The cut-off limits of the GH response to GH-releasing hormone-arginine test related to body mass index

Ginevra Corneli, Carolina Di Somma¹, Roberto Baldelli, Silvia Rovere, Valentina Gasco, Chiara Giulia Croce, Silvia Grottoli, Mauro Maccario, Annamaria Colao¹, Gaetano Lombardi¹, Ezio Ghigo, Franco Camanni and Gianluca Aimaretti

Division of Endocrinology and Metabolic Diseases, Department of Internal Medicine, University of Turin, Corso Dogliotti 14, 10126 Turin, Italy and¹Division of Endocrinology, Federico II University of Naples, Naples, Italy

(Correspondence should be addressed to G Aimaretti; Email: gianluca.aimaretti@unito.it)

Abstract

Objective: The diagnosis of growth hormone (GH) deficiency (GHD) in adults is based on a reduced peak GH response to provocative tests, such as the insulin tolerance test (ITT) and the GH-releasing hormone-arginine (GHRH-ARG) test. However, the cut-off limits of peak GH response in lean subjects are not reliable in obese patients: this is noteworthy since adult GHD is often associated with obesity. Aim of this study was to evaluate the diagnostic cut-off limits of peak GH response to the GHRH-ARG test in overweight and obese as well as in lean population.

Design and methods: The GH responses to the GHRH-ARG test were studied in 322 patients with organic hypothalamic-pituitary disease and in 318 control subjects. Patients were subdivided into two groups on the basis of the number of pituitary hormone deficits, except for GH deficiency: (a) patients with total pituitary hormone deficit (TPHD) and (b) patients without or with no more than two pituitary hormone deficits (PHD). Both patients and control subjects were divided into three subgroups according to body mass index (BMI): lean (BMI $\geq 25$ kg/m$^2$), overweight (BMI $25 < 30$ kg/m$^2$) and obese (BMI $\geq 30$ kg/m$^2$). TPHD patients were assumed to be GH deficient, whereas PHD patients may include subjects with either normal or impaired GH secretion. The statistical analysis was carried out by the Receiver-Operating Characteristic curve analysis (Medcalc 7.2). The diagnostic cut-off points were calculated for lean, overweight and obese subjects to provide optimal separation of GH-deficient patients and control subjects according to two criteria: (1) a balance between high sensitivity and high specificity; (2) to provide the highest pair of sensitivity/specificity values for GH deficiency.

Results: In the lean population the best pair of values, with highest sensitivity as 98.7% and highest specificity as 83.7%, was found using a peak GH cut-off point of 11.5 m$\mu$g/l. In the overweight population the best pair of values, 96.7 and 75.5%, respectively, was found using a peak GH cut-off point of 8.0 m$\mu$g/l. In the obese population the best pair of values, 93.5 and 78.3%, respectively, was found using a peak GH cut-off point of 4.2 m$\mu$g/l. Applying the above mentioned cut-off points, among PHD patients we found that 80 subjects (72%) were GHD whereas 31 (28%) had normal GH secretion.

Conclusions: In conclusion the GHRH-ARG test is a reliable tool for the diagnosis of adult GH deficiency in lean, overweight and obese patients, provided that specific BMI-related cut-off limits are assumed.

European Journal of Endocrinology 153 257–264

Introduction

Growth hormone (GH) deficiency (GHD) is a common finding in adults with acquired hypothalamic-pituitary disorders or is the persistence of a congenital or acquired somatotroph defect diagnosed in childhood (1). GH has important physiological functions in adult humans and several studies have shown that GHD in adult patients is associated with abnormalities in body composition, visceral obesity, metabolic derangements and impaired physical performance (2–7). These impairments improve with recombinant human GH-replacement therapy (rhGH) (2, 4, 6, 8–11).

The clinical features of adult GHD are recognizable but not distinctive, so clinical suspicion must be confirmed by biochemical tests (3, 12–18). Within an appropriate clinical context, GHD in adults may be shown by a single provocative testing, provided that a reproducible test with clear, normative limits is available (1, 18). Insulin-like growth factor-I (IGF-I) and
IGF-binding protein-3 (IGFBP-3), as well as spontaneous GH secretion, are not reliable parameters for the diagnosis of adult GHD (1, 16, 18–20), although the measurement of IGF-I differentiates different degrees of GH deficiency (21).

The insulin tolerance test (ITT) has been indicated as the test of choice (1). A GH peak below 5 μg/l has been reported to be an abnormal response to the ITT (1, 12, 18), whereas severe GHD is defined as a GH response lower than the arbitrary cut-off value of 3 μg/l (1). However, the ITT has been challenged because of its low reproducibility and for its contraindications in several clinical conditions often observed in adult patients with suspected GHD, such as ischaemic heart disease and seizure disorders (1, 16, 22, 23). The GH-releasing hormone-arginine (GHRH-ARG) test represents, at present, one of the best alternatives to the ITT to explore the somatotroph function and for the diagnosis of GHD in non-obese adult patients (1, 16, 18, 24). It is a potent, reproducible and reliable test that shows excellent sensitivity and specificity and is unaffected by gender and aging (1, 16, 18, 25). Moreover, we have recently reported that the GHRH-ARG test can show different degrees of GH deficiency (21). The third and the first centile limits of the peak GH response to the GHRH-ARG test evaluated in a population of 157 normal lean subjects were 16.5 and 9 μg/l, respectively (16). However, in clinical practice the reliability of the GHRH-ARG test, as well as that of ITT, is limited by obesity which is a condition characterized by reduced spontaneous GH secretion and low somatotroph response to all provocative stimuli (26–33). Obesity represents the most important confounding factor for the diagnosis of GHD. In obese subjects the GH response to the GHRH-ARG or insulin tests overlaps frequently with that found in adult patients with GHD (26, 28, 30, 31).

Recently, in a comparative study regarding six different provocative tests for the diagnosis of GHD in adults with hypothalamic-pituitary disease, carefully matched with control subjects, Biller et al. (34) confirmed that the greatest diagnostic accuracy was obtained with the ITT and GHRH-ARG tests. The authors defined the specific cut-off points for each test in order to provide the best pair of sensitivity/specifcity values. For the GHRH-ARG test they found that a peak GH response cut-off limit of 4.1 μg/l provided 95% sensitivity and 91% specificity. This cut-off is lower than that found, but control subjects studied by Biller et al. (34) had a higher body mass index (BMI; 30.3 ± 8.8 kg/m²).

It has to be taken into account that nearly half of the patients with acquired hypothalamic-pituitary disease are overweight or obese (3, 4, 7), indicating the clinical need to define BMI-related cut-off limits of the GH response to provocative tests for the diagnosis of adult GHD. Therefore, the aim of our study was to define appropriate diagnostic cut-off limits of peak GH response to the GHRH-ARG test related to BMI.

Subjects and methods

We studied the GH response to the GHRH-ARG test in 322 patients (174 men and 148 women, age 47.8 ± 1.2 years) with organic hypothalamic-pituitary disease. The most common disorders in this group were hypothalamic-pituitary adenomas (70.8%) or other peri-pituitary tumours (23.0%) requiring neurosurgery and/or radiotherapy. The patients were subdivided into two groups on the basis of the number of pituitary hormone deficits, except for GH deficiency. The first group consisted of 211 patients with total pituitary hormone deficit (TPHD). They all had thyroid-stimulating hormone (TSH), corticotropin and luteinizing hormone (LH) and/or follicle-stimulating hormone (FSH) deficiency; 13 patients also had diabetes insipidus. The second group consisted of 111 patients without or with no more than two pituitary hormone deficits (PHD). Ten patients (9%) were TSH-deficient in isolation, 15 (13.5%) were corticotropin-deficient in isolation, 19 (17.1%) were gonadotropin-deficient in isolation, 11 (9.9%) were both TSH- and corticotropin-deficient, seven (6.3%) were both TSH- and gonadotropin-deficient, 14 (12.6%) were both corticotropin- and gonadotropin-deficient and 35 (31.5%) did not have any pituitary hormone deficiency other than suspected GHD.

All patients were studied during adequate hormonal-replacement therapy, where necessary. Patients treated with dopaminoagonists or with a previous history of acromegaly, active Cushing’s disease, diabetes, malignancy, or renal or hepatic dysfunction were excluded from the study. As a control group we studied 318 normal subjects (147 men and 171 women; age 39.9 ± 0.8 years).

Both patients and control subjects were divided into three subgroups according to BMI: lean (BMI < 25 kg/m²), overweight (BMI ≥ 25 and < 30 kg/m²) and obese (BMI ≥ 30 kg/m²; Table 1). Among the patients, 120 (37.3%) were lean, 120 (37.3%) were overweight and 82 (25.5%) were obese. Among the control subjects, 215 (67.6%), 48 (15.1%) and 55 (17.3%) were lean, overweight and obese, respectively. IGF-I levels in patients and controls are reported in Table 1.

The study was approved by the local Ethics Committee and the patients provided consent.

Methods

After an overnight fast all subjects underwent a GHRH-ARG test (GHRH1-29; GEREE, Serono, Italy; 1 μg/kg i.v. at 0 min; arginine hydrochloride, 0.5 g/kg i.v. over 30 min from 0 to +30 min, up to a maximum of 30 g). Blood samples for GH evaluation were taken every 15 min from 0 to +90 min.

Serum GH was assayed by immunoradiometric method assay (IRMA) method (HGH-CTK IRMA; Diasorin, Saluggia, Italy). All samples from an individual subject were analysed together. The sensitivity of the
Table 1  Gender, age, IGF-I and BMI in patients with TPHD, or without or with no more than two PHD, and control subjects (CS).

<table>
<thead>
<tr>
<th></th>
<th>TPHD</th>
<th>PHD</th>
<th>CS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>211</td>
<td>111</td>
<td>318</td>
</tr>
<tr>
<td>Men</td>
<td>117</td>
<td>57</td>
<td>147</td>
</tr>
<tr>
<td>Women</td>
<td>94</td>
<td>54</td>
<td>171</td>
</tr>
<tr>
<td>Age (years)</td>
<td>47.6±1.1</td>
<td>48.0±1.3</td>
<td>39.9±0.8</td>
</tr>
<tr>
<td>IGF-I (μg/l)</td>
<td>82.1±3.1</td>
<td>116.2±6.3</td>
<td>381.0±13.3</td>
</tr>
<tr>
<td>Lean (BMI &lt; 25 kg/m²)</td>
<td>n = 77; 36.6%</td>
<td>n = 43; 38.7%</td>
<td>n = 215; 67.6%</td>
</tr>
<tr>
<td>Overweight (BMI ≥25 and &lt;30 kg/m²)</td>
<td>n = 82; 38.8%</td>
<td>n = 38; 34.2%</td>
<td>n = 48; 15.1%</td>
</tr>
<tr>
<td>Obese (BMI ≥30 kg/m²)</td>
<td>n = 52; 24.8%</td>
<td>n = 30; 27.0%</td>
<td>n = 55; 17.3%</td>
</tr>
</tbody>
</table>

method was 0.15 μg/l. The inter- and intra-assay coefficients of variation were 3.5–4.4 and 5.1–7.5%, respectively, at GH levels of 1.98–41.92 and 2.99–42.45 μg/l, respectively.

Serum IGF-I was assayed by RIA method (SM-C-RIA-CT; Pantec, Turin, Italy) after acid-ethanol extraction to avoid interference by binding proteins. The sensitivity of the method was 0.1 μg/l. The inter- and intra-assay coefficients of variation were 5.0–9.5 and 8.8–10.8%, respectively, at IGF-I levels of 79.41–712.3 and 79.6–766.4 μg/l, respectively.

Statistical analysis

Results are expressed as mean±S.E.M. of absolute values. Differences between the groups were tested with the Mann–Whitney U test. Correlations were assessed by linear regression analysis.

The lowest limit of normality for peak GH responses to the GHRH-ARG test was defined as the value that provided the best pair of sensitivity/specificity values based on Receiver-Operating Characteristic (ROC) curve analysis (Medcalc 7.2) (35). For the purpose of ROC analysis, we assumed the patients with TPHD to be GHD, according to other studies and to the statement derived from the Growth Hormone Research Society consensus (1, 34, 36, 37). The PHD group (patients without or with no more than two PHD) included subjects with either normal or impaired GH secretion.

ROC curves are constructed by plotting the sensitivity on the ordinate as a function of the complement of specificity for all the possible cut-off values of the diagnostic test. Each point of the ROC curves represents a sensitivity/specificity pair corresponding to a particular decision threshold. The area under the ROC curve (ROC AUC) represents the probability of correctly distinguishing between affected and non-affected individuals. A perfect diagnostic test has an ROC curve that passes through the upper left-hand corner (area under the curve = 1), where the true-positive fraction is 1.0 or 100% (perfect sensitivity) and the false-positive fraction is 0 (perfect specificity). Tests with an area under the curve of less than 0.5 would not discriminate between affected and non-affected subjects (35).

The diagnostic cut-off points were calculated for the lean, overweight and obese groups. To provide optimal separation of GHD and normal subjects we applied two criteria: (1) to minimize misclassification of control subjects and GHD, balancing between high sensitivity and high specificity; (2) to provide the highest pair of sensitivity and specificity values for GHD.

Positive predictive value (PPV), sensitivity and specificity were calculated using the number of patients with true-positive (TP), true-negative (TN), false-positive (FP) and false-negative (FN) results. Sensitivity (calculated as TP/(TP + FN)) and specificity (calculated as TN/(TN + FP)) were defined as the percentages of TPHD patients who had a peak GH value below and above the specific BMI-related cut-off point, respectively. PPV was defined as the likelihood that a subject with a positive test (peak GH below the specific BMI-related cut-off point) was clinically GH-deficient, based on the presence of TPHD (calculated as TP/(TP + FP)).

Results

Age-corrected IGF-I levels were markedly lower in TPHD patients than in control subjects (P < 0.0001). In PHD patients IGF-I levels were intermediate between the values of the other two groups, being significantly higher than in TPHD patients (P < 0.05) and lower than in controls (P < 0.001; Table 1). No differences in IGF-I levels were found among lean, overweight and obese subgroups both in patients and in control subjects.

Peak GH responses to GHRH-ARG test in patients and controls related to BMI are shown in Table 2. The peak always occurred between 30 and 60 min. Peak GH responses were significantly higher in lean than in overweight or obese (P < 0.0001) and in overweight versus obese control subjects (P < 0.0001). On the contrary, there was no difference among lean, overweight and obese TPHD and PHD patients. A significant negative correlation (r = −0.5, P < 0.0001) between peak GH responses and BMI was found in controls but not in patients (Fig. 1).

In the lean population the best pair of values for highest sensitivity, 98.7%, and highest specificity, 83.7%, was found using a peak GH cut-off point of 11.5 μg/l with a
good accuracy of 87.1% (Fig. 2; Table 3). For a sensitivity of 95%, a cut-off point of 7.9 μg/l yielded a specificity of 86.0% and for a specificity of 95%, a cut-off point of 2.3 μg/l yielded a sensitivity of 54.5% (Table 3).

In the overweight population the best pair of values for highest sensitivity, 96.7%, and highest specificity, 75.5%, was found using a peak GH cut-off point of 8.0 μg/l, with a good accuracy of 78.6% (Fig. 3; Table 3). For a sensitivity of 95%, a cut-off point of 7 μg/l yielded a specificity of 76.9% and for a specificity of 95%, a cut-off point of 1.5 μg/l yielded a sensitivity of 52.9% (Table 3).

In the obese population the best pair of values for highest sensitivity, 93.5%, and highest specificity, 78.3%, was found using a peak GH cut-off point of 4.2 μg/l, with a good accuracy as 83.9% (Fig. 4; Table 3). For a sensitivity of 95%, a cut-off point of 4.7 μg/l yielded a specificity of 73.5% and for a specificity of 95%, a cut-off point of 1.0 μg/l yielded a sensitivity of 58.7% (Table 3).

For the lean, overweight and obese populations the cut-off points with the best pair of values for sensitivity/specificity, accuracy, ROC AUC, PPV and negative predictive value (NPV) with their confidence intervals are shown in Table 3. Cut-off points for the sensitivity and the specificity at 95% are also reported.

Applying the above-mentioned cut-off points that minimize misclassification of TPHD and controls, among the PHD group we found that 80 subjects (72.0%) were GHD whereas 31 (28%) had normal GH secretion; specifically, 38 out of 43 (88.3%), 31 out of 38 (81.5%) and 11 out of 30 (36.6%) were GHD in lean, overweight and obese patients, respectively.

In the 80 PHD patients who were found to be GHD, IGF-I levels were significantly lower than in the 31 without GHD (106.7 ± 5.5 versus 142.4 ± 7.7 μg/l, P < 0.0001). Moreover, a significant negative correlation (r = -0.6, P < 0.0001) between peak GH
responses to the GHRH-ARG test and BMI was found in PHD patients without GHD but not in those with GHD. Finally, PHD patients found to be GHD or not had one or two more pituitary hormone deficits with percentages of 77.4 and 48.3%, respectively.

**Discussion**

The diagnosis of GHD in adults is challenging because of the lack of a single specific biological end point, such as growth failure, which is the cardinal clinical sign in paediatric patients. The features of adult GHD are recognizable but not distinctive, so clinical suspicion must be confirmed by biochemical tests (1, 3, 12–18). Consensus guidelines for the diagnosis and treatment of adult GHD recommend provocative testing of GH secretion for patients with hypothalamic-pituitary disease or with childhood-onset GHD; moreover, patients who have undergone cranial irradiation or have a history of head trauma should also be tested (17, 38–40).

The GHRH-ARG and GHRH + GHRP-6 tests represent, at present, the best alternatives to ITT to explore the somatotroph function and to differentiate GHD from normal subjects, provided that appropriate cut-off limits are considered (1, 16–18, 24, 41). In agreement with our previous report (42), as peak GH response always occurred between 30 and 60 min, present results confirm that the test requires only three blood samples for GH evaluation (at 30, 45 and 60 min), avoiding basal determination.

In the diagnostic approach for GH deficiency obesity presents a great limitation since it is characterized by a low response to all provocative stimuli for GH secretion (26–33). In obesity reduction in the half-life of GH (43) as well as a significant decrease in the production and secretion of the hormone have been reported (44). Among potential neuroendocrine causes of the marked impairment in spontaneous and stimulated GH secretion, GHRH hypoactivity has been shown but it is likely that alterations in the modulation of ghrelin, neuropeptide Y and/or leptin could have a role (45). Among metabolic factors, chronic elevation of non-esterified fatty acid levels and hyperinsulinism probably have a key role in reducing GH secretion in obesity (31, 46–48).

In this study clear diagnostic cut-off limits of the GH response to the GHRH-ARG test in lean as well as in

| Table 3 | Cut-off points of peak GH response to the GHRH-ARG test with the best pair of values for sensitivity and specificity, accuracy, ROC AUC, PPV and negative predictive value (NPV), with confidence intervals (CI), in lean, overweight and obese populations; cut-off points for the sensitivity and specificity at 95% are also reported. |
| Lean (BMI < 25 kg/m²) | Overweight (BMI ≥ 25 and < 30 kg/m²) | Obese (BMI ≥ 30 kg/m²) |
| Cut-off point (µg/l) | 11.5 | 8 | 4.2 |
| Sensitivity (%; 95% CI) | 98.7 (92.9–99.8) | 96.7 (93.2–98.6) | 93.5 (82.1–98.6) |
| Specificity (%; 95% CI) | 83.7 (78.6–88.0) | 75.5 (71.1–79.5) | 78.3 (67.9–86.6) |
| Accuracy (%) | 87.1 | 78.6 | 83.9 |
| ROC AUC (95% CI) | 0.943 (0.913–0.965) | 0.922 (0.898–0.941) | 0.917 (0.855–0.958) |
| PPV (%; 95% CI) | 64.4 (0.55–0.72) | 65.9 (0.61–0.78) | 70.5 (0.60–0.97) |
| NPV (%; 95% CI) | 99.5 (0.97–0.99) | 97.9 (0.85–0.98) | 95.6 (0.86–0.97) |
| Cut-off point for sensitivity at 95% | 7.9 | 7.0 | 4.7 |
| Cut-off point for specificity at 95% | 2.3 | 1.5 | 1.0 |
overweight and obese populations were provided using the statistical method of ROC analysis (35). For each group the highest pair of values for sensitivity/specificity, the 95% sensitivity and the 95% specificity cut-off points were calculated. High sensitivity cut-points maximize detection of adult GHD, whereas high specificity cut-points minimize misclassification of normal subjects as GH-deficient. In clinical practice the cut-off points derived by ROC analysis in order to provide the best pair of values for sensitivity/specificity should be used commonly. Interestingly in our study these cut-off limits were found to have a sensitivity similar to the ideal value of 95%, specifically 98.7, 96.7 and 93.5% in lean, overweight and obese populations, respectively. In agreement with some clinicians (34), the 95% sensitivity cut-off point might be considered the most appropriate if patients have a clinical context compatible with GHD.

Present results show that the cut-off limits of peak GH response to the GHRH-ARG test obtained by ROC curve analysis with the best pair of sensitivity/specificity values to minimize misclassification of GHD and control subjects were 11.5, 8.0 and 4.2 \( \mu g/l \) for the lean, overweight and obese populations, respectively. The cut-off point of 11.5 \( \mu g/l \) for the lean population is intermediate between 16.5 and 9.0 \( \mu g/l \), which are the third and first centiles, respectively, of peak GH responses recorded in 157 lean normal subjects (16). The cut-off point of 4.2 \( \mu g/l \) that we found in the obese population is comparable with that (4.1 \( \mu g/l \)) reported by Biller et al. (34) in a population with a BMI of 30.3 \( \pm 5.8 \) kg/m\(^2\), consistent with obesity. The diagnostic reliability of the GHRH-ARG test and of these cut-off limits is confirmed by the good accuracy (87.1, 78.6 and 83.9% in the lean, overweight and obese populations, respectively) and by the high values of ROC AUC (0.943, 0.922 and 0.917 in the lean, overweight and obese populations, respectively).

However, it has to be stressed that the above-mentioned cut-off points referred to the GH assay used in this study. In fact, it is well known that when the same serum sample is tested in different assays, there is a wide variability in the absolute values reported (49, 50). Thus, when using other GH assays the above-mentioned cut-off limits for the peak GH response to the GHRH-ARG test would have to be reconsidered.

The significant inverse correlation that we found between peak GH responses to the GHRH-ARG test and BMI in control subjects was similar to that reported recently by Bonnert et al. (33). In patients no correlation was found and this finding may be explained by the presence of GH deficiency that reduced the peak GH response to the test in lean as well as in overweight and obese subjects.

IGF-I levels have been reported previously as high, normal or low in obese subjects (28, 33, 51–53). In our study we have not found significant difference in IGF-I levels as a function of BMI. However, as expected, we have found a significant difference in IGF-I levels between patients and controls and between TPHD and PHD patients.

Applying the cut-off points defined in the present study, among the patients without or with no more than two PHD (the PHD group) we found that 81.5 and 36.6% of overweight and obese patients, respectively, had GHD. This kind of diagnosis was difficult before, because the specific cut-off limits of peak GH response to the GHRH-ARG test in overweight and obese populations had not been calculated. Similarly to the difference in IGF-I levels recorded between patients with TPHD (the TPHD group) and controls, IGF-I levels were significantly lower in PHD patients who resulted to be GHD than in those without GHD. Moreover, a significant negative correlation between peak GH responses to the GHRH-ARG test and BMI was found in PHD patients without GHD, as well as in control subjects. Finally, it has to be noted that there was a higher percentage of other PHD in PHD patients who resulted to be GHD than in those without GHD. All these findings confirm the reliability of the GHRH-ARG test as diagnostic tool for GH deficiency.

In conclusion, as demonstrated in our previous studies in lean subjects, the GHRH-ARG test is also a reliable diagnostic tool in overweight or obese adult patients suspected of GHD, when appropriate cut-off limits related to BMI are considered. The availability of clear BMI-related cut-off limits is mandatory in common clinical practice in view of the fact that many hypopituitaric patients with suspected GHD are overweight or truly obese.

Acknowledgements

This study was supported by FSMEM and University of Turin, Italy, and was presented in part at the Endocrine Society’s 86th Annual Meeting, New Orleans, LA, USA, in June 2004. The authors wish to thank Dr G Migliaretti for his statistical assistance.

References

6. Carroll PV, Christ ER & Sonksen PH. Growth hormone replacement in adults with growth hormone deficiency: assessment of
38 Kelly DF, Gonzalez IT, Cohan P, Berman N, Swardoff R & Wang C. Hypopituitarism following traumatic brain injury and www.eje-online.org


Received 8 March 2005
Accepted 24 May 2005