Clinical evidence-based cutoff limits for GH stimulation tests in children with a backup of results with reference to mass spectrometry

I V Wagner, C Paetzold¹, R Gausche², M Vogel², A Koerner, J Thiery¹, C G Arsene³, A Henrion³, B Guettler³, E Keller, W Kiess, R Pfaeffle* and J Kratzsch^{1,*}

Centre for Paediatric Research Leipzig (CPL), Hospital for Children and Adolescents, University of Leipzig, Leipzig, Germany, ¹Institute for Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, University of Leipzig, Paul-List-Straße 13-15, 04103 Leipzig, Germany, ²CrescNet GmbH, University of Leipzig, Leipzig, Germany and ³Physikalisch-Technische Bundesanstalt (PTB), Braunschweig, Germany
*(R Pfaeffle and J Kratzsch contributed equally to this work)

Correspondence
should be addressed
to J Kratzsch
Email
juergen.kratzsch@
medizin.uni-leipzig.de

Abstract

Context: Cutoff limits of GH stimulation tests to diagnose GH deficiency (GHD) in children and adolescents are not sufficiently validated by clinical studies due to discrepancies in the performance of GH immunoassays and lack of available study populations.

Objective: We aimed to establish new cutoff limits for GH stimulation tests based on clinical evidence and compared these immunoassay-based values with an antibody-independent mass spectrometric method.

Design and setting: In a retrospective study, GH cutoff limits for eight different immunoassays and isotope dilution mass spectrometry (ID-MS) were calculated from hGH peak concentrations of short-statured children with and without GHD.

Patients: We compared the serum GH peak concentrations at GH stimulation test of 52 short-statured children and adolescents, who have normal GH secretion at initial workup and normal growth in the follow-up, with the serum GH peak concentrations of 44 GHD patients in the same age range, in order to optimize the cutoff limit calculation.

Results: Discriminant analysis of re-measured GH led to a new cutoff limit of 7.09 µg/l using the iSYS assay (IDS) and the limits for the other seven hGH assays varied between 4.32 and 7.77 µg/l. For ID-MS, cutoffs of 5.48 µg/l (22k GH) and 7.43 µg/l (total GH) were ascertained.

Conclusion: The establishment of method-specific clinical evidence-based GH cutoff limits is of importance to ensure adequate clinical diagnosis and treatment of children and adolescents with GHD. ID-MS may become an important tool for providing both reliable and sustainable SI traceability of GH measurements in the future.

European Journal of
Endocrinology
(2014) 171, 389–397

Introduction

The prevalence of growth hormone deficiency (GHD) is estimated to be between 1:4000 and 1:20 000 (1). However, the number of children with short stature visiting clinics of pediatric endocrinology is much higher, although definite numbers do not exist. A correct diagnosis is essential in the evaluation of patients with short stature for suspected GHD (2). A false-positively diagnosed patient will be

subjected unnecessarily to a prolonged, invasive, ineffective, and expensive treatment that might even induce negative side effects. On the other hand, a missed diagnosis will deprive the patient of treatment, resulting in an unfavorable clinical outcome. The first stimulation tests to explore the GH reserve in children were performed during the late 1960s and early 1970s (3, 4, 5, 6, 7). There is a

consensus that a unique cutoff value should be used for various stimulation tests used for laboratory testing of GHD in children so far (e.g. hypoglycemia, arginine, glucagon, etc.) (3, 5, 8, 9). The diagnosis of GHD, which necessitates GH replacement therapy, is confirmed by two GH stimulation tests defined as the peak serum GH concentration below a cutoff value. Current GH cutoff levels for children and adolescents are more or less arbitrarily determined (10, 11, 12). A maximum of 8 µg/l GH has been suggested as a cutoff limit in Germany (13). In addition, inconsistencies within the results obtained using various analytical tools and laboratories are hampering the reliability of this decision (14, 15).

Besides the limited number of clinical studies to support the scientific basis of this value, there are several factors of variation at present that contribute more or less to the variability of GH results, and thus limiting the diagnostic power of the stimulation test. The factors are as follows:

- i) design of the GH immunoassay, preparation of the assay calibrators, specificity of antibodies, matrix differences between calibrators and samples, and interference with endogenous human GH-binding protein (GHBP) (15, 16) and
- ii) factors modifying the GH status such as age, nutritional status (lean body mass and obesity), thyroid hormone function, glucocorticoid levels, pubertal development, and fasting state (2, 3, 15, 17).

Therefore, this article primarily aims to provide new clinical evidence-based cutoff limits obtained using the iSYS assay. iSYS is a recently developed assay system that fulfils current consensus recommendations and presents a useful new tool for reliable measurement of GH (15, 18). Furthermore, this evidence-based cutoff limit is translated into cutoff values for other commercially available immunoassays, using an established linear relationship of results between the iSYS and the other GH assays (19).

Finally, the iSYS assay is compared with isotope dilution mass spectrometry (ID-MS). The measurement of GH by ID-MS has been recently introduced (20, 21) as an alternative to antibody-based methods, and it is exempt from the above-mentioned susceptibility to the interferences by non-target proteins. At the same time, the principle of the method used in ID-MS provides the traceability to the International System of Units (SI), obviating the need to refer to a particular GH-RP (22). The expression of the cutoff values in terms of ID-MS GH concentrations makes them independent of a particular combination of antibody and RP to be used.

Subjects and methods

The study was reviewed and approved by the Ethics Committee of the University of Leipzig (vote number 781). Serum samples of patients from the Department of Pediatric Endocrinology at the University Hospital Leipzig were collected from 2000 to 2011 and the informed consents were obtained. Patients were investigated for the diagnosis of suspected GHD according to the guidelines of the German Society of Pediatrics (13).

Subjects

Clinical data from 359 children and adolescents suspected of suffering from GHD have been traced over a period of at least 12 years in a local computerized auxological database (CrescNet) (23). These patients had to meet the following inclusion criteria: a pathological growth velocity (<25th percentile), a height of <−1.5 SDS or a height of more than 1.5 SDS deviation from their parental target height at initial workup, and growth in the normal or near normal range (Fig. 1 and Table 1). All patients underwent at least one arginine test (0.5 g L-arginine). Those who underwent the second confirmatory test received either glucagon (~60%, 50 mg/kg glucagon) or insulin (~40%, 0.05–0.1 IE/kg insulin). If the maximal GH concentration of both tests was below the cutoff values, the patient was diagnosed with GHD. According to our internal and German guidelines, all girls >11 years or boys >13 years without signs of pubertal development have been primed with sex hormones (either estrogen or testosterone) (13). Patients were excluded when they were suffering from syndromal disease or chronic illness, when the final height was <−1.50 SDS, or when they had been classified as small for gestational age (SGA) at birth. The eligible patients were divided into two groups described below and their sera were further tested in our study.

Cohort of non-GHD children

Of the total number of patients, 109 short children and adolescents were not treated with GH. Subjects were included in the non-GHD group under the condition that GHD had been ruled out at the time of initial clinical workup by the stimulation test and that the normal height was obtained later on based on their growth data retrospectively. After excluding patients suffering from chronic or syndromal diseases such as Turner's syndrome, Noonan syndrome, etc., and those previously classified as SGA, 52 children were available for this cohort of non-GHD children (Fig. 1 and Table 1).

Figure 1

Flow chart of the study cohort. A total of 359 patients were traceable in our CrescNet database that underwent 688 stimulation tests. The flow chart represents the way how our final study cohort was selected.

Stimulation tests were carried out at a mean age of 9.13 ± 3.73 years with a height SDS of -2.18 ± 0.39 and the deviation from parental target height of -0.91 ± 0.72 SDS (mean $\pm$ s.d.). All patients showed a GH concentration higher than $7.63 \mu\text{g/l}$ in the stimulation tests at their diagnostic workups. After an average observation period of 3.77 years, the mean height SDS of this group was at the 10th percentile (1.29 ± 0.46 , ranging from -1.81 to $+0.07$ SDS) and thus comparable to the GH-treated group (-1.04 ± 0.865 ; $P=0.09$). At present, all patients who have already reached their final height have grown to their parental target height range.

Cohort of children with GHD

From the patient cohort, 250 patients were proven to be GH deficient by two pathological stimulation tests at the time of diagnosis and were treated with rhGH (Fig. 1 and Table 1), and 90 subjects were excluded from our study population due to the violation of inclusion/exclusion criteria. Most of them had other medical conditions such as syndromal diseases, SGA, chronic illness, etc. Considering a sufficient amount of remaining serum for re-testing, 44 patients (isolated GHD $n=41$; panhypopituitarism after completed treatment of a tumor, $n=3$) were available for this cohort of children with GHD. At the time of stimulation tests, the mean age of this cohort was 9.92 ± 3.64 years with a height SDS of -2.11 ± 0.65 and deviation from parental target height

SDS of -0.88 ± 0.80 (mean $\pm$ s.d.). All patients showed a GH concentration below $7.32 \mu\text{g/l}$ in two different stimulation tests.

Checking of sample stability over storage

A total of 114 serum samples were randomly selected from the two above-mentioned cohorts. Samples had been stored at $-80 \pm 5^\circ\text{C}$ for more than 11 years.

GH quantification by immunoassays

Originally, samples stored for up to 11 years were measured using the DELFIA method (Perkin-Elmer, Rodgau, Germany). Sensitivity was $0.015 \mu\text{g/l}$. Intra- and interassay coefficient of variation (CV) values were below 6.66% at $2.15 \mu\text{g/l}$, 5.82% at $5.77 \mu\text{g/l}$, and 7.43% at $12.4 \mu\text{g/l}$ ($n=425$). The same kit protocol was applied for the re-measurement of retained samples to check for the stability during long-term storage.

GH measurement in samples specified for the determination of a new cutoff limit for diagnosis of GHD was performed on the automated analyzer iSYS (IDS, Immunodiagnostic Systems). Details about this assay procedure have been recently published (18). Sensitivity was $0.05 \mu\text{g/l}$. Intra- and inter-assay CV values were 4.82% at $2.26 \mu\text{g/l}$, 2.27% at $5.86 \mu\text{g/l}$, and 3.83% at $12.19 \mu\text{g/l}$ ($n=18$).

Clinical Study	I V Wagner and others	Cutoff limits for GHD diagnostics in children	171:3	392
----------------	-----------------------	---	-------	-----

Table 1 Cohort characteristics of selected non-GHD and GHD children and adolescents. Data of the non-GHD cohort ($n=52$) were compared with GHD patients ($n=44$). Mean, s.d., median, and range were given for the measurements of age, height, delta growth velocity, familial target height, IGF, and IGFBP3 as well as for GH maximum at the time of testing compared with iSYS results measured in the year 2012.

Characteristics of subjects	Non-GHD			GHD		
	Mean $\pm$ s.d.	Median	Min–Max	Mean $\pm$ s.d.	Median	Min–Max
Age at stimulation test (years)	9.13 $\pm$ 3.73	8.60	0.75 to 14.73	9.93 $\pm$ 3.68	9.81	3.79 to 16.64
Height at stimulation test (SDS)	−2.18 $\pm$ 0.39	−2.00	−3.07 to 1.46	−2.12 $\pm$ 0.67	−2.10	−4.15 to 0.11
Height at last measurement (SDS)	−1.30 $\pm$ 0.46	−1.40	−1.81 to 0.07	−1.04 $\pm$ 0.87*	−1.08	−2.29 to 1.70
BMI SDS at the time of testing	−1.08 $\pm$ 1.08	−1.04	−3.72 to 2.08	0.18 $\pm$ 1.23	0.22	−2.72 to 2.53
Age at the last measurement (years)	12.91 $\pm$ 3.33	12.99	5.51 to 19.10	15.53 $\pm$ 4.42	15.48	6.91 to 26.88
Delta height (SDS) after the test	0.91 $\pm$ 0.59	0.84	0.11 to 2.84	1.08 $\pm$ 0.94	0.99	−0.64 to 3.82
Familial target height (SDS)	−0.91 $\pm$ 0.72	−0.91	−2.66 to 0.86	−0.89 $\pm$ 0.82	−0.83	−3.45 to 1.59
Height velocity SDS before the test	−0.46 $\pm$ 1.65	−0.55	−5.47 to 4.19	−0.66 $\pm$ 1.83	−0.68	−3.37 to 8.14
IGF1 SDS at the time of testing	−1.22 $\pm$ 1.51	−1.10	−5.01 to 1.59	−1.21 $\pm$ 1.54	−1.09	−6.39 to 1.15
IGFBP3 SDS at the time of testing	−1.10 $\pm$ 1.97	−1.02	−8.15 to 2.59	−1.22 $\pm$ 1.74	−1.05	−7.83 to 0.74
GH max at the time of testing (μ g/l)	12.96 $\pm$ 5.02	12.00	7.63 to 33.80	3.65 $\pm$ 1.87	4.05	0.03 to 7.32
GH max in 2012 (μ g/l) using iSYS	11.83 $\pm$ 4.42	10.89	4.74 to 28.42	3.60 $\pm$ 1.85	3.93	0.05 to 8.07

Max, maximum within the stimulation test. * $P>0.05$ for height SDS at last measurement for non-GHD vs GHD according to the Mann–Whitney U test.

For the translation of this cutoff limit to other frequently used hGH assays, we referred to the data panel of 312 sera of children and adolescents described in detail recently (19). These sera were measured using assays of DELFIA (Perkin-Elmer), Immulite-2000 (Siemens, Munchen, Germany), DxI and IRMA (Beckman-Coulter, Krefeld, Germany), Liaison (Diasorin, Dietzen, Germany), ELISA (Mediagnost, Reutlingen, Germany), and RIA (University of Tuebingen, (10)).

GH quantification by ID-MS

ID-MS analyses were run on a 4000 Q-Trap (Applied Biosystems) using a protocol described previously (20, 21), slightly adapted for the present study. Enzymatic GH fragments T6 and T12 were quantified as surrogates for ‘22 kDa GH’ and ‘whole GH’ respectively. The ID-MS method was validated according to the NCCLS guidelines (24). GH was measured in serum samples at five concentration levels with two replicates at each level. Linearity was found to be excellent (coefficient of determination R^2 , from regression analysis, was 0.99). The average recovery of GH from serum samples was 101.4%, compared with the expected (target) values. The limit of quantification (LoQ, μ g/l) was determined based on the common definition. The overall uncertainty (imprecision, $SD_{\text{total}}=3.5\%$) encompassing within- and between-run variations was estimated by GH measurement in serum samples within the concentration range of 0.5–8.0 μ g/l (21). The ‘between-run precision’ values were 4.2% for T6 and 4.6% for

T12 ($n=6$). All performance data of validation have been detailed previously (20, 21). Particular details about the ID-MS method are provided in the [Supplementary Data](#), see section on [supplementary data](#) given at the end of this article. A subset of serum samples of 30 patients (19 non-GHD and 11 GHD) was used for method comparison and calibration of iSYS against ID-MS. GH concentrations (2.02–20.14 μ g/l) of these sera would cover the concentration range dealt in this study.

Statistical analysis

Descriptive statistics was performed using Excel and Statistica (Statsoft, Tulsa, OK, USA). The MedCalc software (25) was used for statistical evaluation of the decision limits using iSYS-data obtained for both groups of patients, GHD and non-GHD.

Results

Stability of hGH in human serum

The re-measurement of GH (y) in 114 samples that had been stored for ~ 11 years at $-80 \pm 5^\circ\text{C}$ revealed comparable results to the initial measurements (x) before storage. Consequently, the Passing–Bablok regression equation was $y=1.022x-0.032$ μ g/l with no significant deviation from linearity ($P=0.45$; data not shown). The coefficient of correlation according to Spearman was calculated to be $R=0.994$.

iSYS cutoff limit

Based on the classification by auxological criteria and further clinical characteristics, discriminant analysis was applied to assess the diagnostic power of the stimulation test as well as to provide a statistically sound definition of cutoff values. iSYS GH concentrations of non-GHD and GHD patients were compared in Fig. 2a. In the ROC analysis (Fig. 2b), the area under the curve relative to the total area was $99.1 \pm 0.7\%$ and this indicated the average probability for a non-GHD patient to come out with a higher test result when compared with a GHD patient, and *vice versa*. The Youden index (J), indicating the point where the rate of correct classifications was at its maximum, was $7.09 \mu\text{g/l}$. The sum of misclassification rates was estimated to be $(1 - J) \approx 6\%$ at this point. Alternatively, the left-sided 95% reference interval was calculated using a non-parametrical method. This resulted in $7.06 \mu\text{g/l}$ GH as the criterion, which does not, however, make much difference to the Youden index. Considering this, a conservative approach for the final cutoff at $7.09 \mu\text{g/l}$ was reported herein.

Transformation of iSYS cutoff to kits of other manufacturers and to ID-MS

The cutoff limit derived from the above-mentioned iSYS results was translated to the seven other frequently used commercial hGH assays (Table 2). Individual cutoff limits

Table 2 Regression equations (ng/ml) for GH concentrations in serum of children and adolescents measured by several commercially available immunoassay kits and the 'Tuebingen RIA' used as a dependent parameter y were compared with data of the iSYS used as an independent parameter x (17). Coefficients of linear regression r are given in parentheses. The evaluated cutoff limits for each assay were calculated by transforming the iSYS cutoff value of 7.09 ng/ml in each regression equation. GH-RPs used in the individual assay are given as footnotes.

Assay	Regression equation (r)	Cutoff limit ($\mu\text{g/l}$)
Immulite 2000 (Siemens) ^a	$y = 1.031x - 0.455$ ($r = 0.964$)*	7.77
AutoDELFIA (Perkin-Elmer) ^b	$Y = 1.004x + 0.323$ ($r = 0.922$)*	7.44
iSYS (IDS) ^a	NA	7.09
Liaison (DiaSorin) ^a	$Y = 0.823x + 0.412$ ($r = 0.919$)*	6.25
RIA (in-house Tuebingen) ^a	$y = 0.565 + 1.271$ ($r = 0.818$)*	5.28
Dxl (Beckmann-Coulter) ^a	$Y = 0.689x + 0.271$ ($r = 0.880$)*	5.15
ELISA (Mediagnost) ^a	$y = 0.678 + 0.327$ ($r = 0.869$)*	5.14
BC-IRMA (Beckmann-Coulter) ^c	$y = 0.622x - 0.096$ ($r = 0.927$)*	4.32

* $P < 0.0001$.

^aWHO 2. IS 98/574.

^bWHO 1. IS 80/505.

^cWHO 1. IS 88/624.

Figure 2

(a) iSYS-GH test results for GHD- and non GHD children and adolescents; (b) Receiver Operating Characteristic (solid graph) for the stimulation test. Horizontal line in panel (a) marks the cut-off limit where the rate of correct classifications is at its maximum (Youden index $J = 0.939$).

were obtained based on proved linear inter-relationships in the range of $3\text{--}10 \mu\text{g/l}$ for the comparison between iSYS results and the hGH methods of DELFIA, Immulite-2000, Dxl, Liaison, Tuebingen RIA, Mediagnost ELISA, and Beckmann-Coulter IRMA (data not shown and (18)). It should be noted that these numbers are referring to the standard protocols as provided by the manufacturers and particularly to the use of the specified GH-RPs.

Plots of ID-MS results against iSYS and regression lines for the 30 sera are shown in Fig. 3. The antibody-independent ID-MS method enables highly selective acquisition of the 22 kDa form using the enzymatic GH cleavage product T6 as a probe, apart from a larger quantity of 'all GH' isoforms, obtained by quantification of T12 (see Supplementary Data). Statistically significant correlation between iSYS and ID-MS results (both T6 and T12) was found with $R > 0.94$ ($P < 0.001$). Linear models were calculated by (unweighted) Deming regression analysis ($T6 = -0.029 \mu\text{g/l} + 0.777 \times \text{iSYS}$ and $T12 = 1.373 \mu\text{g/l} + 0.854 \times \text{iSYS}$) for the prediction of

Figure 3

GH-quantification by ID-MS ($n=30$). Regression graphs are shown at top and residuals plots at the bottom. Results for iSYS- and ID-MS are compared. Calibration in Figure 3 (a) is T6-based and in Figure 3 (b) T12-based.

ID-MS results from iSYS measurements, which led to the transformed cutoffs of $5.48 \mu\text{g/l}$ for the T6 and $7.43 \mu\text{g/l}$ for the T12 fragment respectively. These data suggest the highly quantitative agreement of immunological and mass spectrometric measurements.

Discussion

Stability of sera

To ensure the comparability of the current iSYS data to the GH data at initial workup in 2000 and 2001, it was necessary to check the stability of GH over storage. Therefore, a subset of samples, which originally had been measured by DELFIA assay, were re-measured using the same method, before changing to iSYS. To the best of our knowledge, this is the first study describing the long-term stability of GH at $-80 \pm 5^\circ\text{C}$. There is only one 30-year-old study analyzing the stability of protein hormones that could be detected by RIA after repeated freeze-thaw cycles, but there are no long-term data available (26). According to our results, the 22 kDa GH isoform, which is specifically detected by the applied DELFIA assay, appears to be stable for a period of 10 years at -80°C with regard to its immunological properties.

A new cutoff limit based on a cohort of non-GHD children and adolescents

Most of the attempts to establish GH cutoff limits for children and adolescents are 30–40 years old and they mostly tried to test the different stimulation tests and different methods rather than to define a unique cutoff concentration (3, 4). There are too many variables that prevent the definition of a single cutoff value. Different types of stimulation test, heterogeneity in assay characteristics, and the performance of analysts can yield different results in GH determination (15, 16). Moreover, many modifying factors (e.g. age, body mass, nutritional status, puberty, fasting state, etc.) may have an effect on circulating GH concentrations (2, 3, 15, 17). In the last 40 years, several studies described the establishment of cutoff values in adults (2, 27, 28), but only a paucity of data exists for the important group of children and adolescents (for review, see (29)).

The most suitable pediatric approach was delivered by Binder *et al.* (10) and they defined GH cutoff concentrations in patients based on their auxology (10). GH deficiency was defined solely by analyzing the growth dynamics before and during GH therapy. Thereby, three inclusion criteria were used to define GHD based on only auxology: non-familial short stature, pre-treatment catch-down growth, and treatment catch-up growth. Accordingly, a cutoff limit was calculated on the basis of the 95th percentile of the GHD cohort. To validate these data, sera of children with Turner's syndrome ($n=34$) or SGA ($n=74$) were used instead of sera of apparent non-GHD children. One further limitation is that GH was measured only by an in-house RIA of the Tuebingen Children's Hospital and not by other commercially available tests. According to their suggestion, GH reference data should be established in the same experimental setting presented in this article (10). But this idea seems not to be practical.

By contrast, our approach was based on the cutoff limit of a non-GHD cohort of children and adolescents who presented a low height SDS and/or a low growth velocity at diagnoses. We used the iSYS GH assay that proved to be a new and sensitive method to measure GH levels in serum samples of pediatric patients (17). It should be recognized that the results of this cohort do not exactly reflect the condition of healthy children and adolescents. However, it is ethically impossible to test perfectly healthy children of normal height. In addition, our approach of comparing data of short patients without GHD with data of GHD patients would more closely reflect the reality of the everyday clinical setting. As it was relatively difficult to

gather non-GHD patients with stimulation test results, we had to accept a small sample size of 52. For the evaluation of this cutoff limit, we further used randomly selected sera from GHD patients. The cohort size of 44 constitutes a first set for evaluation and the results, and this should be validated with a higher number of patients in the near future. Nevertheless, the calculated cutoff limit for the iSYS method of 7.09 µg/l is in good agreement with the target cut-point of 6.7 µg/l (upper limit 8.0 µg/l), which has been recently recommended in Germany (13). The use of a unique cut-point presumes the hypothesis that all three stimulants used in the GH stimulation tests should have the same potency to induce GH secretion. Such a perception is common in the literature, observing that there are no systematic investigations available so far. Poor reproducibility of GH stimulation tests with each of the three stimulants justified this hypothesis for practical use (30, 31).

Transformation of data to other commercially available GH kits

Owing to the limited amount of available serum, it was impossible to re-measure all the samples with all of the GH testing methods listed in Table 2. Therefore, linear relationships, previously established by Müller *et al.* (19), were used for the calculation of cutoff limits of six commercially available immunoassays as well as for the Tuebingen RIA (19). However, the use of linear statistics for the transformation of these data has a common disadvantage in that a bias becomes noticeable with increasing deviation from the linear regression line. The latter has to be at least taken into account when interpreting the data from the 'Tuebingen RIA', which correlates with iSYS with a correlation coefficient *R* of only 0.818. However, GH concentrations in sera of children had a largely linear correlation between individual assays within a range of 3–10 µg/l according to the Passing–Bablok method (data not shown), which covers our calculated cutoff limits broadly. This finding suggests that there is no detectable systematic bias between the individual assays. Moreover, the bias observed as scatter from the regression line between various analytical methods appears to be distinctively lower than the differences between the calculated cutoff limits ranging from 4.32 to 7.77 µg/l. These differences display the high variability of immunoassay results for one sample and underline that assay-specific cutoffs are mandatory for the diagnosis of GHD. Interestingly, our approach revealed a calculated cutoff for the 'Tuebingen in-house RIA' as 5.28 µg/l, which was somewhat different from the cutoff of

6.60 µg/l, which was determined by the approach of Binder *et al.* (10) as the 95th percentile for GHD children. This difference may indicate the restricted diagnostic sensitivity of GH stimulation tests, which might be further crucial even by an improved approach in determining the cutoff values for the clinical diagnostics of GHD in children and adolescents.

Comparison of immunoassay results with ID-MS

Mass spectrometry provides an independent approach to determine GH without major downsides of antibody-based assays. This is accomplished by the definition of enzymatic GH cleavage products (i.e. T6 and T12) unambiguously coding for the particular GH forms of interest (i.e. 22 kDa and total), in combination with extremely selective recognition of these products in mass spectrometry, even from complex matrix such as serum digests. (For a more detailed outline of the method principle, see [Supplementary Data](#).) For instance, it has been recently demonstrated that GHBP does not interfere with GH results in ID-MS (20). Similarly, by the measurement principle, the validity of results will not be compromised by dimers or by higher GH aggregates that may be present. In addition, it will not be compromised by accompanying proteins in general. At the same time, the use of isotopically labeled internal standards (U-¹⁵N recombinant 22 kDa GH) renders high reliability in this method. Most importantly, traceability to the International System of Units (SI) (21) is obtained with reference to the amount of substance of amino acids in value assignment to the GH calibrator stock solution used. This is crucial for the long-term validity of the cutoff values obtained in this study, and it makes ID-MS results independent of a particular GH-RP to be used. Similarly, by linking the cutoff values to ID-MS as described, the cutoff values will be applicable even after the expiry of the commercial kits/antibodies that were used to establish them.

Regarding the iSYS assay, the comparison with the ID-MS results suggests that concentration of the 22 kDa GH form is somewhat overestimated and it appears that the total amount of GH hormone is reflected in a roughly unbiased manner. A definitive statement as to this, however, would require a separate study, preferably with a further higher number of sera included. In general, the comparison of ID-MS data with the results from GH immunochemical assays is an option to detect the gaps in the traceability chain, which is frequently encountered in practice (32).

For the purpose of this study, the sample size was limited to $n=30$, owing to the considerably increased demand in the sample-processing time of ID-MS compared with current commercial assays and a limited volume of serum for both analyses. In line with our expectation, the ratio of cutoff concentrations obtained for T6 and T12 (5.48/7.43) matched fairly well with the estimated fraction of the 22 kDa GH with respect to total GH in the pituitary (73%) (33). The T6- and T12-based cutoff levels should be applicable as reference points for determining appropriate decision limits with a (hypothetical) future antibody-based kit that requires nothing but calibration of assay results against T6 or T12 (depending on the selectivity of that antibody).

In conclusion, we present the cutoff decision limits for the laboratory diagnostics of GHD in children and adolescents established with stored sera from stimulation tests of non-GHD and GHD children. The results are SI traceable with reference to ID-MS and close to the suggested data of the national guidelines. This procedure is thought to be an important step to improve the reliability of the diagnosis of GHD in children and adolescents.

Supplementary data

This is linked to the online version of the paper at <http://dx.doi.org/10.1530/EJE-14-0165>.

Declaration of interest

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

Funding

The development of mass spectrometric methods used in this study for the determination of serum growth hormone was supported by the fund from the European Community's Seventh Framework Programme, ERA-NET Plus, under the iMERA-Plus (project-grant agreement number 217257).

Acknowledgements

The authors thank Rüdiger Ohlendorf for performing mass spectrometric amino acid analyses.

References

- Wales JK, Wit JM & Rogol AD. *Pediatric Endocrinology and Growth*. London: Saunders, 2003.
- Eskes SA, Tomaso NB, Endert E, Geskus RB, Fliers E & Wiersinga WM. Establishment of reference values for endocrine tests. Part VII: growth hormone deficiency. *Netherlands Journal of Medicine* 2009 **67** 127–133.
- Frasier DS. A preview of growth hormone stimulation tests in children. *Pediatrics* 1974 **53** 929–937.
- Kaplan SL, Abrams CAL, Bell JJ, Conte FA & Grumbach MM. Growth and growth hormone. Changes in serum level of growth hormone following hypoglycemia in 134 children with growth retardation. *Pediatric Research* 1968 **2** 43–63. (doi:10.1203/00006450-196801000-00005)
- Youlton R, Kaplan SL & Grumbach MM. Growth and growth hormone. IV. Limitations of the growth hormone response to insulin and arginine and of the immunoreactive insulin response to arginine in the assessment of growth hormone deficiency in children. *Pediatrics* 1969 **43** 989–1004.
- Brook CGD & Hindmarsch PC. Tests for growth hormone secretion. *Archives of Disease in Childhood* 1991 **66** 85–87. (doi:10.1136/adc.66.1.85)
- Stimmler L & Brown GA. Growth hormone secretion provoked by insulin-induced hypoglycaemia in children of short stature. *Archives of Disease in Childhood* 1969 **42** 232–238. (doi:10.1136/adc.42.223.232)
- Carel JC, Tresca J-P, Letrait M, Chaussain J-L, Lebouc Y, Job J-C & Coste J. Growth hormone testing for the diagnosis of growth hormone deficiency in childhood: a population register-based study. *Journal of Clinical Endocrinology and Metabolism* 1997 **82** 2117–2121. (doi:10.1210/jcem.82.7.4106)
- Roth J, Glick SM, Yalow RS & Berson SA. Secretion of human growth hormone: physiologic and experimental modification. *Metabolism* 1963 **12** 577–579.
- Binder G, Huller E, Blumenstock G & Schweizer R. Auxology-based cut-off values for biochemical testing of GH-secretion in childhood. *Growth Hormone & IGF Research* 2011 **2** 212–218. (doi:10.1016/j.ghir.2011.05.007)
- Binder G. Growth hormone deficiency: new approaches to the diagnosis. *Pediatric Endocrinology Reviews* 2011 **1** 535–537.
- Werther GA. Growth hormone measurements versus auxology in treatment decisions: the Australian experience. *Journal of Pediatrics* 1996 **128** 47–51. (doi:10.1016/S0022-3476(96)70011-2)
- Binder G. Guideline of the German Society of Endocrinology/Study Group Pediatric Endocrinology/Diabetology. Diagnostics of growth hormone deficiency in childhood and adolescence. Extensive version. See <http://www.uni-duesseldorf.de/AWMF/ii>.
- 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. *Journal of Clinical Endocrinology and Metabolism* 2000 **85** 3990–3993.
- Clemmons DR. Consensus statement on the standardization and evaluation of growth hormone and insulin-like growth factor assays. *Clinical Chemistry* 2011 **57** 555–559. (doi:10.1373/clinchem.2010.150631)
- Strasburger CJ & Bidlingmaier M. How robust are laboratory measures of growth hormone status? *Hormone Research* 2005 **64** 1–5. (doi:10.1159/000087745)
- Bidlingmaier M & Freda PU. Measurement of human growth hormone by immunoassays: current status, unsolved problems and clinical consequences. *Growth Hormone & IGF Research* 2010 **20** 19–25. (doi:10.1016/j.ghir.2009.09.005)
- Manolopoulou J, Alami Y, Petersenn S, Schopohl J, Wu Z, Strasburger CJ & Bidlingmaier M. Automated 22-kD growth hormone-specific assay without interference from Pegvisomant. *Clinical Chemistry* 2012 **58** 144–156. (doi:10.1373/clinchem.2012.188128)
- Müller A, Scholz M, Blankenstein O, Binder G, Pfäffle R, Körner A, Kiess W, Heider A, Bidlingmaier M, Thierry J *et al.* Harmonization of growth hormone measurements with different immunoassays by data adjustment. *Clinical Chemistry and Laboratory Medicine* 2011 **49** 1135–1142.
- Arsene CG, Henrion A, Diekmann N, Manolopoulou J & Bidlingmaier M. Quantification of growth hormone in serum by

- isotope dilution mass spectrometry. *Analytical Biochemistry* 2010 **401** 228–235. (doi:10.1016/j.ab.2010.03.005)
- 21 Arsene CG, Schulze D, Kratzsch J & Henrion A. High sensitivity mass spectrometric quantification of serum growth hormone by amphiphilic peptide conjugation. *Journal of Mass Spectrometry* 2012 **47** 1554–1560. (doi:10.1002/jms.3094)
 - 22 Burkitt WI, Pritchard C, Arsene C, Henrion A, Bunk D & O'Connor G. Toward Système International d'Unité-traceable protein quantification: from amino acids to proteins. *Analytical Biochemistry* 2008 **376** 242–251. (doi:10.1016/j.ab.2008.02.010)
 - 23 Hoepffner W, Pfaeffle W, Gausche R, Meigen C & Keller E. Early detection growth disorders with the CrescNet system at the Leipzig treatment center. *Deutsches Ärzteblatt* 2011 **8** 123–128.
 - 24 CLSI. Preliminary Evaluation of Quantitative Clinical Laboratory Measurement Procedures; Approved Guidelines-Third Edition. CLSI document EP10-AS-AMP. Wayne PA: Clinical and Laboratory Standards Institute, 2014.
 - 25 MedCalc 12.5.0.0 – 64-bit, ©1993–2013 MedCalc-Software bvba; MedCalc software, Acacialaan 22, B-8400 Ostend, Belgium. <http://www.medcalc.org>.
 - 26 Bauman JE. Stability of radioimmunoassayable steroid and protein hormones after repeated freeze–thaw cycles. *Clinical Chemistry* 1982 **28** 2336–2337.
 - 27 Corneli G, Di Somma C, Baldelli R, Rovere S, Gasco V, Croce CG, Grottoli S, Maccario M, Colao A, Lombardi G *et al.* The cut-off limits of the GH response to GH-releasing hormone-arginine test related to body mass index. *European Journal of Endocrinology* 2005 **153** 257–264. (doi:10.1530/eje.1.01967)
 - 28 Klose M, Stochholm K, Janukonyte J, Lehman Christensen L, Frystyk J, Andersen M, Laurberg P, Christiansen JS & Feldt-Rasmussen U. Prevalence of posttraumatic growth hormone deficiency is highly dependent on the diagnostic set-up: results from The Danish National Study on Posttraumatic Hypopituitarism. *Journal of Clinical Endocrinology and Metabolism* 2014 **99** 101–110. (doi:10.1210/jc.2013-2397)
 - 29 Ranke MB. Growth hormone deficiency: diagnostic principles and practice. In *Diagnostics of Endocrine Function in Children and Adolescents*, 4th edn, pp 115–124. Eds MB Ranke & PE Mullis. Basel: Karger, 2011.
 - 30 Hilczer M, Smoczynska J, Stawarska R & Lewinski A. Stability of IGF-I concentration despite divergent results of repeated GH stimulating tests indicates poor reproducibility of test results. *Endocrine Regulations* 2006 **40** 37–45.
 - 31 Fisker S, Jørgensen JO & Christiansen JS. Variability in growth hormone stimulation tests. *Growth Hormone & IGF Research* 1998 **8** 31–35. (doi:10.1016/S1096-6374(98)80006-3)
 - 32 Boulo S, Hanisch K, Bidlingmaier M, Arsene CG, Panteghini M, Auclair G, Sturgeon C, Schimmel H & Zegers I. Gaps in the traceability chain of human growth hormone measurements. *Clinical Chemistry* 2013 **59** 1074–1082. (doi:10.1373/clinchem.2012.199489)
 - 33 Arsene CG, Ohlendorf R, Burkitt W, Pritchard C, Henrion A, O'Connor G, Bunk DM & Güttler B. Protein quantification by isotope dilution mass spectrometry of proteolytic fragments: cleavage rate and accuracy. *Analytical Chemistry* 2008 **80** 4154–4160. (doi:10.1021/ac7024738)

Received 26 February 2014

Revised 22 May 2014

Accepted 24 June 2014