH stimulation, particularly in borderline cases (35,38)	
IGF-1 generation test (short course of GH, measure pre–post IGF-1 ± IGFBP-3)	Dynamic test of peripheral GH sensitivity and capacity to generate IGF-1 in response to exogenous GH (41,42)	Useful in differentiating GH insensitivity vs GHD vs secondary causes of low IGF-1; can help predict growth response to GH therapy in some cohorts; helpful in selected syndromic or severe IGF-1 deficiency cases (41,42)	Time-consuming and resource-intensive; protocols not standardized; sensitivity and specificity variable; limited normative data; poor performance for milder GH insensitivity; not ideal as routine GHD screening; recent reviews emphasize significant limitations (41,42)	Reserved for specialized centers and selected patients where GH insensitivity, severe primary IGF-1 deficiency, or atypical response to GH is suspected; adjunct, not replacement, for GH stimulation and clinical assessment (41,42)	(41,42)
Nocturnal spontaneous GH secretion assessment (frequent overnight sampling)	Detailed physiologic profile of GH pulsatility, amplitude, and coupling to sleep stages (13,16–18,20,22,23)	Provides rich information on true physiologic GH secretion; useful in research and to understand complex or atypical cases; can help differentiate neurosecretory dysfunction from classic GHD in specialized settings (13,16–18,20,23)	Labor-intensive, costly, and impractical for routine use; uncomfortable for children; significant intra-individual variability; low specificity for GHD vs normal variants; interpretation requires advanced deconvolution analysis (13,20,23)	Primarily a research or tertiary-center tool; rarely indicated in clinical practice except in highly selected cases where conventional tests are inconclusive and a detailed physiological profile is needed (13,20,23)	(13,16–18,20,22,23)

Table 5: Shows that each diagnostic method captures a different dimension of the GH–IGF-1 axis, explaining why no single test can reliably diagnose GHD on its own. GH stimulation tests remain the essential confirmatory tool but are limited by assay variability, BMI and pubertal effects, and only moderate specificity. IGF-1 and IGFBP-3 offer convenient screening markers with good specificity, though insufficient sensitivity to exclude GHD. Combined indices such as the IGF-1/IGFBP-3 ratio enhance discrimination in borderline cases but require standardized assays. The IGF-1 generation test and nocturnal GH profiling provide valuable physiological insight but are too resource-intensive for routine use. Overall, the table emphasizes the need for an integrated diagnostic approach that uses biomarkers to guide testing and relies on stimulation tests, interpreted in clinical context, for confirmation.

Table 6 Changes in IGF-1, IGFBP-3, and GH peak across Tanner stages in healthy children and adolescents

Tanner stage	IGF-1 pattern (age/sex-adjusted)	IGFBP-3 pattern	GH secretion / GH peak pattern (spontaneous or stimulated)	Clinical interpretation	Ref
I (prepubertal)	IGF-1 is low–normal, gradually increasing with age; levels are substantially below pubertal peak, typically around –1 to 0 SDS in healthy children (43,44,46,47).	IGFBP-3 is also low–moderate, increasing steadily with age but with less steep change than IGF-1; values are clearly lower than in pubertal stages (43,44,45,46).	GH is already secreted in pulsatile fashion with stable frequency but relatively small pulse amplitude and lower 24-h output; stimulated GH peaks are usually adequate with standard cutoffs (16–18).	Auxology (height SDS, growth velocity) is paramount; both IGF-1 and IGFBP-3 can be in the "normal" range despite early GHD; low-normal levels must be interpreted with growth data and bone age.	(16–18,43–47)
II (early puberty)	IGF-1 begins a steep upward trajectory, rising into mid-normal or high-normal SDS; in many cohorts, early pubertal IGF-1 is already clearly higher than in Tanner I, especially in girls (43,44,21,47).	IGFBP-3 increases further but more modestly than IGF-1; the IGF-1/IGFBP-3 ratio starts to rise, reflecting increasing GH drive (43,45,47).	GH pulse amplitude increases and 24-h secretion rises, but the pubertal peak in GH output is not yet reached; GH responses to stimulation tests often begin to increase vs prepuberty, particularly in boys entering puberty (16–18,20,22,24).	Rising IGF-1 and IGFBP-3 help confirm true pubertal onset; lack of increase with clear pubertal signs may suggest GH–IGF-1 axis impairment. However, overlap with late Tanner I remains; cut-offs must remain cautious.	(16–18,20,21,22,24,43–47)
III (mid-puberty)	IGF-1 reaches or approaches peak values in many cohorts; Korean and Turkish reference data show highest median or near-highest IGF-1 around Tanner III–IV, with girls peaking 1–2 years earlier than boys (43,44,18,47,48).	IGFBP-3 is high, often near its pubertal maximum but rising less steeply than IGF-1; the IGF-1/IGFBP-3 ratio peaks around Tanner III–IV, indicating maximal bioactive IGF-1 (43,45,47,49).	GH pulsatility shows markedly increased pulse amplitude and 24-h secretory mass, with relatively stable frequency; GH peaks after stimulation are often highest in mid-puberty; spontaneous nocturnal peaks are tightly coupled to slow-wave sleep (16–18,20,22,24,48,49).	This stage is the physiologic "window of maximal GH–IGF-1 drive" and growth velocity; failure of IGF-1 and GH peaks to rise appropriately in a growing child at Tanner III strongly suggests true GHD or GH insensitivity rather than constitutional delay alone.	(16–18,20,22,24,43–45,47–49)

IV (late puberty, peak growth spurt)	Multiple reference models (Korean, Turkish, European) show highest or plateaued IGF-1 SDS at Tanner IV, followed by a decline thereafter; girls generally reach peak ~2 years earlier than boys (43,44,21,47,48).	IGFBP-3 peaks around Tanner IV–V, with some cohorts showing maximal concentrations at Tanner IV in boys and Tanner V in girls; IGF-1/IGFBP-3 ratio remains high (43,45,47,49).	GH pulse amplitude and integrated 24-h secretion remain very high, close to Tanner III levels; spontaneous nocturnal GH peaks and stimulated peaks are robust, though inter-individual variability increases (16–18,20,22,24,48,49).	Tanner III–IV represent the best physiological "stress test" of the GH–IGF-1 axis; markedly low IGF-1 and/or subnormal GH peaks here are highly suspicious, provided BMI, nutrition, and chronic disease are accounted for.	(16–18,20,22,24,43–45,47–49)
V (late/postpubertal)	IGF-1 begins to decline from its pubertal peak, though still higher than prepubertal values; there is a gradual fall toward young adult levels with advancing age (21,43,44,47).	IGFBP-3 also declines slightly from its peak; in some cohorts, IGFBP-3 continues to be relatively high into Tanner V with a slower fall than IGF-1, leading to a modestly reduced IGF-1/IGFBP-3 ratio (43–45,47,49).	GH pulse amplitude and 24-h output decrease from pubertal maxima, and GH peaks on stimulation tests may be lower than at Tanner III–IV despite normal adult GH status; sleep-associated peaks diminish as slow-wave sleep declines (20,21,22,24,48).	Falling IGF-1 and GH output in Tanner V reflect physiologic maturation, not pathology. Using pubertal peak cut-offs for older adolescents can lead to overdiagnosis of GHD; interpretation should transition gradually toward adult reference ranges.	(20–22,24,21,43–45,47–49)

Table 6 illustrates the tightly coordinated rise and fall of IGF-1, IGFBP-3, and GH secretory patterns across the five Tanner stages, emphasizing why pubertal status must be central to interpreting biochemical tests for GHD. From Tanner I to IV, there is a progressive and physiologically predictable increase in IGF-1, IGFBP-3, GH pulse amplitude, and stimulated GH peaks, with the strongest GH–IGF-1 activation occurring during Tanner III–IV. These stages represent the most informative "physiologic stress test" of the axis, where failure of IGF-1 or GH to rise appropriately is most suggestive of true GHD rather than constitutional delay. By Tanner V, all markers decline toward adult levels, making low or modest GH peaks physiologic rather than pathological. Overall, the table reinforces that Tanner-stage–adjusted interpretation is essential for avoiding misdiagnosis and that pubertal stage profoundly modifies both biochemical expectations and diagnostic accuracy.

Table 7 Values and levels of GH Pulsatility, IGF-1 and IGFBP-3 Across Tanner Stages

Tanner stage	GH pulse frequency (pulses / 24 h)	GH pulse amplitude / 24-h secretion	Typical IGF-1 range (ng/mL)*	Typical IGFBP-3 range (ng/mL)*	Key notes	Ref
I – Prepubertal	~4–6	Low; small burst mass and lower 24-h GH output despite normal frequency	Girls: 49–342 Boys: 63–279	~3000–4000	Baseline prepubertal pattern; IGF-1 and IGFBP-3 low-normal. Auxology and growth velocity remain the most important discriminators of GHD vs normal variant.	(16–18,43–47)
II – Early puberty	~5–7	Increasing amplitude and secretory mass; frequency relatively stable	Girls: 115–428 Boys: 75–420	~3500–4500	Start of pubertal GH–IGF-1 rise; IGF-1 and IGFBP-3 begin to climb above prepubertal values, especially in girls entering puberty.	(16–18,20,21,43–47)
III – Mid-puberty	~5–7	Marked increase in pulse amplitude and 24-h GH secretion; near-maximal burst mass	Girls: 145–760 Boys: 94–765	~4000–5000	Period of maximal GH–IGF-1 drive and rapid growth. Failure of IGF-1 and GH to rise appropriately at this stage is highly suspicious for GHD (after accounting for BMI and comorbidities).	(16–18,20,22,24,43–45,47–49)
IV – Late puberty (peak growth spurt)	~5–7	High (~17 ng/mL bursts); 24-h GH secretion at or near peak	Girls: 224–586 Boys: 223–578	~4500–5200	Peak GH amplitude and secretion; IGF-1 and IGFBP-3 plateau at their highest levels before beginning to decline toward Tanner V / young adulthood.	(16–18,20,22,24,43–45,47–49)
V – Late/postpubertal	~4–6	Declining amplitude and 24-h GH output as slow-wave sleep and pubertal drive wane	Girls: 188–512 Boys: 227–518	~4000–4800	IGF-1 and IGFBP-3 fall from pubertal peaks toward adult ranges; lower GH peaks here are physiologic, so using Tanner III–IV cut-offs in Tanner V can overdiagnoses GHD.	(20–22,21,43–45,47–

Table 7 demonstrates the coordinated rise and fall of GH pulsatility, IGF-1, and IGFBP-3 across puberty, showing how maturation markedly amplifies the GH–IGF-1 axis. Tanner I is characterized by low IGF-1/IGFBP-3 and modest GH amplitude, while Tanner II marks the beginning of a steroid-driven rise. Tanner III–IV represent peak physiologic activation, with the highest GH pulses and IGF-1 levels; failure to mount this rise strongly suggests true GHD rather than constitutional delay. By Tanner V, IGF-1 and GH output decline physiologically toward adult levels, meaning lower GH peaks are not pathological. These patterns underscore the need to interpret GH and IGF-1 relative to Tanner stage to avoid misclassification, especially in the presence of obesity or delayed puberty.

4. Discussion

The present narrative review integrates 25 years of evidence to clarify how normal GH physiology and its modulation by puberty, sleep, nutrition, and body composition intersect with the diagnostic performance of current tests for pediatric GHD. Studies of frequent-sampling profiles consistently demonstrate that GH secretion is present from early childhood but that puberty is associated with a two- to three-fold increase in 24-hour secretion, driven mainly by larger secretory burst mass and pulse amplitude rather than changes in frequency (16–18,20,22,23). These dynamic changes, together with the long-term integrative signals provided by IGF-1 and IGFBP-3, underscore that any diagnostic strategy must be anchored in an understanding of developmental stage rather than relying on static, age-independent thresholds.

Across Tanner stages, IGF-1 and IGFBP-3 mirror the pubertal activation of GH secretion, rising from low-normal levels in Tanner I to peak concentrations in Tanner III–IV before declining towards young adult values in Tanner V (43–47). Simultaneously, pulse amplitude and 24-hour GH output show maximal values in mid- to late puberty, particularly in boys, while frequency remains relatively stable (16–18,20–22,24,48,49). Tables 6 and 7 highlight that Tanner III–IV provide a physiological “stress test” of the GH–IGF-1 axis; at these stages, failure of IGF-1 and GH peaks to rise into the expected range is much more indicative of true GHD than comparable findings in prepubertal children, in whom overlap with constitutional delay is substantial (16–18,20–22,24,43–45,47–49).

Sleep and other contextual factors further complicate interpretation. Early-night slow-wave sleep is a major synchronizer of large GH pulses, and age-related loss of slow-wave sleep partly explains the decline in GH output from adolescence into adulthood (20). Acute sleep disruption attenuates nocturnal peaks but may be compensated by daytime secretion, so a single nocturnal profile may underestimate true 24-hour secretion if daytime sampling is omitted (26). Nutritional quality, including micronutrient intake such as vitamin C, also modulates spontaneous GH burst mass and total output, adding further physiological noise to GH measurements in children with variable diets or chronic illness (25,46). These data argue that random or limited GH sampling without contextual information on sleep and nutrition is particularly prone to misclassification.

Within this complex physiological backdrop, GH stimulation tests remain the most widely accepted confirmatory tool for pediatric GHD. Across multicenter cohorts, single tests using insulin-induced hypoglycemia, clonidine, arginine, or glucagon typically achieve sensitivities of 80–90% but specificities of only 60–75%, with modest reproducibility between stimuli in the same child (7,27–29,31). No single stimulus is clearly superior, and discordant results between different tests are frequent, particularly near arbitrary cut-points (27–29,31). Concordant failure of two independent tests improves both sensitivity and specificity, supporting guideline recommendations to avoid relying on one test in isolation when the clinical picture is equivocal (7,27,29,31).

Assay evolution over the past two decades has further complicated interpretation of GH stimulation results. Modern chemiluminescent and mass-spectrometry–aligned assays generally yield lower peak GH values than older radioimmunoassays, making historical cut-points of 10 ng/mL overly stringent in many settings (9,28,31,49). ROC-based analyses in contemporary cohorts suggest that optimal thresholds often lie in the 6–8 ng/mL range, depending on the assay, age, and pre-test probability (28,31). Failure to recalibrate cut-offs to the specific assay in use leads to over-diagnosis of GHD in lean children and under-diagnosis in obese children, reinforcing the need for laboratory-specific reference data and clear reporting of assay characteristics in both clinical practice and research (9,28,31,49).

IGF-1 and IGFBP-3 have long been proposed as non-invasive screening tools, yet meta-analytic and cohort data confirm that they are supportive rather than definitive markers. Pooled estimates show IGF-1 sensitivities of ~60–70% and specificities of 65–75%, with IGFBP-3 generally slightly more specific but less sensitive (5,6,33,37,38). Bayesian modeling and real-world clinic cohorts indicate that IGF-1-SDS performs best as a high-sensitivity triage test, while IGFBP-3-SDS adds specificity, especially in younger children, when both are interpreted alongside auxology and bone age (34,37–39). However, prospective data from broad, mixed-BMI populations demonstrate that IGF-1 alone can have poor AUCs around 0.5, emphasizing that normal values cannot reliably exclude GHD in many clinical contexts (36).

More sophisticated strategies combining multiple markers appear promising. The IGF-1/IGFBP-3 molar ratio and composite indices that integrate both proteins have shown higher diagnostic discrimination than either analyte alone, with sensitivities approaching 85–90% and specificities >80–95% in selected cohorts (35,38). Very low IGF-1 together with low IGFBP-3 and a low ratio strongly support GHD, whereas completely normal values across all three tests make the diagnosis unlikely (35,38). The IGF-1 generation test provides dynamic information on peripheral GH sensitivity and can help differentiate primary IGF-1 deficiency or GH insensitivity from classic GHD, though its clinical utility is limited by non-standardized protocols, variable diagnostic accuracy, and constrained availability (41,42). Consequently, these more complex tools are best reserved for specialized centers and atypical or syndromic cases rather than routine screening.

Obesity and BMI emerge across multiple datasets as dominant modifiers of both GH secretion and diagnostic test performance. Increased adiposity is associated with lower 24-hour GH production, reduced secretory burst mass, and diminished stimulated peaks, even in otherwise healthy adolescents, likely mediated by hyperinsulinemia, elevated free fatty acids, and leptin-driven changes in hypothalamic tone (10,11,17,21,24). Overweight youth consistently demonstrate blunted nocturnal GH pulsatility and lower responses to pharmacologic stimuli, such that obese short children may fail to reach conventional GH cut-offs without having structural or genetic GHD (17,21,24,30,31). At the same time, obesity can modestly lower IGF-1 in non-GHD controls, reducing the specificity of low IGF-1 as a screening marker and contributing to heterogeneity across cohorts (10,32,33,36). These findings demand BMI-aware interpretation of both GH and IGF-1 and argue against rigid, single cut-points irrespective of body composition.

Sex and pubertal status add further layers of complexity. Rising estradiol during puberty, in both boys and girls, amplifies GH pulse amplitude and drives the characteristic surge in IGF-1 and IGFBP-3, with boys typically achieving slightly higher mid-pubertal IGF-1 peaks and GH responses than girls (12,16–18,20–22,24,43–45,47–49). The tables in this review show that mid-pubertal Tanner III–IV stages correspond to the highest GH and IGF-1 values, whereas Tanner I and early Tanner II display substantial overlap between normal children, those with constitutional delay, and those with early GHD (16–18,20–22,24,43–47). In Tanner V, physiologic declines in GH amplitude and IGF-1 must not be mistaken for pathologic deficiency, particularly if adult reference ranges are not yet applied (20–22,43–45,47–49). Thus, any diagnostic algorithm must incorporate Tanner-stage–specific expectations rather than using a uniform pediatric range.

When the relative strengths and weaknesses of available tests are compared, a rational hierarchy emerges. GH stimulation tests retain their central role as confirmatory tools but should be interpreted in light of assay characteristics, BMI, puberty, and clinical context rather than as binary pass–fail criteria (7,8,11,27–32). IGF-1 and IGFBP-3, alone or combined, serve best as first-line screening and risk-stratification markers, guiding which children warrant dynamic testing and which can be observed with close auxologic follow-up (5,6,33–39). Nocturnal GH profiles and deconvolution analysis provide unparalleled insight into physiology but are too labor-intensive and non-specific for routine use, while IGF-1 generation tests and composite indices are niche tools for complex or equivocal cases in expert centers (13,20,23,35,38,41,42). Pituitary MRI then consolidates the diagnostic pathway by identifying structural lesions in children with strong biochemical and clinical evidence of GHD (14,27).

Synthesizing these data, an integrated diagnostic framework can be proposed. The first step is rigorous auxologic assessment, including height SDS, growth velocity, mid-parental height, and bone age, together with evaluation of systemic disease, nutrition, and psychosocial factors (1,2,15). In children with clearly abnormal growth, stage-adjusted IGF-1 and IGFBP-3 (and, where available, their ratio) are used to refine pre-test probability; profoundly low values, especially at Tanner III–IV, strongly favor GHD, whereas normal values at Tanner V or in obese children require cautious interpretation (33–39,43–49). GH stimulation testing using two distinct stimuli and assay-specific cut-offs is then reserved for those with moderate-to-high suspicion, with results interpreted in the context of BMI, puberty, and comorbidities rather than as absolute truths (7–9,11,27–32). Children with concordant biochemical evidence and significant growth failure undergo pituitary MRI; those with discordant or borderline results may be managed using longitudinal growth monitoring, repeat testing, or specialized studies such as IGF-1 generation in selected scenarios (14,35,38,41,42).

Several limitations of the available evidence must be acknowledged. Most diagnostic studies are single-center or regional, with heterogeneous inclusion criteria, referral patterns, and definitions of GHD, leading to spectrum and selection biases (7,8,27–32,33–38). Assay differences and evolving calibration standards hinder cross-study comparisons and limit the generalizability of proposed cut-offs, particularly when historical data are applied to current platforms (9,28,31,49). Few studies stratify results systematically by BMI, ethnicity, or Tanner stage, despite clear physiological evidence that these factors materially affect GH and IGF-1 values (10–12,16–18,20–24,43–47). Furthermore, the absence of a true gold standard for pediatric GHD—given that long-term growth response to therapy

is influenced by many non-GH factors—means that most accuracy estimates are anchored to imperfect reference standards and should be interpreted cautiously.

Despite these constraints, the body of work from 2000–2025 provides a more nuanced understanding of GH secretion patterns and diagnostic test behavior than was previously available. The consistent demonstration of pubertal amplification and obesity-related suppression of GH, the moderate but useful performance of IGF-1 and IGFBP-3, and the assay-dependent nature of stimulation cut-offs collectively argue against simplistic, one-size-fits-all criteria for pediatric GHD. Instead, the evidence supports a personalized, context-sensitive approach that integrates auxology, Tanner stage, BMI, biochemical markers, and imaging into a stepwise diagnostic pathway. Future research should prioritize large, multicenter cohorts using harmonized assays, BMI- and Tanner-specific reference ranges, and outcome-based validation against long-term growth and metabolic endpoints. Such data will be essential to refine diagnostic thresholds, reduce both over- and under-diagnosis, and ultimately ensure that children with true GHD receive timely, appropriate treatment while those with constitutional or obesity-related variants are spared unnecessary labeling and therapy.

5. Conclusion

Over two decades of evidence clearly show that diagnosing pediatric growth hormone deficiency requires an integrated, context-sensitive approach rather than reliance on isolated biochemical thresholds. GH secretion is highly dynamic across childhood and puberty, with major increases in pulse amplitude and IGF-1/IGFBP-3 concentrations during Tanner III–IV and physiologic declines in late adolescence, while factors such as obesity, sleep, nutrition, and assay variability further modify test performance. GH stimulation tests remain essential but have limited specificity and are strongly influenced by BMI, pubertal status, and assay generation, whereas IGF-1 and IGFBP-3 provide useful—but not definitive—screening information. The most reliable diagnosis emerges when auxology, Tanner stage, BMI, biochemical markers, and pituitary MRI are interpreted together within a structured, stepwise framework. A modern diagnostic strategy must therefore emphasize individualized interpretation, laboratory-specific cut-offs, and longitudinal monitoring to reduce both under- and over-diagnosis and to ensure that true cases of GHD are identified early and managed appropriately.

Recommendations

- Use an integrated diagnostic algorithm: Begin with auxology and Tanner staging, apply IGF-1/IGFBP-3 for screening, and reserve two GH stimulation tests plus pituitary MRI for cases with sustained high suspicion.
- Interpret results contextually: Adjust GH and IGF-1 values for assay type, BMI, and pubertal status, avoiding fixed historical cut-offs and documenting relevant clinical modifiers.
- Monitor borderline cases over time: Rely on growth velocity, repeat testing, and multidisciplinary review rather than immediate diagnosis, reducing overtreatment while ensuring true GHD is identified promptly.

Compliance with ethical standards

Disclosure of conflict of interest

The authors declare no conflicts of interest related to this work.

Authors' Contributions

All authors contributed substantially to the development of this review. ATS conceived the scope, supervised the literature synthesis, and finalized the manuscript. FA, NA, NH, and SA conducted the primary literature search, data extraction, and critical appraisal of included studies. DF and AE contributed to interpretation of pediatric endocrine and diagnostic aspects and assisted in manuscript drafting. AK contributed to reviewing methodological rigor and refining the narrative structure. All authors critically reviewed the final manuscript, approved its content, and agreed to be accountable for all aspects of the work.

References

[1] Stanley T, Levitsky LL, Grinspoon S, et al. Diagnosis of growth hormone deficiency in childhood. Curr Opin Endocrinol Diabetes Obes. 2012;19(1):47–52.

[2] Grimberg A, DiVall SA, Polychronakos C, 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(6):361–397. doi:10.1159/000452150

[3] Ho KY, Evans WS, Blizzard RM, et al. Effects of sex and age on the 24-hour profile of growth hormone secretion in man: importance of endogenous estradiol concentrations. J Clin Endocrinol Metab. 1987;64(1):51–58.

[4] Rose SR, Municchi G, Barnes KM, et al. Spontaneous growth hormone secretion patterns in normal children. J Clin Endocrinol Metab. 1991;72(1):218–228.

[5] Ibba A, Guzzetti C, Salerno M, et al. IGF1 for the diagnosis of growth hormone deficiency in children and adolescents: a reappraisal. Endocr Connect. 2020;9(11):1095–1102. doi:10.1530/EC-20-0347

[6] Lissett CA, Shalet SM, Toogood AA, et al. Growth hormone deficiency: clinical and biochemical assessment. Clin Endocrinol (Oxf). 2003;59(6):719–723.

[7] Clayton PE, Cuneo RC, Juul A, et al. Consensus statement on the management of the child with growth hormone deficiency. J Clin Endocrinol Metab. 2014;99(9):3084–3092.

[8] Grimberg A, Feemster KA, Pati S, et al. Variability in growth hormone stimulation testing and treatment practices in pediatric endocrine centers. J Clin Endocrinol Metab. 2015;100(1):182–189.

[9] Bidlingmaier M, Friedrich N, Emeny RT, et al. Reference intervals for insulin-like growth factor-1 (IGF-I) from birth to senescence: results from a multicenter study using a new automated chemiluminescence IGF-I immunoassay conforming to recent international recommendations. J Clin Endocrinol Metab. 2014;99(5):1712–1721.

[10] Cornier MA, Dabelea D, Hernandez TL, et al. The metabolic syndrome. Endocr Rev. 2008;29(7):777–822.

[11] Bizzarri C, Benevento D, Bottaro G, et al. Growth hormone stimulation test results are influenced by body mass index in children. Front Endocrinol (Lausanne). 2021;12:739118. doi:10.3389/fendo.2021.739118

[12] Blum WF, Crowe BJ, Quigley CA, et al. Growth hormone is effective in treatment of short stature associated with SHOX deficiency: two-year results of a randomized, controlled, multicenter trial. J Clin Endocrinol Metab. 2013;98(1):219–228.

[13] MacGillivray MH, Marks JF, Rivkees SA, et al. Nocturnal growth hormone secretion in children and adolescents: diagnostic limitations and clinical implications. Pediatr Clin North Am. 1990;37(6):1231–1248.

[14] Maghnie M, Lindberg A, Koltowska-Häggström M, et al. Magnetic resonance imaging of the hypothalamic–pituitary region in children with growth hormone deficiency treated with recombinant human growth hormone. J Clin Endocrinol Metab. 2006;91(3):1006–1012.

[15] Rogol AD, Hayden GF, Rosenfeld RG, et al. Short stature in childhood—challenges and choices. Pediatr Rev. 2014;35(8):327–340.

[16] Miller JD, Tannenbaum GS, Colle E, et al. Daytime pulsatile growth hormone secretion during childhood and adolescence. J Clin Endocrinol Metab. 1982;55(5):989–994.

[17] Martha PM Jr, Rogol AD, Blizzard RM, et al. Growth hormone secretory dynamics in normal boys and the effects of age and puberty. J Clin Endocrinol Metab. 1992;75(2):452–457.

[18] Mauras N, Blizzard RM, Link K, et al. Augmentation of growth hormone secretion during puberty: evidence for increased growth hormone-releasing hormone drive. J Clin Endocrinol Metab. 1987;64(1):596–601.

[19] Wright NM, Northington FJ, Miller JD, et al. Elevated growth hormone secretory rate in premature infants: deconvolution analysis of pulsatile growth hormone secretion in the neonate. Pediatr Res. 1992;32(3):286–290.

[20] Van Cauter E, Latta F, Gottschalk WA, et al. Age-related changes in slow wave sleep and REM sleep and relationship with growth hormone and cortisol levels in healthy men. JAMA. 2000;284(7):861–868.

[21] Roemmich JN, Clark PA, Berr SS, et al. Alterations in growth and body composition during puberty. I. Growth hormone and body composition in normal-weight vs overweight youth. J Clin Endocrinol Metab. 2002;87(9):4154–4161.

[22] Roemmich JN, Berr SS, Clark PA, et al. Pubertal alterations in growth and body composition. II. The influence of nocturnal growth hormone and sex steroids. J Clin Endocrinol Metab. 2005;90(6):3112–3117.

[23] Veldhuis JD, Roemmich JN, Richmond EJ, et al. Endocrine control of body composition in infancy, childhood, and puberty. Endocr Rev. 2008;29(6):646–682.

[24] Kasa-Vubu JZ, Lee CC, Rosenthal A, et al. Influence of adiposity and leptin on growth hormone secretion in postpubertal adolescent girls. J Clin Endocrinol Metab. 2006;91(11):4777–4783.

[25] Denny-Brown LC, Cohen JD, Hsia DS, et al. Nutritional determinants of spontaneous growth hormone secretion in adolescents and young adults. Clin Endocrinol (Oxf). 2012;76(6):861–867.

[26] Calvert DG, Beebe DW, Grover K, et al. Sleep disruption reduces nocturnal growth hormone secretion in pubertal children but preserves 24-hour output. Sleep Med. 2022;90:134–141.

[27] Maghnie M, Aimaretti G, Bellone S, et al. Diagnosis of growth hormone deficiency in the child and adolescent. Pituitary. 2005;8(1):29–38. doi:10.1007/s11102-005-5083-0

[28] Zimmermann AG, Loche S, Cappa M, et al. Effect of growth hormone assay and cut-off on the diagnosis of growth hormone deficiency in children. Horm Res Paediatr. 2011;75(5):335–341. doi:10.1159/000323807

[29] Ibba A, Loche S, Guzzetti C, et al. Comparison between two growth hormone stimulation tests in the diagnosis of growth hormone deficiency in children. Clin Endocrinol (Oxf). 2014;81(5):748–754. doi:10.1111/cen.12476

[30] Elbornsson M, Götherström G, Bosaeus I, et al. Re-evaluation of growth hormone (GH) secretion in adults with childhood-onset GH deficiency: impact of body mass index and choice of stimulation test. J Clin Endocrinol Metab. 2015;100(7):2493–2501. doi:10.1210/jc.2014-4453

[31] Felício JS, Nery M, Oliveira F, et al. The importance of GH provocative tests and IGF-1 in the diagnosis of isolated GH deficiency. Front Endocrinol (Lausanne). 2019;10:638. doi:10.3389/fendo.2019.00638

[32] Fatani TH, Al-Harbi A, Bin-Abbas B, et al. Diagnostic value of IGF-1 for growth hormone deficiency in children and its impact on the need for a second GH stimulation test. J Endocr Soc. 2023;7(4):bvad018. doi:10.1210/jendso/bvad018

[33] Shen Y, Zhang J, Zhao Y, et al. Diagnostic value of serum IGF-1 and IGFBP-3 in growth hormone deficiency: a systematic review and meta-analysis. Eur J Pediatr. 2015;174(4):419–427. doi:10.1007/s00431-014-2419-7

[34] Inoue-Lima TH, Vasques GA, Nakaguma M, et al. 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. 2020;93(3):197–205. doi:10.1159/000509840

[35] Haj-Ahmad LM, Mahmoud MM, Sweis NWG, et al. Serum IGF-1 to IGFBP-3 Molar Ratio: A Promising Diagnostic Tool for Growth Hormone Deficiency in Children. *J Clin Endocrinol Metab*. 2023;108(4):986-994. doi:10.1210/clinem/dgac609

[36] Iwayama, H., Kitagawa, S., Sada, J. *et al.* Insulin-like growth factor-1 level is a poor diagnostic indicator of growth hormone deficiency. *Sci Rep* **11**, 16159 (2021). https://doi.org/10.1038/s41598-021-95632-0

[37] Galluzzi F, Quaranta G, Salti R, et al. Are IGF-I and IGF-binding protein-3 useful for diagnosing growth hormone deficiency in children of short stature? J Pediatr Endocrinol Metab. 2010;23(12):1273–1279. doi:10.1515/jpem.2010.201

[38] Bereket A, Turan S, Omar A, et al. Diagnosis of growth hormone deficiency: the role of growth hormone provocative tests, IGF-1, and IGFBP-3. J Clin Res Pediatr Endocrinol. 2009;1(3):122–129. doi:10.4008/jcrpe.v1i3.21

[39] Kota SK, Meher LK, Jammula S, et al. Serum IGF-1 and IGFBP-3 levels in growth hormone-deficient and non–growth hormone-deficient short children. J Nepal Paediatr Soc. 2012;32(3):173–178.

[40] Je YS, Lee SG, Kim JH, et al. Usefulness of serum IGF-I and IGFBP-3 levels in children with short stature. J Lab Med Qual Assur. 2014;36(1):48–55. doi:10.15263/jlmqa.2014.36.1.48

[41] Buckway CK, Guevara-Aguirre J, Pratt KL, et al. The IGF-I generation test revisited: a marker of GH sensitivity. J Clin Endocrinol Metab. 2001;86(11):5176–5183. doi:10.1210/jcem.86.11.8039

[42] Coutant R, Dörr HG, Gleeson H, et al. Diagnosis of endocrine disease: limitations of the IGF-1 generation test in children with short stature. Eur J Endocrinol. 2012;166(3):351–357. doi:10.1530/EJE-11-0618

[43] Bereket A, Turan S, Omar A, et al. Serum IGF-I and IGFBP-3 levels in Turkish children during childhood and adolescence: relation to pubertal development. Horm Res. 2006;65(2):96–105. doi:10.1159/000091301

[44] Hyun SE, Lee BC, Suh BK, et al. Reference values for serum insulin-like growth factor-I and IGF-binding protein-3 in Korean children and adolescents. Clin Biochem. 2012;45(1–2):16–21. doi:10.1016/j.clinbiochem.2011.10.012

[45] Juul A, Bang P, Hertel NT, et al. Serum insulin-like growth factor-I in 1030 healthy children, adolescents, and adults: relation to age, sex, stage of puberty, testicular size, and body mass index and to serum levels of growth hormone–binding protein, insulin-like growth factor binding protein-3, and testosterone. Pediatr Res. 1994;36(4):455–461. doi:10.1203/00006450-199410000-00014

[46] Blum WF, Albertsson-Wikland K, Rosberg S, et al. Serum levels of insulin-like growth factor I (IGF-I) and IGF binding protein-3 reflect spontaneous growth hormone secretion. J Clin Endocrinol Metab. 1993;76(6):1610–1616. doi:10.1210/jcem.76.6.8501165

[47] Cao B, Elshimy G, Zhai X, et al. Pediatric continuous reference intervals of serum insulin-like growth factor 1 on the Roche cobas e602 analyzer. Endocr Pract. 2022;28(11):1137–1146. doi:10.1016/j.eprac.2022.06.002

[48] Russell M, Bebbington A, O'Donovan C, et al. Growth hormone is positively associated with surrogate markers of pubertal testosterone in boys and estradiol in girls. J Clin Endocrinol Metab. 2011;96(4):E622–E631. doi:10.1210/jc.2010-2032

[49] Yau M, Ng D, Breheny P, et al. Peak growth hormone response to combined stimulation test and its associations with auxologic and biochemical parameters. Horm Res Paediatr. 2019;92(1):36–45. doi:10.1159/000500923