

10.4274/jcrpe.galenos.2019.2019.0176

review

**A proposal for the interpretation of the serum IGF-I concentration as part of laboratory screening in children with growth failure**

**Short Title: Interpreting serum IGF-I in short children**

Jan M Wit<sup>1</sup>, Martin Bidlingmaier<sup>2</sup>, Christiaan de Bruin<sup>1</sup>, Wilma Oostdijk<sup>1</sup>

<sup>1</sup>Department of Paediatrics, Leiden University Medical Center, Leiden, the Netherlands

<sup>2</sup>Endocrine Research Laboratories, Medizinische Klinik und Poliklinik IV, Klinikum der Universität München, 80336 Munich, Germany

No grants or fellowships supporting the writing of the paper

ORCID numbers:

J.M.Wit: 0000-0002-1715-5020

M.Bidlingmaier: 0000-0002-4681-6668

C.de Bruin: 0000-0003-1478-2261

W.Oostdijk: 0000-0003-0180-2470

**Abstract**

The serum IGF-I concentration is commonly used as a screening tool for growth hormone deficiency (GHD), but there is no consensus on the cut-off limit of IGF-I standard deviation score (SDS) to perform GH stimulation tests for confirmation or exclusion of GHD. We argue that the cut-off limit is dependent on the clinical pre-test likelihood of GHD and propose a diagnostic strategy in which the cut-off limit varies between none to -2 SDS.

**Keywords:** short stature, growth disorders, IGF-I, IGFBP-3

Corresponding Author and name and address of person to whom reprint requests should be addressed :

Prof. dr. Jan M. Wit,

Department of Paediatrics

Leiden University Medical Center

Albinusdreef 2

2333 ZA Leiden,

Netherlands

Tel: +31 71 5262824  
Fax: +31 71 5248198  
E-mail: j.m.wit@lumc.nl  
Submitted: 04-Nov-2019  
Accept: 14-Dec-2019

## **Introduction**

In guidelines on the diagnostic approach to children referred to a paediatrician or paediatric endocrinologist for short stature and/or growth faltering (from now on referred to as “growth failure”, abbreviated as GF), including the consensus paper on Idiopathic short stature (ISS) (1) and the recent Dutch guideline (2), it is advised to perform laboratory screening for potential subclinical pathological causes, including serum IGF-I (with or without serum IGFBP-3), in order to screen for growth hormone deficiency (GHD). During the pre-final phase of the Dutch guideline, when authorization was sought from representatives of the Paediatric Association of the Netherlands and other specialist societies, the working group was asked to add specific instructions for the general paediatrician at which cut-off point of IGF-I a GH stimulation test should be performed to confirm or exclude GHD. Since a literature search did not provide a clear answer to this question, we performed a stepwise analysis of the decision procedure, taking established clinical epidemiological techniques into consideration. This led to several subsequent versions of an addendum to the Dutch Guideline on Triage and Diagnosis of Growth disorders in children (2), which were reviewed by members of the Section of Paediatric Endocrinology and the Growth Hormone Advisory Group of the Paediatric Association of the Netherlands and discussed at several meetings. In the present commentary we present an English adaptation of the addendum to the official guideline that will be published in Dutch on the internet.

### **Definition, subcategories and varying levels of uncertainty of GHD**

Impaired secretion of GH in children is causally related to impaired growth, anthropometric characteristics (e.g., normal proportions, relatively large head) and changes of body composition (e.g., sarcopenia, excess of body fat and low bone mineral density) as well as functional abnormalities (e.g., hypoglycaemia) (3, 4). However, for practical and financial reasons it is troublesome to measure the GH secretion over 24 hours in individual patients, so that as a proxy indicator of spontaneous GH secretion the GH response to a stimulus is commonly used. The various problems of such stimulation tests are well known (5) and also low serum levels of GH-

dependent proteins (IGF-I and IGFBP-3) cannot be considered gold standards.

GHD can be subcategorized in acquired and non-acquired (congenital) forms (4). The acquired form is usually caused by space-occupying processes in the brain, such as tumours (e.g., craniopharyngioma), but can also be the consequence of brain trauma, infections, irradiation, histiocytosis or vascular anomalies. The congenital form can be subdivided into three subgroups (6). The first is an (almost) certain form where the etiology is known, for example if GHD is associated with cerebral or facial malformations, anatomic pituitary or hypothalamic abnormalities detected with magnetic resonance imaging (MRI), other pituitary hormone deficiencies, or established genetic causes of GHD, such as mutations of *GHI*, *GHRHR* or *GHSR* in isolated GHD (IGHD) and of multiple transcription factors in multiple pituitary hormone deficiency (MPHD). In line with Ranke et al (6) this etiologic subgroup is abbreviated as cGHD. The second and most frequent subgroup is idiopathic isolated GHD (IIGHD). The third subgroup is growth hormone neurosecretory dysfunction (NSD), characterised by clinical features (including a “typical” growth curve) and low serum IGF-I which lead to suspicion of GHD, in combination with a normal GH peak at a stimulation test. In the original papers it was shown that such children had a low spontaneous GH secretion over 24 hours in contrast to a normal GH peak in response to a GH stimulation test (7-9). Due to the practical and financial hurdles of performing 24-hour GH profiles, this diagnosis has also been assumed without proof of a decreased spontaneous GH secretion (6).

The diagnostic process of acquired GHD and cGHD is straightforward, although sometimes with considerable diagnostic delay, but this does not apply to IIGHD and NSD. Retesting of the hypothalamic-pituitary-GH/IGF-I axis in puberty and after reaching adult height of patients with IIGHD leads in most cases to normal results regarding the GH peak in a GH stimulation test (10-12). This phenomenon is usually interpreted as an initial false-positive result of GH testing, although a transient form of GHD cannot be excluded. The diagnosis of NSD has been controversial from the beginning, although an adequate growth response to GH treatment of children diagnosed as NSD has been documented (7-9).

### **Diagnostic approach with respect to GHD in children referred for growth failure**

There is no doubt that the diagnostic process for children referred for GF should be aimed at detecting and treating all children with acquired GHD

or cGHD as soon as possible. In the remaining children (IIGHD and NSD), which constitute the majority of cases (6)), the clinician has to face the challenge to distinguish as best as possible IIGHD or NSD from other causes of GF or short stature of unknown origin (idiopathic short stature, ISS). It is particularly difficult to distinguish IIGHD from children with the non-familial form of ISS with maturational delay (1, 3, 4). Many prepubertal short children who later present delayed puberty, have a prepubertal growth pattern characterised by a low height velocity (decreasing height SDS) (13, 14). However, the diagnosis of “Constitutional Delay of Growth and Puberty” (CDGP) can only be diagnosed with certainty if the onset of puberty (Tanner genital or breast stage 2) is indeed delayed (boys >14 years, girls >13 years).

For an estimate of the probability of GHD in a child with GF (in comparison to another or unknown cause), the clinician needs to take the following steps:

- 1) Be aware of the *general* prevalence of GHD in children referred for GF in the specific clinic;
- 2) estimate the *individual pre-test likelihood of GHD* in the patient, based on weighing the diagnostic clues from the medical history (including family history), physical examination, growth curve, laboratory screening and skeletal maturation;
- 3) perform *laboratory screening* using serum IGF-I and (in young children) IGFBP-3, and interpret the result taking into consideration age, sex, pubertal status and body mass index
- 4) calculate the *post-test likelihood of GHD* based on serum IGF-I (with or without IGFBP-3) and individual pre-test likelihood;
- 5) if the post-test likelihood is considered sufficiently high, perform GH-provocation tests (*GH-stimulation tests*);
- 6) If the results of the GH stimulation test are compatible with GHD, perform an MRI of the hypothalamus-pituitary area (in specific cases genetic testing can also be considered);
- 7) if the GH peak is normal in the setting of a severely decreased serum IGF1, consider performing an IGF-I generation test (to differentiate between rare syndromes with either normal or decreased GH sensitivity) and/or genetic tests, preferably after consultation of a paediatric endocrinologist and/or clinical geneticist knowledgeable of genetic causes of GF.

In this Critical Review we shall discuss the first four steps. For the remaining three steps we refer to previous documents, e.g. (15-18).

## **Prevalence of GHD and its subcategories in children referred for growth failure**

The prevalence of GHD in the general population of children is uncertain. Estimates vary from 1:3.400 to 1:30.000 (19-23). If one estimates a prevalence of 1:5000, and assumes that almost all cases present with a height below -2 SDS (2.3<sup>rd</sup> percentile), the prevalence of GHD in short children in the population would be approximately 1%. The reported prevalence of GHD in children referred for GF to a paediatric (endocrine) clinic also varies. In Dutch studies the prevalence was between 0.5-3% (24-26) and in other studies prevalences between 0 and 23% were reported (27-34). For this analysis we estimated the general prevalence of GHD in children referred for GF at 2%, in line with a recent publication (35), but this percentage may differ between a general and academic paediatric clinic. Furthermore, the variation in diagnostic approaches, for example whether sex hormone priming is used when a GH stimulation test is performed in the pre- to early-pubertal age range (36), can also be expected to have a significant impact on the percentage of GHD diagnosed in referred children.

We have not found data on the proportion of acquired and congenital GHD, but based on our clinical experience acquired GHD is rare. Within the group of congenital GHD an indication of the relative proportions can be obtained from the report on 50 years' experience in Tübingen: out of 636 patients, 122 had cGHD (19.2%), 455 IIGHD (71.7%) en 58 (9.1%) NSD (6).

## **Estimate of the individual pre-test likelihood of GHD in a patient referred for growth failure**

A full clinical assessment of the child will lead to an estimate of the individual clinical pre-test likelihood of GHD, based on an inventory of all relevant clinical features. These comprise diagnostic clues from the medical history, physical examination, growth curve, general laboratory screening and skeletal age. Although objective data are lacking, we believe that the relative weight of these diagnostic clues is different. **Table 1** shows the diagnostic clues and our personal opinion on their relatively importance. Future prospective studies are needed to collect observational data on the relative weight of the various diagnostic clues. Based on the balance of positive and negative diagnostic clues the clinician can estimate the clinical pre-test likelihood of GHD in an individual patient. Thereafter, the result of serum IGF-I (and IGFBP-3) can be interpreted against the background of the pre-test likelihood of GHD.

The estimate of the pre-test likelihood of GHD is obviously a continuum and subjective, since objective data are lacking of the proportional weighing of diagnostic clues. For practical purposes, we propose five categories of pre-test likelihood:

1. High pre-test likelihood ( $\approx 50\%$ ), if there are several positive clues and no negative clues
2. Moderate pre-test likelihood ( $\approx 20\%$ ), if there are some positive clues for GHD, and no negative clues
3. Rather low pre-test likelihood ( $\approx 10\%$ ), if there are few positive clues for GHD and no negative clues
4. Low pre-test likelihood ( $\approx 5\%$ ), if there are few positive clues and at least one negative clue for GHD.
5. Very low pre-test likelihood ( $\leq \approx 2\%$ ), if there are no positive clues and/or  $\geq 1$  negative clues for GHD; this is roughly equal to (or less than) the estimated prevalence in children referred for growth failure.

In case of a high pre-test likelihood of GHD the clinician will directly perform an MRI of the pituitary/hypothalamic area, particularly in neonates or if there is clinical suspicion of acquired GHD. In neonates with clinical and biochemical features of hypothalamic/pituitary deficiency an abnormal brain MRI in combination with a random serum GH concentration  $< 7 \mu\text{g/L}$  in the first week of life has been proposed as sufficient to make the diagnosis of GHD (37). A low serum IGF-I supports the diagnosis, but its specificity at that age is low because the lower limit of the reference range is then below the sensitivity of the assay. A low IGFBP-3 ( $< -2.0$  SDS) is also supportive, and in general its specificity appears to be higher than that of IGF-I (38-40).

Also when an acquired form of GHD is suspected, the clinician will promptly perform a brain MRI, in addition to further diagnostic investigations into hypothalamic/pituitary function. Such investigations will usually be repeated following neurosurgical or oncological treatment. During follow-up of neuro-oncological patients with an increased risk of developing GHD, a recent international guideline advised to perform GH stimulation tests when there is clinical suspicion of GHD, irrespective of the result of serum IGF-I, because in such patients a low serum IGF-I appears to have a low sensitivity (41). It has been suggested that if multiple pituitary axes are affected, a GH stimulation test can be omitted because of the high pre-test likelihood of GHD (41). In patients with possible IIGHD, a brain MRI is only performed after GHD has been confirmed by GH stimulation tests.

### **Laboratory screening using serum IGF-I and interpreting the result adjusting for age, sex, pubertal status and body mass index**

The serum concentration of IGF-I is not only influenced by GH secretion, but also by biological factors, such as nutritional status, age, sex, and pubertal stage, as well as analytical factors. This implies that if the serum IGF-I concentration is measured in the context of laboratory screening as an indicator of GHD in children with GF, the result should first be adjusted for confounding factors.

In most countries different commercial IGF-I assays are used with a considerable inter-assay variation. In our country, the major part of the inter-assay variation of 5-20% is compensated because of a national harmonization program. For each assay there is also some intra-assay variation (e.g., a coefficient of variation of 6-8% in the assay used in Tübingen (6) and a total coefficient of variation 3.4%–8.7% in the iSYS assay (42)), so that repetition of an IGF-I determination should be considered, particularly if there is a discrepancy between the clinical features and the laboratory results. Given the inter-assay variation, it is advisable to use the same assay during follow-up of an individual patient. For a proper interpretation of the serum IGF-I concentration, the first step is to express IGF-I as standard deviation score (SDS) for age and sex. Obviously, for this purpose adequate reference data from a large population of healthy children are needed. Given the differences between assays, reference data should be made for each assay separately.

It is generally assumed that in prepubertal children in the age range where puberty usually has not yet started IGF-I SDS can be calculated for age and sex. However, the maximum age for this approach is arbitrary. One could argue that the 10<sup>th</sup> percentile of the age at start of puberty in the population would be a rational choice (9 years in girls and 10 years in boys in our country), but other investigators suggested a cut-off of 8 and 9 years, respectively (35).

Regarding adjustment for pubertal status in older children, the situation is more complex, and a further adjustment step has to be performed. For this purpose, there are essentially two possible approaches. For a few assays reference data have been reported on IGF-I per pubertal stage (for example (42-44)), but the assays used in the two oldest reports were calibrated to old standards. IGF-I percentiles according to pubertal (Tanner) stage for the IDS iSYS assay are shown in **Table 2** (42).

If an assay is used for which such reference data are unavailable, one can calculate SDS for skeletal age (35), because a (relatively) delayed puberty is usually associated with a (relatively) delayed skeletal age. However, its

predicted power is reportedly slightly lower than that of the SDS for pubertal stage (35).

Since the expression of IGF-I as an SDS is more informative than a percentile position, we calculated mean, SD and cut-off limits for various SDS for children with different Tanner stages based on the raw data that were used for the construction of reference data for the IDS iSYS assay (42) (**Table 3**).

A further complication when assessing serum IGF-I occurs if BMI SDS is low or high (45). In a child with a low BMI SDS, another sign of undernutrition, or a recent disease associated with decreased appetite, IGF-I is relatively low. This would imply that in such children a low serum IGF-I has a relatively low predictive value. If feasible, one may wish to first improve the nutritional condition before repeating further IGF-I measurements. In a child or adult with a high BMI SDS, GH secretion is usually low in contrast to a serum IGF-I in the upper half of the reference range. In overweight children, one could therefore expect that the optimal cut-off point of IGF-I for the decision to perform a GH stimulation test may be higher than in children with a normal BMI SDS. Because of the lack of observational data a formal adjustment of the IGF-I result for BMI is not possible, so that the clinician only can use his/her subjective clinical judgement.

When a low serum IGF-I is found in a child with a low clinical pre-test likelihood we suggest to repeat serum IGF-I, preferably in combination with a determination of serum IGFBP-3, before one decides to perform a GH stimulation test. If IGFBP-3 SDS is remarkably lower than IGF-I SDS, one should consider the possibility of a homozygous mutation of *IGFALS* (46).

#### **Calculation of the post-test likelihood of GHD based on serum IGF-I (with or without IGFBP-3) and individual pre-test likelihood of GHD**

In a previous study by our group (47), various parameters for quantifying the diagnostic value of serum IGF-I en IGFBP-3 and other markers were investigated in a Dutch cohort of children from 4 years of age with GHD or ISS. The optimal cut-off point for IGF-I for the diagnosis of GHD was -0.83 [with an Area Under the Curve (AUC) of 0.80 in the ROC analysis], but specificity was low at that point (47%). The optimal cut-off point for IGFBP-3 was -0.47 SDS (AUC 0.69) with a specificity of 22%. In the same publication data were shown on the frequencies of a serum IGF-I or IGFBP-3 SDS below  $<-2$ ,  $<-1$  and  $<0$  in children with either GHD or ISS, as well as the positive likelihood ratio (LR+) and negative likelihood ratio (LR-) (**Table 4**).

For IGF-I, the sensitivity (65% ) and specificity (78%) of a cut-off at -2 SDS in the Dutch study (47) were similar to those reported in other studies, as summarized in a meta-analysis (39), where average sensitivity and specificity of IGF-I were 0.66 and 0.69, with positive and negative likelihood ratios similar to those in the Dutch study. Also for IGFBP-3 similar findings were reported in the Dutch study (sensitivity 53%, specificity 81%) as in the later meta-analysis (50% and 79%). Noteworthy, a low IGFBP-3 is less sensitive but more specific than a low IGF-I. In the literature there is no consensus whether IGFBP-3 should be added to an IGF-I measurement in the screening phase (48). Because of the higher sensitivity of IGF-I (38, 39), and in order to reduce expenses, in the recent Dutch guideline we decided to limit IGFBP-3 determinations in this phase to children below 3 years of age, because in that age range the 3<sup>rd</sup> percentile of the reference range of IGF-I is close to the detection limit of most assays (40).

We also advised to add an IGFBP-3 determination when an IGF-I measurement is repeated, for example in children with a low IGF-I but a low pre-test likelihood of GHD. In such children, a low IGFBP-3 (because of its high specificity) considerably increases the likelihood of GHD (or a homozygous *IGFALS* defect). Also if a GH stimulation test is performed, we advise to repeat IGF-I and IGFBP-3 determinations at baseline.

A so far unexplored issue is which cut-off for serum IGF-I should be used for the decision to perform a GH stimulation test. Most papers suggest that IGF-I should be below -2 SDS, but it is obvious that this would imply that approximately 35% of children with GHD would not be detected. In one paper a cut-off of -1 SDS was suggested (49). However, according to general clinical epidemiological principles, the diagnostic value of a test depends on the clinical pre-test likelihood, which implies that different cut-off limits should be used for different ranges of pre-test likelihood. We therefore calculated the post-test likelihood depending on pre-test likelihood (in five arbitrary categories) and IGF-I result (<2, <-1 and <0 SDS) (Table 5). Based on this post-test likelihood, the clinician can make the decision if a GH stimulation test would be indicated to confirm or reject the diagnosis of GHD. We appreciate that the decision about which likelihood is sufficiently high to warrant GH stimulation tests is subjective. For the sake of argument we assumed that a post-test likelihood of >10% would be sufficient to decide to perform GH stimulation tests. The consequences of this analysis for different categories are described in the following paragraph. We emphasize that the IGF-I SDS that is used for decision-making should have been adjusted as well as possible for

confounding factors such as puberty and nutritional status, although we acknowledge that in contrast to the adjustment for pubertal stage, there are no numerical data that enable the recalculation of IGF-I SDS adjusted for nutritional status.

For each of the five categories of clinical pre-test likelihood a rational choice of cut-off for serum IGF-I can be made, as explained below. A schematic representation of this advice is shown in **Table 6**.

*High ( $\geq 50\%$ ) clinical pre-test likelihood:* In such children a GH stimulation test is indicated regardless of the result of IGF-I. If IGF-I is  $< 0$  SDS all post-test likelihoods are high ( $> 52-70\%$ ) and even in the rare cases with an IGF-I above 0 SDS (4% in the Dutch study (47)) the post-test likelihood will remain far above 10%. In fact, in children treated for brain tumours, for example using irradiation, or obese children, the IGF-I result is not a determining factor for performing a GH stimulation test (41).

*Moderate (20%) clinical pre-test likelihood:* At a cut-off of 0 SDS the post-test likelihood is approximately equal to the pre-test likelihood, and at a cut-off of -1 SDS it is 30%. We suggest that in such cases a serum IGF-I  $< 0$  SDS can be used for the decision to perform a GH stimulation test.

*Rather low (10%) clinical pre-test likelihood:* At a cut-off point of -1.0 SDS the post-test likelihood increases from 10 to 16%, which is in our opinion sufficient to perform a GH stimulation test. When IGF-I is between -1 and 0 SDS the post-test likelihood is not different from the pre-test likelihood and we assume that the clinician's subjective assessment will influence the decision to perform further testing.

*Low (5%) or very low ( $\leq 2\%$ ) clinical pre-test likelihood:* At a low pre-test likelihood the post-test likelihood is only above 10% at a cut-off point of -2.0 SDS, while at a very low likelihood the post-test likelihood remains below 10% at all IGF-I results. We suggest that at (very) low pre-test likelihood a rational first step would be to repeat an IGF-I determination, in combination with serum IGFBP-3. If this IGF-I result is in line with the first one and is supported by a low or low-normal IGFBP-3 SDS, further investigations are warranted to diagnose either GHD or rare syndromes characterized by low IGF-I and IGFBP-3. Such syndromes can be divided into two subgroups: those with normal sensitivity to GH [NSD, bio-inactive GH (Kowarski syndrome) or *GHSR* mutation] and those with low or absent sensitivity to GH (GH resistance, such as in children with mutations of *GHR*, *IGFALS*, *STAT5B*, *STAT3* and *IGF1*) (50). In order to differentiate between the two subgroups one can consider to perform an IGF-generation test, in spite of its imperfect diagnostic value (51). In a short child with an IGF-I  $< -2$  SDS and GH peak  $> 10$  ng/mL one can also consider to consult a

paediatric endocrinologist and/or clinical geneticist to discuss performing a growth-specific whole exome based gene panel (50, 52).

Obviously, one should realize that the estimate of the pre-test likelihood of GHD is subjective, and probably will vary between clinicians. The same applies to the cut-off point of post-test likelihood which is considered to be sufficiently high for performing a GH stimulation test.

### **Interpretation of a serum IGF-I >0 SDS in a short child**

A serum IGF-I >0 SDS can be seen in children with ISS, although mean IGF-I is approximately -1 SDS. If such IGF-I concentrations are found in a short child with low or low-normal birth size or low or low-normal head circumference, this is suggestive for a mutation or deletion of *IGF1R* (for a clinical score, see (53)). Also in children with Silver-Russell syndrome (caused by various (epi)genetic disorders, including *IGF2* and other gene mutations), a *PAPPA2* mutation, Bloom syndrome and *IGF1* mutations IGF-I can be increased or in the upper half of the reference range (54).

### **Conclusion**

In the child referred for short stature and/or growth faltering determination of serum IGF-I is a useful component of laboratory screening. The result should be expressed as SDS for age and sex and adjusted to physiologic (particularly pubertal stage and nutritional status) and analytic factors, and interpreted according to the clinical pre-test likelihood of GHD. The post-test likelihood can guide the clinician's decision to perform GH stimulation tests.

### **Statements**

#### **Acknowledgement**

We acknowledge the contributions of members of the Section of Paediatric Endocrinology of the Paediatric Association of the Netherlands to various versions of the Addendum to the Dutch Guideline on Triage and Diagnosis of Growth disorders in children which is at the base of this Commentary (B.Bakker, H.Claahsen-van de Grinten, S.E.Hannema, G.A.Kamp, K.Pijnenburg-Kleizen, H.M.van Santen, T.C.J.Sas, E.L.R.Van den Akker, P.Voorhoeve, M.J.E.Walenkamp). We also thank the clinical chemists dr. S.Kos-Sanders and dr. B.E.P.B. Ballieux for useful comments.

#### **Disclosure Statement**

JMW is member of scientific advisory boards of Merck, Lumos, Agios and Aeterna Zentaris and received Speaker's fees from Pfizer, Versartis, Sandoz, Lilly, Merck, JCR, Ipsen, Novo Nordisk and Ferring. MB is member of scientific advisory boards of Genescience, OPKO and Sandoz and has received speaker's fees from Diasorin, IDS, IPSEN, Novartis,

Pfizer and Roche. CdB received speaker's fees from Pfizer, NovoNordisk, Sandoz, Biomarin and Ferring. WO received Speaker's fee from Ferring BV.

### **Funding Sources**

The authors received no funding of any research relevant to their study.

### **Author Contributions**

JMW prepared subsequent versions of the manuscript. MB calculated IGF-I SDS positions for pubertal stages and reviewed and revised subsequent versions of the manuscript. CdB revised subsequent versions of the manuscript. WO chaired the working group of the Guideline and was involved in reviewing various versions of the manuscript. All authors agree with the submitted manuscript.

### **Ethics**

Ethics Committee Approval: not applicable

Informed Consent: not applicable

### **Authorship Contributions**

Surgical and Medical Practices: not applicable

Concept: JMW, WO

Design: JMW

Data Collection or Processing: MB and JMW

Analysis or Interpretation: JMW, MB, CdB, WO

Literature Search: JMW

Writing: JMW, WO, CdB, MB

**Conflict of Interest:** No conflict of interest

**Financial Disclosure:** No financial disclosure

JMW prepared subsequent versions of the manuscript. MB calculated IGF-I SDS positions for pubertal stages and reviewed and revised subsequent versions of the manuscript. CdB revised subsequent versions of the manuscript. WO chaired the working group of the Guideline and was involved in reviewing various versions of the manuscript. All authors agree with the submitted manuscript.

### **References**

1. Cohen P, Rogol AD, Deal CL, Saenger P, Reiter EO, Ross JL, et al. Consensus statement on the diagnosis and treatment of children with idiopathic short stature: a summary of the Growth Hormone Research Society, the Lawson Wilkins Pediatric Endocrine Society, and the European Society for Paediatric Endocrinology Workshop. *J Clin Endocrinol Metab.* 2008;93:4210-7.
2. Wit JM, Kamp GA, Oostdijk W, on behalf of the Dutch Working Group on T, Diagnosis of Growth Disorders in C. Towards a Rational and

Efficient Diagnostic Approach in Children Referred for Growth Failure to the General Paediatrician. *Horm Res Paediatr.* 2019;91:223-40.

3. Wit JM, Ranke MB, Kelnar CJH. ESPE Classification of Paediatric Endocrine Diagnoses. *Horm Res.* 2007;68(suppl 2):1-120.
5. Rosenfeld RG, Albertsson-Wikland K, Cassorla F, Frasier SD, Hasegawa Y, Hintz RL, et al. Diagnostic controversy: the diagnosis of childhood growth hormone deficiency revisited. *J Clin Endocrinol Metab.* 1995;80:1532-40.
6. Ranke MB, Schweizer R, Binder G. Basal characteristics and first year responses to human growth hormone (GH) vary according to diagnostic criteria in children with non-acquired GH deficiency (naGHD): observations from a single center over a period of five decades. *J Pediatr Endocrinol Metab.* 2018;31:1257-66.
7. Spiliotis BE, August GP, Hung W, Sonis W, Mendelson W, Bercu BB. Growth hormone neurosecretory dysfunction. A treatable cause of short stature. *JAMA.* 1984;251:2223-30.
8. Bercu BB, Diamond FB, Jr. Growth hormone neurosecretory dysfunction. *Clin. Endocrinol. Metab.* 1986;15:537-90.
9. Bercu BB, Shulman D, Root AW, Spiliotis BE. Growth hormone (GH) provocative testing frequently does not reflect endogenous GH secretion [published erratum appears in *J Clin Endocrinol Metab* 1987 Feb;64(2):382]. *J. Clin. Endocrinol. Metab.* 1986;63:709-16.
10. Darendeliler F, Spina I, Bas F, Bundak R, Isguven P, Arslanoglu I, et al. Reevaluation of growth hormone deficiency during and after growth hormone (GH) treatment: diagnostic value of GH tests and IGF-I and IGFBP-3 measurements. *J Pediatr Endocrinol. Metab.* 2004;17:1007-12.
11. Bizzarri C, Pedicelli S, Boscherini B, Bedogni G, Cappa M, Cianfarani S. Early retesting by GHRH + arginine test shows normal GH response in most children with idiopathic GH deficiency. *J Endocrinol Invest.* 2015;38:429-36.
12. Penta L, Cofini M, Lucchetti L, Zenzeri L, Leonardi A, Lanciotti L, et al. Growth Hormone (GH) Therapy During the Transition Period: Should We Think About Early Retesting in Patients with Idiopathic and Isolated GH Deficiency? *Int J Environ Res Public Health.* 2019;16.
13. Rekers-Mombarg LT, Cole TJ, Massa GG, Wit JM. Longitudinal analysis of growth in children with idiopathic short stature. *Ann. Hum Biol.* 1997;24:569-83.
14. Reinehr T, Hoffmann E, Rothermel J, Lehrian TJ, Binder G. Characteristic dynamics of height and weight in preschool boys with

constitutional delay of growth and puberty or hypogonadotropic hypogonadism. *Clin Endocrinol (Oxf)*. 2019.

15. 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.
16. Ranke MB, Wit JM. Growth hormone - past, present and future. *Nat Rev Endocrinol*. 2018;14:285-300.
17. Collett-Solberg PF, Jorge AAL, Boguszewski MCS, Miller BS, Choong CSY, Cohen P, et al. Growth hormone therapy in children; research and practice - A review. *Growth Horm IGF Res*. 2019;44:20-32.
18. 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*. 2019:1-14.
19. Vimpani GV, Vimpani AF, Lidgard GP, Cameron EH, Farquhar JW. Prevalence of severe growth hormone deficiency. *Br. Med. J*. 1977;2:427-30.
20. Rona RJT, J.M. Aetiology of idiopathic growth hormone deficiency in England and Wales. *Arch Dis Child*. 1977;52:197-208.
21. Lindsay R, Feldkamp M, Harris D, Robertson J, Rallison M. Utah Growth Study: growth standards and the prevalence of growth hormone deficiency. *J. Pediatr*. 1994;125:29-35.
22. Thomas M, Massa G, Craen M, de Zegher F, Bourguignon JP, Heinrichs C, et al. Prevalence and demographic features of childhood growth hormone deficiency in Belgium during the period 1986-2001. *Eur J Endocrinol*. 2004;151:67-72.
23. Schweizer R, Blumenstock G, Mangelsdorf K, Eehalt S, Rossner L, Dorn T, et al. Prevalence and incidence of endocrine disorders in children: results of a survey in Baden-Wuerttemberg and Bavaria (EndoPrIn BB) 2000-2001. *Klin Padiatr*. 2010;222:67-72.
24. Grote FK, Oostdijk W, De Muinck Keizer-Schrama SM, Van Dommelen P, Van Buuren S, Dekker FW, et al. The diagnostic work up of growth failure in secondary health care; an evaluation of consensus guidelines. *BMC. Pediatr*. 2008;8:21.
25. Stalman SE, Hellinga I, van DP, Hennekam RC, Saari A, Sankilampi U, et al. Application of the Dutch, Finnish and British Screening Guidelines in a Cohort of Children with Growth Failure. *Horm Res Paediatr*. 2015;84:376-82.

26. Stalman SE, Hellings I, Wit JM, Hennekam RC, Kamp GA, Plotz FB. Growth failure in adolescents: etiology, the role of pubertal timing and most useful criteria for diagnostic workup. *J Pediatr Endocrinol Metab.* 2016;29:465-73.
27. Shu SG, Chen YD, Chi CS. Clinical evaluation of short children referred by school screening: an analysis of 655 children. *Acta Paediatr Taiwan.* 2002;43:340-4.
28. Strufaldi MW, Silva EM, Puccini RF. Follow-up of children and adolescents with short stature: the importance of the growth rate. *Sao Paulo Med J.* 2005;123:128-33.
29. Bhadada SK, Bhansali A, Kochhar R, Menon AS, Sinha SK, Dutta P, et al. Does every short stature child need screening for celiac disease? *J Gastroenterol Hepatol.* 2008;23:e353-6.
30. Bhadada SK, Bhansali A, Ravikumar P, Kochhar R, Nain CK, Dutta P, et al. Changing scenario in aetiological profile of short stature in India-growing importance of celiac disease: a study from tertiary care centre. *Indian J Pediatr.* 2011;78:41-4.
31. Papadimitriou A, Douros K, Papadimitriou DT, Kleanthous K, Karapanou O, Fretzayas A. Characteristics of the short children referred to an academic paediatric endocrine clinic in Greece. *J Paediatr Child Health.* 2012;48:263-7.
32. Sisley S, Trujillo MV, Khoury J, Backeljauw P. Low Incidence of Pathology Detection and High Cost of Screening in the Evaluation of Asymptomatic Short Children. *J Pediatr.* 2013;163:1045-51.
33. Singh P, Sharma PK, Agnihotri A, Jyotsna VP, Das P, Gupta SD, et al. Coeliac disease in patients with short stature: A tertiary care centre experience. *Natl Med J India.* 2015;28:176-80.
34. Lashari SK, Korejo HB, Memon YM. To determine frequency of etiological factors in short statured patients presenting at an endocrine clinic of a tertiary care hospital. *Pak J Med Sci.* 2014;30:858-61.
35. Inoue-Lima TH, Vasques GA, Scalco RC, Nakaguma M, Mendonca BB, Arnhold IJP, et al. IGF-1 assessed by pubertal status has the best positive predictive power for GH deficiency diagnosis in peripubertal children. *J Pediatr Endocrinol Metab.* 2019;32:173-9.
36. Lodefalk M, Nilsson O. To Prime or Not to Prime - Is That Still a Question? A Comment on the US Guidelines on Growth Hormone and Insulin-Like Growth Factor-I Treatment in Children and Adolescents. *Horm Res Paediatr.* 2017;88:179-80.
37. Binder G, Weidenkeller M, Blumenstock G, Langkamp M, Weber K, Franz AR. Rational approach to the diagnosis of severe growth

hormone deficiency in the newborn. *J Clin Endocrinol Metab.* 2010;95:2219-26.

38. Sizonenko PC, Clayton PE, Cohen P, Hintz RL, Tanaka T, Laron Z. Diagnosis and management of growth hormone deficiency in childhood and adolescence. Part 1: diagnosis of growth hormone deficiency. *Growth Horm. IGF. Res.* 2001;11:137-65.

39. Shen Y, Zhang J, Zhao Y, Yan Y, Liu Y, Cai J. Diagnostic value of serum IGF-1 and IGFBP-3 in growth hormone deficiency: a systematic review with meta-analysis. *Eur J Pediatr.* 2015;174:419-27.

40. Blum WF, Alherbish A, Alsagheir A, El Awwa A, Kaplan W, Koledova E, et al. The growth hormone-insulin-like growth factor-I axis in the diagnosis and treatment of growth disorders. *Endocr Connect.* 2018;7:R212-R22.

41. Sklar CA, Antal Z, Chemaitilly W, Cohen LE, Follin C, Meacham LR, et al. Hypothalamic-Pituitary and Growth Disorders in Survivors of Childhood Cancer: An Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab.* 2018;103:2761-84.

42. Bidlingmaier M, Friedrich N, Emeny RT, Spranger J, Wolthers OD, Roswall J, 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;99:1712-21.

43. Lofqvist C, Andersson E, Gelander L, Rosberg S, Blum WF, Albertsson WK. Reference values for IGF-I throughout childhood and adolescence: a model that accounts simultaneously for the effect of gender, age, and puberty. *J Clin Endocrinol Metab.* 2001;86:5870-6.

44. Elmlinger MW, Kuhnel W, Weber MM, Ranke MB. Reference ranges for two automated chemiluminescent assays for serum insulin-like growth factor I (IGF-I) and IGF-binding protein 3 (IGFBP-3). *Clin Chem. Lab Med.* 2004;42:654-64.

45. Hawkes CP, Grimberg A. Insulin-Like Growth Factor-I is a Marker for the Nutritional State. *Pediatr Endocrinol Rev.* 2015;13:499-511.

46. Domene HM, Hwa V, Argente J, Wit JM, Camacho-Hubner C, Jasper HG, et al. Human acid-labile subunit deficiency: clinical, endocrine and metabolic consequences. *Horm Res.* 2009;72:129-41.

47. Rikken B, van Doorn J, Ringeling A, Van den Brande JL, Massa G, Wit JM. Plasma levels of insulin-like growth factor (IGF)-I, IGF-II and IGF-binding protein-3 in the evaluation of childhood growth hormone deficiency. *Horm. Res.* 1998;50:166-76.

48. Kayemba-Kay's S, Epstein S, Hindmarsh P, Burguet A, Ingrand P, Hankard R. Does plasma IGF-BP3 measurement contribute to the diagnosis of growth hormone deficiency in children? *Ann. Endocrinol (Paris)*. 2011;72:218-23.
49. Ranke MB, Schweizer R, Elmlinger MW, Weber K, Binder G, Schwarze CP, et al. Significance of basal IGF-I, IGFBP-3 and IGFBP-2 measurements in the diagnostics of short stature in children. *Horm Res*. 2000;54:60-8.
50. 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.
51. Coutant R, Dorr HG, Gleeson H, Argente J. Diagnosis of endocrine disease: limitations of the IGF1 generation test in children with short stature. *Eur. J Endocrinol*. 2012;166:351-7.
52. Dauber A, Rosenfeld RG, Hirschhorn JN. Genetic evaluation of short stature. *J Clin Endocrinol Metab*. 2014;99:3080-92.
53. Walenkamp MJE, Robers JML, Wit JM, Zandwijken GRJ, van Duyvenvoorde HA, Oostdijk W, et al. Phenotypic features and response to growth hormone treatment of patients with a molecular defect of the IGF-1 receptor. *J Clin Endocrinol Metab*. 2019.
54. Finken MJJ, van der Steen M, Smeets CCJ, Walenkamp MJE, de Bruin C, Hokken-Koelega ACS, et al. Children Born Small for Gestational Age: Differential Diagnosis, Molecular Genetic Evaluation, and Implications. *Endocr Rev*. 2018;39:851-94.

#### **Uncategorized References**

4. Quigley CAR, M.B. 2016;Pages. Accessed at Growth Analyser at <http://www.icped.org>.

**Table 1. Positive and negative clinical clues for GHD**

	<b>Positive clues</b>	<b>Negative clues</b>
<b>History</b>	<ul style="list-style-type: none"> <li>• fatigue</li> <li>• decrease of muscle power</li> <li>• decrease of concentration</li> <li>• symptoms of other pituitary deficiencies*</li> <li>• positive family history for GHD*</li> <li>• consanguinity</li> <li>• breech delivery</li> <li>• prolonged postnatal hypoglycaemia*</li> <li>• prolonged neonatal jaundice*</li> <li>• neurologic symptoms*</li> <li>• history of meningitis, cranial irradiation or brain injury*</li> </ul>	Symptoms compatible with alternative explanation of growth failure
<b>Physical examination</b>	<ul style="list-style-type: none"> <li>• lobulated abdominal fat</li> <li>• frontal bossing</li> <li>• high-pitched voice</li> <li>• cryptorchidism</li> <li>• neurologic signs*</li> <li>• vision abnormalities*</li> <li>• optic hypoplasia or other abnormalities at fundoscopy*</li> <li>• micropenis*</li> </ul>	<ul style="list-style-type: none"> <li>• dysmorphisms*</li> <li>• dysproportion*</li> <li>• underweight</li> <li>• any signs compatible with alternative explanation of growth failure</li> </ul>
<b>Growth</b>	<ul style="list-style-type: none"> <li>• growth deceleration*</li> <li>• absence of growth acceleration despite progression of puberty*</li> </ul>	no or little growth faltering*
<b>Laboratory screening</b>	<ul style="list-style-type: none"> <li>• Central hypothyroidism or deficiency of other pituitary axes*</li> <li>• central diabetes insipidus*</li> </ul>	Laboratory findings suggestive for an alternative explanation of growth failure (one of the secondary)

		growth disorders or Turner syndrome)*
<b>Radiograph of hand and wrist</b>	Bone age delay*	Bone age close to chronological age or advanced*

\*Diagnostic clues with relatively high weight in the opinion of the authors

**Table 2. Percentiles of IGF-I (ng/mL) according to sexual maturation, based on 854 samples (age 0-20 yrs), as measured with iSYS (42)\*.**

Tanner	Age	IGF-I, ng/mL, percentiles				
		2.5%	25%	50%	75%	97.5%
<b>males</b>						
<b>I</b>	6.1-12.9	81.3	132.5	160.0	187.9	255.3
<b>II</b>	8.1-14.8	106.2	212.4	276.9	331.8	432.3
<b>III</b>	10.9-16.0	244.9	341.2	407.2	449.0	511.4
<b>IV</b>	12.4-17.1	222.6	364.5	439.0	492.4	577.7
<b>V</b>	13.5-20.0	227.4	308.6	355.7	412.3	517.8
<b>females</b>						
<b>I</b>	5.8-12.1	85.9	152.6	187.7	235.3	323.0
<b>II</b>	9.3-14.1	117.5	190.0	247.3	323.2	451.3
<b>III</b>	9.3-15.1	258.3	335.5	382.8	430.8	528.5
<b>IV</b>	11.8-16.6	224.2	339.8	378.3	437.5	585.8
<b>V</b>	12.5-19.9	188.2	277.4	339.1	394.9	511.6

\*For results expressed in nmol/L, multiply by 0.131

**Table 3. Cut-off points at -2, -1, 0 +1 en +2 SDS of serum IGF-I (ng/mL) according to sexual maturation as measured with iSYS (Bidlingmaier, personal communication 2019)\***

		<b>IGF-I (ng/mL) SD and SDS positions based on Box-Cox transformation</b>					
<b>Tanner stage</b>	<b>age range</b>	<b>SD<sup>^</sup></b>	<b>-2.0 SDS</b>	<b>-1.0 SDS</b>	<b>0.0 SDS</b>	<b>+1.0 SDS</b>	<b>+2.0 SDS</b>
<b>boys</b>							
<b>I</b>	6.1-12.9	44.0	82.0	117.7	161.5	205.4	257.1
<b>II</b>	8.1-14.8	83.2	110.5	192.3	275.6	358.9	443.2
<b>III</b>	10.9-16.0	69.7	230.9	329.2	396.5	464.2	517.8
<b>IV</b>	12.4-17.1	90.1	211.1	340.8	426.7	513.6	582.2
<b>V</b>	13.5-20.0	74.7	226.9	288.1	362.5	437.0	525.5
<b>girls</b>							
<b>I</b>	5.8-12.1	60.0	88.5	134.0	193.64	253.3	327.1
<b>II</b>	9.3-14.1	90.3	101.6#	167.1	256.59	346.7	460.8
<b>III</b>	9.3-15.1	69.0	246.5	313.6	382.67	451.7	522.3
<b>IV</b>	11.8-16.6	86.1	235.2	304.8	390.25	475.9	577.9
<b>V</b>	12.5-19.9	82.1	184.7	256.6	338.74	420.9	512.3

\*For results expressed in nmol/L, multiply by 0.131

<sup>^</sup> Here, the untransformed SD is shown. For calculations of the various SDS cut-off points the Box-Cox transformation was used.

#This values is lower than the P2.5 obtained by the Harrell-Davis transformation (117.5 ng/mL) used in Bidlingmaier et al (42). For clinical practice, one could consider an IGF-I <117.5 in a girls with Tanner breast stage II as roughly equivalent to <-2.0 SDS.

**Table 4. Frequency of children with GHD or ISS with a serum IGF-I or IGFBP-3 <-2, <-1 en <0 SDS, and positive and negative likelihood ratios (LR+ en LR-) (47).**

<b>IGF-I SDS</b>	<b>&lt;-2</b>	<b>&lt;-1</b>	<b>&lt;0</b>	<b>IGFBP-3 SDS</b>	<b>&lt;-2</b>	<b>&lt;-1</b>	<b>&lt;0</b>
<b>GHD</b>	65%	86%	96%	<b>GHD</b>	53%	78%	96%
<b>ISS</b>	22%	50%	88%	<b>ISS</b>	19%	56%	88%
<b>LR+</b>	3.0	1.7	1.1	<b>LR+</b>	2.8	1.4	1.1
<b>LR-</b>	0.44	0.29	0.33	<b>LR-</b>	0.58	0.51	0.33

**Table 5. Post-test likelihood of GHD according to clinical pre-test likelihood and serum IGF-I\***

<b>Category of pre-test likelihood</b>	<b>IGF-I SDS cut-off</b>	<b>LR+</b>	<b>Pre-test likelihood</b>	<b>Pre-test odds</b>	<b>Post-test odds</b>	<b>Post-test likelihood</b>
<b>High (<math>\geq 50\%</math>)</b>	<0	1.1	50%	1	1.1	<b>52%</b>
	<-1	1.7		1	1.7	<b>63%</b>
	<-2	3.0		1	3.0	<b>70%</b>
<b>Moderate (20%)</b>	<0	1.1	20%	0.25	0.275	<b>21.5%</b>
	<-1	1.7		0.25	0.425	<b>30%</b>
	<-2	3.0		0.25	0.75	<b>43%</b>
<b>Rather low (10%)</b>	<0	1.1	10%	0.11	0.12	<b>11%</b>
	<-1	1.7		0.11	0.19	<b>16%</b>
	<-2	3.0		0.11	0.33	<b>25%</b>
<b>Low (5%)</b>	<0	1.1	5%	0.05	0.055	5%
	<-1	1.7		0.05	0.085	8%
	<-2	3.0		0.05	0.158	<b>14%</b>
<b>Very low (<math>\leq 2\%</math>)</b>	<0	1.1	2%	0.02	0.022	2%
	<-1	1.7		0.02	0.034	3.3%
	<-2	3.0		0.02	0.06	5.7%

\*The following formulas were used: Pre-test odds = pre-test likelihood/(1 minus pre-test likelihood). Post-test odds = Positive likelihood (LR+) x pre-test odds. Post-test likelihood = post-test odds/(post-test odds + 1). Percentages >10% are in bold print.

**Table 6. Suggested IGF-I cut-off limits in the screening phase to guide the decision to perform GH stimulation tests**

<b>Clinical pre-test likelihood of GHD</b>	<b>IGF-I SDS cut-off for GH stimulation tests</b>
High ( $\geq 50\%$ )	no
Moderate (20%)	<0 SDS
Rather low (10%)	<-1 SDS (or eventually <0 SDS)
Low (5%) or very low ( $\leq 2\%$ )	<-2 SDS; consider alternative diagnosis