Original Paper

HORMONE RESEARCH IN PÆDIATRICS

Horm Res Paediatr
DOI: 10.1159/000493133

Received: May 8, 2018
Accepted: August 20, 2018
Published online: October 4, 2018

Impact of the Choice of IGF-I Assay and Normative Dataset on the Diagnosis and Treatment of Growth Hormone Deficiency in Children

Maarten A.C. Broeren[a, b] Johannes G. Krabbe[a, c] Lianne S. Boesten[d]
Anita C.S. Hokken-Koelega[e, f] Yolanda B. de Rijke[a]

[a]Department of Clinical Chemistry, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands; [b]Department of Clinical Chemistry, Máxima Medical Center, Veldhoven-Eindhoven, The Netherlands; [c]Department of Clinical Chemistry and Laboratory Medicine, Medisch Spectrum Twente, Medlon BV, Enschede, The Netherlands; [d]Department of Clinical Chemistry, Ysselland Ziekenhuis, Capelle aan de IJssel, The Netherlands; [e]Dutch Growth Research Foundation, Rotterdam, The Netherlands; [f]Department of Pediatrics, Subdivision of Endocrinology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

Keywords
Insulin-like growth factor I · Growth hormone deficiency in children · *Z*-score · Normative dataset · Harmonization

Abstract
Background: The analysis of insulin-like growth factor I (IGF-I) is an important tool for pediatricians in the diagnosis and treatment of growth hormone deficiency in children. However, significant differences exist in IGF-I assays and normative datasets, which can have important clinical consequences. ***Methods:*** IGF-I analyses were performed using the IDS-iSYS platform on 1,897 samples from pediatric patients (0.5–18 years old). *Z*-scores were calculated based on normative IGF-I data from Bidlingmaier et al. (SD-BM) [J Clin Endocrinol Metab. 2014 May;99(5):1712–21] and normative IGF-I data from the IGF-I harmonization program in the Netherlands (SD-NL). The differences in *Z*-scores were analyzed at relevant clinical decision points (–2 SD, +2 SD). These normative datasets were also compared to normative data reported by Elmlinger et al. [Clin Chem Lab Med. 2004;42(6):654–64]. ***Results:*** The difference in *Z*-score between SD-BM and SD-NL was highest in males between 0 and 3 years old, exceeding 2 SD. Clinically relevant discordance between both *Z*-scores at –2 and +2 SD was found in 12.7% of all samples. The IGF-I levels at –2 and +2 SD reported in the normative dataset of Elmlinger et al. were up to 100% higher than the IGF-I levels reported by Bidlingmaier et al. or the Dutch harmonization program. ***Conclusion:*** Pediatricians and laboratory specialists should be aware of relevant differences that can exist between IGF-I assays and normative data. Well-defined pediatric reference ranges for the IDS-iSYS platform are highly desirable.

Introduction

Insulin-like growth factor I (IGF-I) is involved in many biochemical processes in the human body. It has its function in different steps of cell proliferation, differentiation, and apoptosis and has an insulin-like metabolic function [1]. Moreover, IGF-I is believed to play a role in brain function, aging, and longevity [2, 3]. IGF-I is highly dependent on growth hormone (GH) secretion and exerts most of its cellular effects by binding to the IGF-I

M.A.C. Broeren and J.G. Krabbe contributed equally.

E-Mail karger@karger.com
www.karger.com/hrp

Prof. Dr. Yolanda B. de Rijke
Department of Clinical Chemistry, Erasmus MC, University Medical Center Rotterdam
Wytemaweg 80
NL–3015 CN Rotterdam (The Netherlands)
E-Mail y.derijke@erasmusmc.nl

receptor. IGF-I analysis is used in the assessment of acromegaly and GH deficiency (GHD) as well as in the monitoring of GH therapy [4–6]. IGF-I serum levels are influenced by factors including age, puberty, nutritional status, liver function, and binding protein levels. Moreover, several IGF-I gene polymorphisms are related to (physiological) variations of serum levels of IGF-I [7].

For pediatricians, measurement of IGF-I is relevant for the investigation of the GH-IGF-I axis in children with short stature or decelerating growth. The use of *Z*-scores is helpful, since it is a standardized way of displaying the deviation of IGF-I concentrations from an age-related population. *Z*-scores lower than –2 SD strongly suggest an abnormality in the GH-IGF-I axis when other causes of low IGF-I values have been excluded, while *Z*-scores above 0 SD are likely to have normal GH provocative test results, making further evaluation of the GH-IGF-I axis unnecessary [6, 8, 9]. Still, normal values of IGF-I can occur in GHD. As a true gold standard for GHD is lacking, the diagnosis of GHD is challenging and requires an integrated assessment of all available data, including assessment of IGF-I levels [8, 10, 11]. Measurements of IGF-I levels are also used to monitor GH treatment. IGF-I levels generally increase upon increasing the GH dose, making IGF-I a useful biomarker for GH exposure. High IGF-I levels were found to be associated with increased cancer and cardiovascular disease mortality [12], and it is advised to lower the GH dose when IGF-I levels increase above a *Z*-score of +2 SD [6].

Reliable IGF-I assays and well-derived normative data are a prerequisite for successful use of IGF-I measurements in clinical practice. Still, it is recognized that significant differences can exist in both the results of IGF-I assays and in normative datasets [13–15]. Differences between the results of IGF-I assays are caused by several factors, including variability in assay design, the use of different calibration standards, interference from IGF-I-binding proteins, or differences in epitope specificity of the used antibodies [14]. As a consequence, reference values for IGF-I are usually assay-specific and can differ between laboratories [16]. Factors that can result in differences between normative data include number of subjects, ethnicity, nutritional status, preanalytical variability, or the statistical methodology used to convert IGF-I data to *Z*-scores [14, 17].

In the Endocrine Laboratory of Erasmus Medical Center, a change in IGF-I assay from IMMULITE (Siemens, Munich, Germany) to iSYS (Immunodiagnostic Systems Ltd., Tyne and Wear, UK) was implemented in 2013 due to a global supply disruption of IMMULITE reagent [18, 19]. A change in age- and sex-specific normative data was implemented simultaneously: the reference ranges determined by Elmlinger et al. [20] on the IMMULITE platform were replaced by the reference ranges for the iSYS assay as reported by Bidlingmaier et al. [21]. From that moment on, significantly higher *Z*-scores were reported in both pediatric and adult patient groups [22]. In the pediatric ward this change had relevant clinical consequences. Fewer children were diagnosed as GH deficient because they did not meet a required criterion (IGF-I *Z*-score below –2 SD). Moreover, the dosage of GH was reduced in a significant number of children as *Z*-scores more frequently exceeded +2 SD. The most pronounced differences were observed in children between 0 and 5 years and children between 10 and 18 years.

Age- and sex-specific reference ranges for IGF-I were also determined in the Netherlands in 2014 based on a national harmonization program, in analogy to the harmonization program for the analysis of GH [23]. A goal of this program was to decrease the variability in IGF-I results between immunoassays of different manufacturers by applying a correction factor to the IGF-I concentrations measured in each laboratory. The factor is obtained from analysis of a harmonization sample with a fixed IGF-I value. After correction of the IGF-I results in each laboratory, harmonized reference values can be applied for every assay.

We hypothesized that application of the Dutch harmonized reference values might improve the IGF-I *Z*-scores in pediatric patient samples, resulting in lower *Z*-scores than those obtained in the past using the IMMULITE IGF-I assay and concomitant reference values. In this study, we compared pediatric IGF-I *Z*-scores obtained from the assay-specific reference values published by Bidlingmaier et al. to the harmonized reference ranges from the Netherlands. These normative datasets were also compared to the previously used normative data reported by Elmlinger et al. The differences are discussed in light of the clinical consequences for the diagnosis and treatment of GHD in children.

Subjects and Methods

Subjects

We retrospectively reviewed the data of pediatric patients for whom IGF-I analysis was requested by their pediatrician from January 2014 to June 2015. Patients younger than 0.5 years or older than 18 years were excluded from analysis. IGF-I results below the detection limit of the assay were excluded from analysis. The ethnic background of patients was diverse, but the majority was Cau-

casian. IGF-I analysis was requested to evaluate a possible disorder in GH secretion or sensitivity and to monitor both GH and IGF-I replacement therapy. The participants remained anonymous during the database analysis.

Specimens and Assay

Serum was collected via capillary blood sampling or venipuncture. The samples were allowed to stand at room temperature for 30 min. They were then centrifuged at 2,163 *g* at 4 °C for 10 min. Analysis of IGF-I was performed within 8 h from blood collection. IGF-I was measured using a solid-phase, enzyme-labeled chemiluminescent immunometric assay (iSYS; Immunodiagnostic Systems Ltd.) according to the manufacturer's instructions. The assay is calibrated against the reference standard NIBSC 02/254. Within-laboratory, day-to-day precision was 9.7% at the level of 3.5 nmol/L and 4.2% at the level of 19.7 nmol/L.

Statistics and Normative Data

Z-scores for the measured IGF-I levels on the iSYS platform were based on age- and sex-specific normative data as reported by Bidlingmaier et al. These age- and sex-specific normative data were derived from 4,100 subjects (1,809 males, 2,291 females) from different cohorts with ages ranging from 0 to 18 years. The largest number of samples (n = 1,360) came from the Canadian Laboratory Initiative on Pediatric Reference Interval Database (CALIPER) conducted at the Hospital for Sick Children (Toronto, ON, Canada). This population was ethnically diverse and consisted of children attending dentistry, orthopedic, and plastic surgery clinics. The other cohorts included healthy children from Canada, Denmark, Sweden, Germany, Austria, and Greece. The samples were analyzed at six different sites using the iSYS platform. The recombinant NIBSC standard 02/254 was used for calibration [21].

In our study, *Z*-scores were calculated using the formula $Z = [(\text{IGF-I} / M)^{L} - 1] / (LxS)$ [21, 24, 25]. Age was calculated as [(date of blood withdrawal – date of birth) / 365.25]. IGF-I units were converted using the formula IGF-I (ng/mL) × 0.1307 = IGF-I (nmol/L). *Z*-scores based on the normative data of Bidlingmaier et al. were reported as "SD-BM."

Z-scores were also calculated from normative data of the Dutch harmonization program for IGF-I. The Dutch harmonization program for IGF-I has been implemented in analogy to the harmonization of GH measurements in the Netherlands, which was initiated to reduce between-laboratory variability of GH results at the clinical decision point of 20 mIU/L. In this program, a commutable serum pool was used as a consensus reference material. Each laboratory analyzed GH from this pool and calculated a harmonization factor f defined as $GH_{pool} / GH_{measured}$. Subsequently, "harmonized" GH results from patients were obtained by multiplication of GH results with the harmonization factor [23]. For IGF-I this methodology was also applied. The harmonization factor for IGF-I was defined as $IGF\text{-}I_{pool} / IGF\text{-}I_{measured}$, and subsequent IGF-I results were multiplied by the laboratory-specific harmonization factor. With a consensus value for $IGF\text{-}I_{pool}$ of 23.0 nmol/L, a harmonization factor of 1.11 was found from January 2014 to June 2015 on the iSYS platform. These harmonized IGF-I results were used in conjunction with harmonized normative data obtained in the Netherlands from 207 healthy subjects (106 males, 101 females) between 0 and 18 years old. The Dutch normative data were obtained using the Siemens IMMULITE 1000 platform standardized against the second WHO IS 87/518 and were harmonized according to the procedure described above. *Z*-scores were calculated using the mathematical LMS procedure as described above and were reported as "SD-NL." Differences in *Z*-scores between SD-BM and SD-NL were calculated for each sample and evaluated per age group (0.5–4.9 years, 5.0–9.9 years, and 10.0–18.0 years) and sex. Discordance between the *Z*-scores was evaluated at the clinical decision points of –2 SD and +2 SD. Discordance at –2 SD was defined as SD-BM ≤–2 SD while SD-NL was >–2 SD or vice versa; discordance at +2 SD was defined as SD-BM ≤+2 SD while SD-NL was >+2 SD or vice versa. Differences between the normative data of the Dutch harmonization program and the data reported by Bidlingmaier et al. and Elmlinger et al. were investigated by graphical comparison. The age- and sex-specific normative data from Elmlinger et al. were based on serum samples from 785 healthy subjects (382 males, 403 females) younger than 19 years. All samples were analyzed using the Siemens IMMULITE IGF-I assay standardized against the second WHO IS 87/518 and were analyzed at a single site [20]. Data were analyzed using Excel 2013 (Microsoft).

Results

In total, 1,897 samples (1,054 from males, 843 from females) were included for IGF-I analysis. The difference in *Z*-score between SD-BM and SD-NL is depicted in Figure 1. In males, SD-BM is lower than SD-NL for all ages. In males between 0 and 3 years, SD-BM can be more than 2 SD lower than SD-NL. With increasing age the difference between both *Z*-scores decreases slowly, but does not disappear. In females, the trend with increasing age seems similar to that in males, albeit with a smaller difference between both *Z*-scores. There seems to be no big difference between SD-BM and SD-NL in females 4–10 years old. Thereafter, SD-BM seems to be lower than SD-NL again.

Figure 2 shows the correlation between SD-BM and SD-NL in different age groups. The discordant results at –2 SD and +2 SD are depicted in blue. For males, a bias towards higher SD-NL values was observed in the age range of 0–4.9 years, resulting in a significant number of discordant results at –2 SD and +2 SD. A similar trend was observed at higher age, though less pronounced. For males, no discordance was found at +2 SD in the age range of 5.0–9.9 years. In females, a similar trend was observed, but the bias was smaller. In the age range of 5.0–9.9 years there seemed to be no bias. Overall, 130 results (6.9%) were discordant at –2 SD, most of which were males. At +2 SD, 110 results (5.8%) were discordant (Table 1).

A direct comparison of the normative data from Bidlingmaier et al. and the Dutch harmonization program is displayed in Figure 3. In males, the –2 SD line for

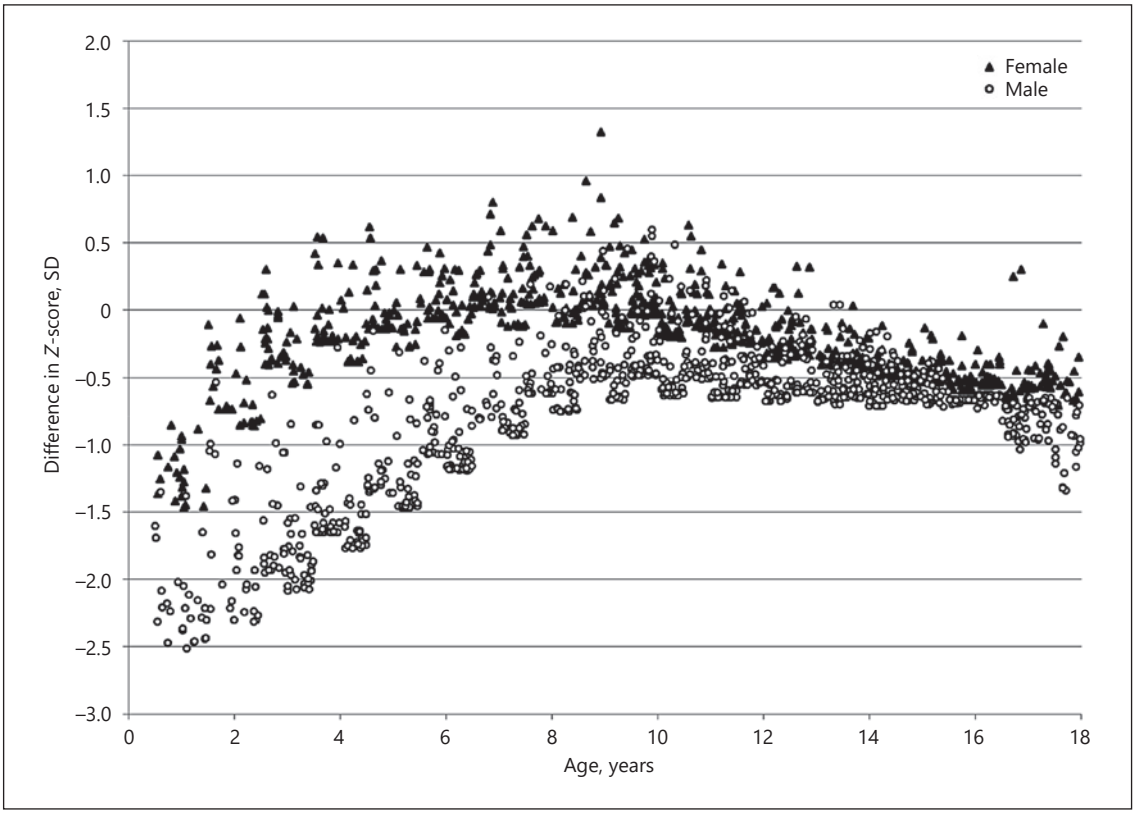

Fig. 1. Difference in *Z*-score (SD-BM minus SD-NL harmonization) against age for males (circles) and females (triangles).

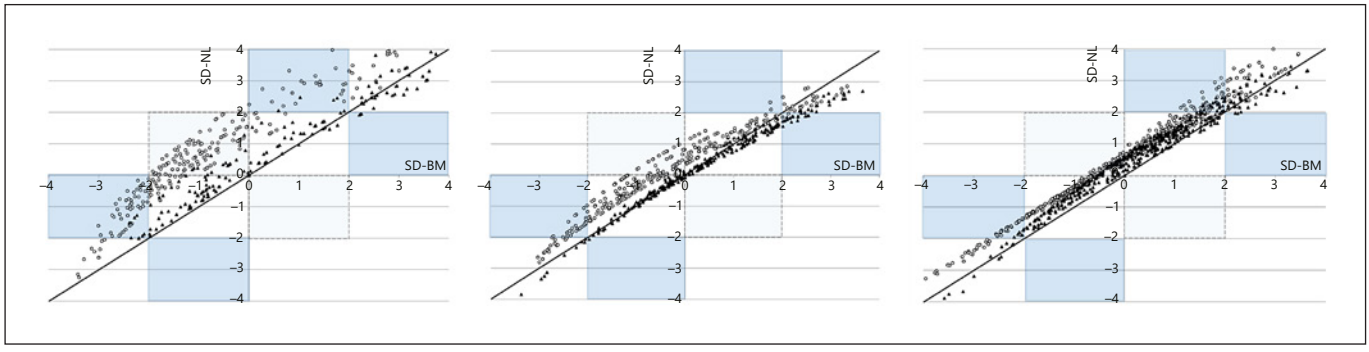

Fig. 2. SD-NL versus SD-BM in different age groups. Left: 0.5–4.9 years. Middle: 5.0–9.9 years. Right: 10–18 years. Males: circles; females: triangles. Black line: $y = x$. The discordant areas at –2 SD, 0 SD, and +2 SD are highlighted.

the Dutch cohort ran at lower IGF-I concentrations than the –2 SD line of the Bidlingmaier cohort over the complete age range. The maximum difference was 4 nmol/L. In females, the two –2 SD lines ran at virtually identical IGF-I concentrations. In males between 0 and 7 years, the +2 SD line of the Bidlingmaier reference values ran at higher IGF-I concentrations than the +2 SD line of the Dutch cohort, with a maximum difference of 12.5 nmol/L at year 1. For females, the difference between the +2 SD curves was smaller.

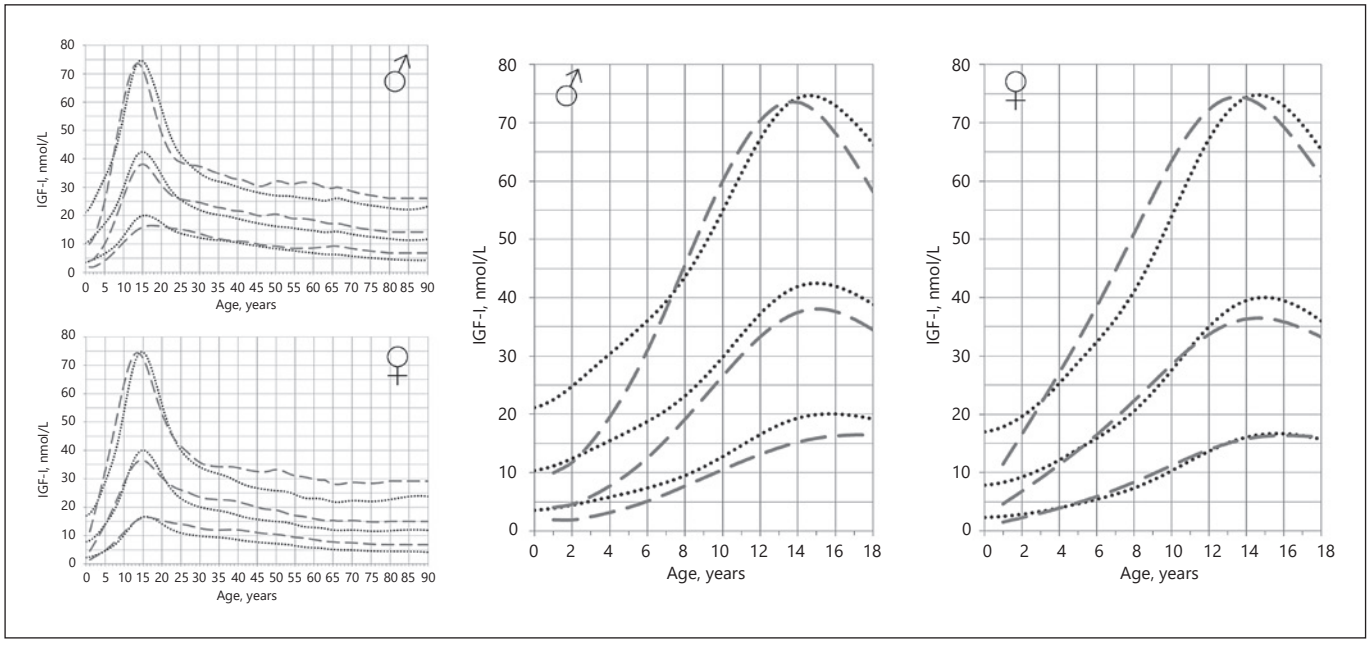

Fig. 3. Comparison of the Bidlingmaier (dotted lines) and Dutch harmonized (dashed lines) reference curves (–2 SD, mean, +2 SD) for IGF-I in males and females (left). The pediatric ranges are depicted in more detail (right).

Table 1. Overview of the number of discordant results of SD-NL versus SD-BM for males and females

Discordance	Males	Females	Total
<–2 SD	113 (10.7%)	17 (2.0%)	130 (6.9%)
–2 SD to 0 SD	209 (19.8%)	71 (8.4%)	280 (9.5%)
>+2 SD	58 (5.5%)	52 (6.2%)	110 (5.8%)
All	380 (36.1%)	140 (16.6%)	520 (27.4%)

Relative numbers were calculated based on the total amount of samples from males (n = 1,054), females (n = 843), or both (n = 1,897).

In Figure 4, the normative data from Bidlingmaier et al. and the Dutch harmonization program are compared to the normative data from Elmlinger et al., which were used in conjunction with the IMMULITE IGF-I assay before the introduction of the iSYS IGF-I assay. In both males and females, the normative IGF-I results from Elmlinger et al. were significantly higher than the normative results from Bidlingmaier et al. and the Dutch harmonization program at 0–6 and 12–18 years.

Discussion

In this study, we investigated the difference in Z-scores obtained with the Bidlingmaier reference values (SD-BM) and the Dutch harmonized reference values (SD-NL) for IGF-I in children. This comparison was initiated by the observation that a switch in IGF-I assay from IMMULITE to iSYS and concomitant normative data in our laboratory resulted in a substantial increase in IGF-I Z-scores in children that did not seem to be in agreement with other clinical findings. Similar trends were found in adults [22]. We hypothesized that application of the harmonized reference values in the Netherlands might result in a decrease in Z-scores, comparable to the earlier assay and normative data of Elmlinger et al. Interestingly, the opposite appeared to be the case (Fig. 1).

The Z-scores of males were higher for the harmonized reference values than for the reference values of Bidlingmaier et al. over the whole age range, with the biggest difference between 0 and 6 years. For girls the difference was smaller. A direct comparison of Z-scores is useful as it can give an indication whether differences are clinically relevant. The comparison of Z-scores depicted in Figure 2 again shows that the majority of samples had a higher Z-score for SD-NL, but also indicates that a sub-

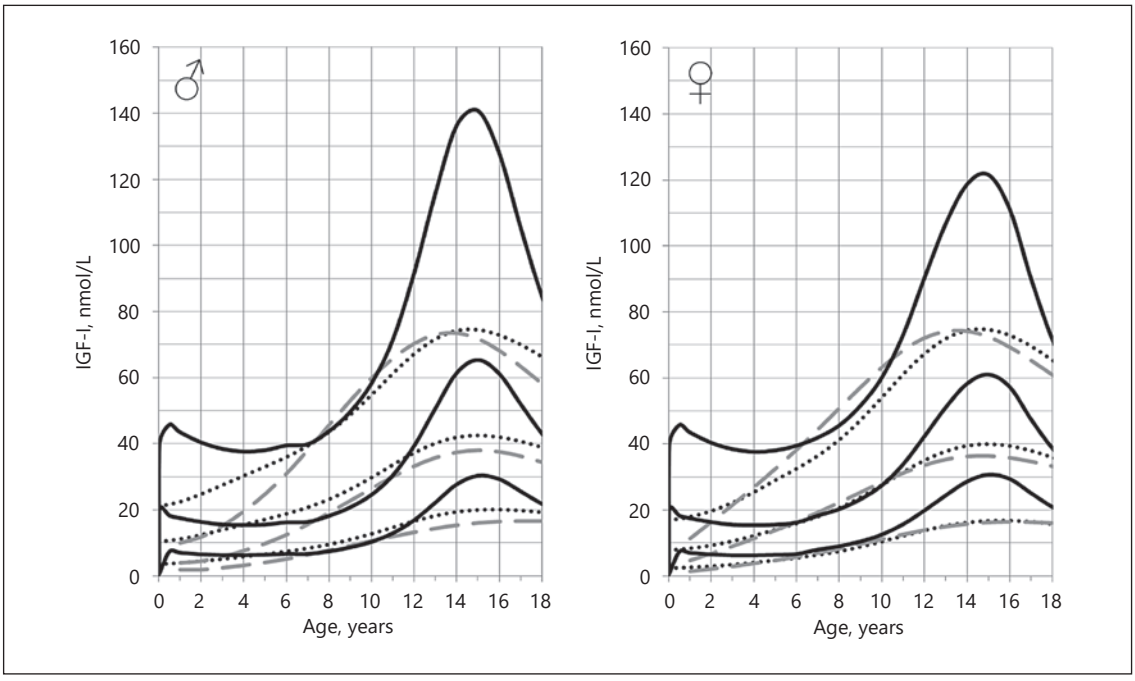

Fig. 4. Comparison of the Elmlinger reference curves (solid lines) for IGF-I to the Bidlingmaier (dotted lines) and Dutch harmonized reference curves (dashed lines) in males (left) and females (right).

stantial group of samples was abnormal in only one of the two normative datasets (Table 1). These discordant results are most likely clinically relevant. At the clinical decision point –2 SD, it is implied that a diagnosis of GHD is made less frequently when the harmonized reference values are used. As no gold standard is present, the diagnosis of GHD is difficult and should be based on an integrated assessment that can include auxology, clinical judgement, provocative GH testing, neuroimaging, or genetic testing next to IGF-I measurements. Still, IGF-I measurements play an important role in this assessment, especially when the overall picture is not clear-cut. For example, in the Netherlands an IGF-I *Z*-score below –2 SD is a prerequisite for starting GH treatment when a partial response in provocative GH testing is found. Additional discordant results appear when the range between –2 and 0 SD is taken into account. Generally, provocative GH testing is only indicated when IGF-I *Z*-scores are <0 SD and other causes of low IGF-I values have been excluded. Here, there is a large group of samples with SD-NL above 0 SD while SD-BM is below 0 SD (Table 1; Fig. 2). For these patients, provocative GH testing is not considered when the harmonized reference ranges are applied, while it is not rejected when the Bidlingmaier reference ranges are applied. At +2 SD, differences between normative data are relevant during GH treatment. GH treatment is monitored by means of IGF-I measurements, and dose reduction is advised when IGF-I *Z*-scores exceed +2 SD [6]. Based upon the discordant results at +2 SD, dose reduction of GH is required less frequently when the Bidlingmaier reference ranges are applied. Importantly, we can only conclude that clinically relevant differences are present between both normative datasets. It is not possible to conclude which of the normative datasets are better in clinical practice, as we only compared IGF-I *Z*-scores and did not take into account other parameters that are relevant for the diagnosis and treatment of GHD.

The most important reason for the observed differences in *Z*-scores between SD-NL and SD-BM is the difference in normative datasets (Fig. 3). Still, it is important to realize that the harmonization factor of 1.11 induces an additional difference in *Z*-scores. At –2 SD for example, the harmonization factor slightly increases the difference between SD-NL and SD-BM in males, as IGF-I results are multiplied by 1.11 before calculation of SD-NL, while this is not done for calculation of SD-BM. The harmonization factor causes the slight difference between SD-NL and SD-BM at –2 SD for females, as the normative datasets are virtually identical for females at –2 SD.

The observed differences between SD-NL and SD-BM are relevant, but smaller than the differences that were observed when the switch in IGF-I assay in our laboratory was implemented in 2013. The change in IGF-I assay from IMMULITE to iSYS was accompanied by a change from normative data reported by Elmlinger et al. to normative data of Bidlingmaier et al. The normative data from Elmlinger et al. are very different from the normative data of Bidlingmaier et al., especially at 0–5 years and at 12–18 years (Fig. 4). At 0–5 years, IGF-I values at –2 SD and +2 SD can be up to 100% higher for the normative dataset of Elmlinger et al. Also, IGF-I levels at the pubertal peak are 50–100% higher for the Elmlinger dataset. These differences are only slightly reduced by the difference in immunoassay. It is known that the IMMULITE IGF-I assay reports higher IGF-I values than the iSYS assay. Bidlingmaier et al. compared both assays and found that IGF-I concentrations were approximately 20% higher on the IMMULITE assay [21]. Chanson et al. [16] derived normative data for both assays in adults and found that IGF-I levels at +2 SD are 25–35% higher for the IMMULITE assay than for the iSYS assay. Interestingly, they observed a much smaller difference between both assays at –2 SD, implying that the bias might be nonlinear and dependent on the magnitude of IGF-I. Overall, these results demonstrate that a switch from IMMULITE to iSYS, together with a switch in normative data from Elmlinger et al. to Bidlingmaier et al., was accompanied by a substantial increase in IGF-I *Z*-score in pediatric males and females, especially between 0 and 5 years and around puberty.

Reliable IGF-I assays and well-derived normative data are crucial for the diagnosis and treatment of GHD in children. Still, significant differences in current commercial IGF-I assays are present, even when manufacturers adhere to the recommendations of the Consensus Group on Standardization and Evaluation of GH and IGF-I assays [13]. For example, Chanson et al. [16] found a moderate overall agreement between commercial IGF-I assays based on weighted kappa coefficients, while most assays were calibrated against the most recent IS 02/254 WHO reference standard. These differences between assays are not necessarily problematic and can be dealt with by using well-derived assay-specific reference ranges. Unfortunately, IGF-I reference ranges are very difficult to establish due to the large number of subjects required as well as many other factors that influence IGF-I concentrations, such as ethnicity, nutritional status, body mass index, or medication. As a consequence, heterogeneity in the study population can significantly influence normative IGF-I levels. Varewijck et al. [17] compared IGF-I *Z*-scores based on normative data of the VARIETE cohort and normative data of Bidlingmaier et al. in GH-deficient adults. Both *Z*-scores were derived from the same iSYS IGF-I assay. Interestingly, the mean *Z*-score using the VARIETE cohort was 1 SD lower than the mean *Z*-score based on the Bidlingmaier cohort. Most likely, the difference in selection criteria for subject inclusion in both cohorts is an important cause of the observed disagreement. Participants in the VARIETE cohort were healthy French adult volunteers, selected by strict inclusion and exclusion criteria, while subjects in the Bidlingmaier cohort were included based on different inclusion criteria from 12 different cohorts from Europe, Canada, and the United States. Considering the clinical consequences of the observed differences, a thorough selection of subjects for the normative dataset seems pivotal.

Also for the pediatric reference ranges reported by Bidlingmaier et al., different cohorts were used that applied different inclusion criteria and were reported as ethnically diverse [21]. This might cause heterogeneity in the obtained reference ranges, and it can be questioned whether these reference ranges can be extrapolated to the Dutch population. This can only be answered by comparison to well-derived pediatric normative data based on a random selection of individuals of the background population determined with the iSYS assay. Unfortunately, pediatric reference ranges for IGF-I are even more difficult to establish than adult reference ranges due to the strong variation in IGF-I concentrations during childhood and puberty.

Harmonization can be a tool to reduce differences in test results for laboratory measurements that lack a reference measurement procedure [26, 27]. For example, harmonization of GH measurements has proven to reduce between-laboratory variation and to improve comparability in the confirmation of GHD by provocative GH testing [23, 28]. However, it is important to emphasize that the clinical decision point in provocative GH testing is a fixed, single cutoff. Harmonization of IGF-I results is much more complicated as clinical decision points vary significantly for age and sex and are present at both –2 SD and +2 SD. Considering the challenges that accompany the harmonization procedure of laboratory measurements, including analytical specificity for the analyte and commutability of reference material [26], it is doubtful whether harmonization of IGF-I results based on a single multiplication factor is appropriate for the diagnosis and treatment of GHD in children. To the best of our knowledge, there is no study that has demonstrated the applicability of harmonized IGF-I results in clinical practice.

We cannot conclude which of the two normative datasets is most applicable in clinical practice. In our opinion, both normative datasets have drawbacks that hamper optimal use of IGF-I measurements in the diagnosis and treatment of GHD in children. Well-derived pediatric normative data for the iSYS IGF-I assay are lacking and are, considering the clinical relevance, highly desirable. Analysis of IGF-I with mass spectrometry can circumvent some of the problems that are inherent to immunoassays, e.g., interference of IGF-binding proteins or cross-reactivity of IGF-II [29, 30]. Mass spectrometry might be a promising technique to reduce variability in IGF-I analysis and to improve standardization across platforms, provided that calibration procedures are harmonized [31]. The development of a robust IGF-I reference method is clinically relevant and should be the focus of future research.

In conclusion, we compared IGF-I *Z*-scores in pediatric patient samples using normative datasets from Bidlingmaier et al. and the IGF-I harmonization program in the Netherlands. In a substantial group of samples, clinically relevant differences were observed for the diagnosis and treatment of GHD. Both normative IGF-I datasets for the iSYS assay are very different from the normative dataset reported by Elmlinger et al. for the IMMULITE IGF-I assay, which is not explained by differences between both assays. Pediatricians and laboratory specialists should be aware of relevant differences that can exist between IGF-I assays and normative data.

Statement of Ethics

This study was approved by the Medical Ethics Committee of the Erasmus Medical Center (Rotterdam, The Netherlands).

Disclosure Statement

The authors have no conflicts of interest to disclose.

References

1 Jones JI, Clemmons DR. Insulin-like growth factors and their binding proteins: biological actions. Endocr Rev. 1995 Feb;16(1):3–34.
2 Russo VC, Gluckman PD, Feldman EL, Werther GA. The insulin-like growth factor system and its pleiotropic functions in brain. Endocr Rev. 2005 Dec;26(7):916–43.
3 Junnila RK, List EO, Berryman DE, Murrey JW, Kopchick JJ. The GH/IGF-1 axis in ageing and longevity. Nat Rev Endocrinol. 2013 Jun;9(6):366–76.
4 Katznelson L, Laws ER Jr, Melmed S, Molitch ME, Murad MH, Utz A, et al.; Endocrine Society. Acromegaly: an endocrine society clinical practice guideline. J Clin Endocrinol Metab. 2014 Nov;99(11):3933–51.
5 Molitch ME, Clemmons DR, Malozowski S, Merriam GR, Vance ML; Endocrine Society. Evaluation and treatment of adult growth hormone deficiency: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2011 Jun;96(6):1587–609.
6 Grimberg A, DiVall SA, Polychronakos C, Allen DB, Cohen LE, Quintos JB, et al.; Drug and Therapeutics Committee and Ethics Committee of the Pediatric Endocrine Society. 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):361–97.
7 Jung SY, Rohan T, Strickler H, Bea J, Zhang ZF, Ho G, et al. Genetic variants and traits related to insulin-like growth factor-I and insulin resistance and their interaction with lifestyles on postmenopausal colorectal cancer risk. PLoS One. 2017 Oct;12(10):e0186296.
8 Growth Hormone Research Society; GH Research Society. Consensus guidelines for the diagnosis and treatment of growth hormone (GH) deficiency in childhood and adolescence: summary statement of the GH Research Society. J Clin Endocrinol Metab. 2000 Nov;85(11):3990–3.
9 Cohen P, Rogol AD, Deal CL, Saenger P, Reiter EO, Ross JL, et al.; 2007 ISS Consensus Workshop participants. 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 Nov;93(11):4210–7.
10 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 May;80(5):1532–40.
11 Stanley T. Diagnosis of growth hormone deficiency in childhood. Curr Opin Endocrinol Diabetes Obes. 2012 Feb;19(1):47–52.
12 Burgers AM, Biermasz NR, Schoones JW, Pereira AM, Renehan AG, Zwahlen M, et al. Meta-analysis and dose-response metaregression: circulating insulin-like growth factor I (IGF-I) and mortality. J Clin Endocrinol Metab. 2011 Sep;96(9):2912–20.
13 Clemmons DR. Consensus statement on the standardization and evaluation of growth hormone and insulin-like growth factor assays. Clin Chem. 2011 Apr;57(4):555–9.
14 Frystyk J, Freda P, Clemmons DR. The current status of IGF-I assays – a 2009 update. Growth Horm IGF Res. 2010 Feb;20(1):8–18.
15 Hawkes CP, Grimberg A. Measuring growth hormone and insulin-like growth factor-I in infants: what is normal? Pediatr Endocrinol Rev. 2013 Dec;11(2):126–46.
16 Chanson P, Arnoux A, Mavromati M, Brailly-Tabard S, Massart C, Young J, et al.; VARIETE Investigators. Reference Values for IGF-I Serum Concentrations: Comparison of Six Immunoassays. J Clin Endocrinol Metab. 2016 Sep;101(9):3450–8.
17 Varewijck AJ, van der Lely AJ, Neggers SJCMM, Hofland LJ, Janssen JAMJL: Disagreement in normative IGF-I levels may lead to different clinical interpretations and GH dose adjustments in GH deficiency. Clin Endocrinol (Oxf). 2018 Mar;88(3):409–14.

18 Siemens Healthcare Diagnostics. IMMULITE®, IMMULITE® 1000, IMMULITE® 2000, IMMULITE® 2000 XPi: All IMMULITE platforms for IGF-I shift in patient medians and supply disruption. Urgent field safety notice No. 4005. http://webarchive.nationalarchives.gov.uk/20150113101428/http://www.mhra.gov.uk/home/groups/fsn/documents/fieldsafetynotice/con207161.pdf.
19 Algeciras-Schimnich A, Bruns DE, Boyd JC, Bryant SC, La Fortune KA, Grebe SK. Failure of current laboratory protocols to detect lot-to-lot reagent differences: findings and possible solutions. Clin Chem. 2013 Aug;59(8):1187–94.
20 Elmlinger MW, Kühnel W, Weber MM, Ranke MB. Reference ranges for two automated chemiluminescent assays for serum insulin-like growth factor I (IGF-I) and IGF-binding protein 3 (IGFBP-3). Clin Chem Lab Med. 2004;42(6):654–64.
21 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 May;99(5):1712–21.
22 Varewijck AJ, Lamberts SW, van der Lely AJ, Neggers SJ, Hofland LJ, Janssen JA. The introduction of the IDS-iSYS total IGF-1 assay may have far-reaching consequences for diagnosis and treatment of GH deficiency. J Clin Endocrinol Metab. 2015 Jan;100(1):309–16.
23 Ross HA, Lentjes EW, Menheere PM, Sweep CG; Endocrinology Section and project group "Calibration 2000" of the SKML (Dutch Foundation for Quality Assessment in Clinical Laboratories). Harmonization of growth hormone measurement results: the empirical approach. Clin Chim Acta. 2014 May;432:72–6.
24 Cole TJ, Green PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. Stat Med. 1992 Jul;11(10):1305–19.
25 Yee TW. Quantile regression via vector generalized additive models. Stat Med. 2004 Jul;23(14):2295–315.
26 Greg Miller W, Myers GL, Lou Gantzer M, Kahn SE, Schönbrunner ER, Thienpont LM, et al. Roadmap for harmonization of clinical laboratory measurement procedures. Clin Chem. 2011 Aug;57(8):1108–17.
27 Van Uytfanghe K, De Grande LA, Thienpont LM. A "Step-Up" approach for harmonization. Clin Chim Acta. 2014 May;432:62–7.
28 Müller A, Scholz M, Blankenstein O, Binder G, Pfäffle R, Körner A, et al. Harmonization of growth hormone measurements with different immunoassays by data adjustment. Clin Chem Lab Med. 2011 Jul;49(7):1135–42.
29 Hines J, Milosevic D, Ketha H, Taylor R, Algeciras-Schimnich A, Grebe SK, et al. Detection of IGF-1 protein variants by use of LC-MS with high-resolution accurate mass in routine clinical analysis. Clin Chem. 2015 Jul;61(7):990–1.
30 Hawkes CP, Murray DM, Kenny LC, Kiely M, O'B Hourihane J, Irvine AD, et al. Correlation of Insulin-Like Growth Factor-I and -II Concentrations at Birth Measured by Mass Spectrometry and Growth from Birth to Two Months. Horm Res Paediatr. 2018;89(2):122–31.
31 Cox HD, Lopes F, Woldemariam GA, Becker JO, Parkin MC, Thomas A, et al. Interlaboratory agreement of insulin-like growth factor 1 concentrations measured by mass spectrometry. Clin Chem. 2014 Mar;60(3):541–8.