Effect of Sex and Assay Method on Serum Concentrations of Growth Hormone in Patients with Acromegaly and in Healthy Controls

Helene Markkanen,1* Tuula Pekkarinen,2 Matti J. Välimäki,3 Henrik Alfthan,1 Ritva Kauppinen-Mäkelin,4 Timo Sane,3 and Ulf-Håkan Stenman1

Background: Diagnosis and follow-up of acromegaly is based on measurements of serum growth hormone (GH) concentrations during an oral glucose tolerance test (OGTT). A nadir value <1 µg/L is commonly used to define a normal response, but some authors suggest lower cutoff values.

Methods: To compare the results and subsequent patient classification obtained with 3 GH assays, we obtained basal serum samples from 78 apparently healthy adult controls (43 women and 35 men; median age, 32.5 years) and from 71 treated (44 women and 27 men; median age, 55.2 years) and 7 untreated acromegaly patients (4 women and 3 men; median age, 54.6 years), and OGTT was performed on all patients and on 72 of the 78 controls. GH was determined by 2 immunometric assays—a double monoclonal (AutoDELFIA; Wallac) and a monopolyclonal (Immulite 2000; DPC) assay—and in a limited set of samples by an RIA (Spectria RIA; Orion).

Results: There was a strong correlation (r = 0.995; P <0.001) between the 2 immunometric methods, but the results obtained with the Immulite 2000 were, on average, 1.4-fold higher than those obtained with the AutoDELFIA. At concentrations around the cutoff (1 µg/L), however, the difference was ~2-fold. Overall, the Orion RIA method also showed a good correlation (r = 0.951–0.959) with the other methods, but it did not measure concentrations <2 µg/L. Women had higher basal and OGTT nadir GH concentrations than men.

Conclusion: Reference intervals should be determined separately for each method, and the need for establishing sex-specific reference values should be investigated.

© 2006 American Association for Clinical Chemistry

Hypersecretion of growth hormone (GH)5 as the cause of acromegaly is confirmed by measuring serum GH during an oral glucose tolerance test (OGTT). Sensitive immunometric GH assays using monoclonal antibodies have been used to establish consensus guidelines for diagnosis of acromegaly (1, 2). Although the overall correlation between GH assays is good, differences greater than 2-fold between various methods have been observed (3, 4). Generally, competitive RIAs yield higher results than immunometric assays and do not reliably measure low GH concentrations (<1 µg/L), which is a significant disadvantage when evaluating the nadir GH concentrations during the OGTT (5). The detection limits of most immunometric methods are well below 1 µg/L, but there are also differences in the GH concentrations measured by various immunometric assays (6).

The heterogeneity of circulating GH is an important reason for the between-method discrepancies. In addition to the major form of GH [relative molecular mass (M_r) 22 000] and the splice variant with a relative molecular mass of 20 000, many other forms of GH occur in the circulation. Up to one half of GH in serum may be bound to 2 GH-binding proteins, which alter the metabolic clearance and biological activity of GH (7). GH in these complexes is differently recognized by various assays (8–10). Furthermore, the use of different standards (i.e.,

1 Department of Clinical Chemistry and 3 Division of Endocrinology, Department of Medicine, Helsinki University Hospital, Helsinki, Finland. 2 Department of Internal Medicine, Peijas Hospital, Vantaa, Finland. 4 Department of Internal Medicine, Jorvi Hospital, Espoo, Finland. 1* Address correspondence to this author at: Department of Clinical Chemistry, Helsinki University Hospital, FIN-00290 Helsinki, Finland. Fax 358-9-477-74906; e-mail helene.markkanen@helsinki.fi. Received September 9, 2005; accepted December 29, 2005. Previously published online at DOI: 10.1373/clinchem.2005.060236

5 Nonstandard abbreviations: GH, growth hormone; OGTT, oral glucose tolerance test; IS, International Standard; IRP, International Reference Preparation; and hGH, human growth hormone.
pituitary or recombinant GH) and differences in antibody specificity for the various GH isoforms contributes to the discrepancies (4, 11, 12). Most currently used assays are calibrated against the pituitary-derived International Standard (IS) 80/505 from the WHO. This and the older pituitary-derived International Reference Preparation (IRP 66/217) contain a range of GH variants, and their GH content has been estimated on the basis of bioactivity and expressed in international units (IU). The latest standards consisting of M, 22 000 recombinant human GH (IS 88/624 and IS 98/574) are primarily calibrated in mass units (13, 14).

In spite of the differences in results obtained by different assays, fixed decision limits are used in recommendations for establishment of diagnosis and for definition of remission in patients with acromegaly. In the consensus statement issued in 2000 (2), a nadir concentration of 1 µg/L in the OGTT, as measured by a sensitive immunometric assay, was used to define normalization of GH secretion, but this limit has recently been criticized for being too high (15). Differences in basal serum GH concentrations between healthy men and women have been observed in several studies (16–20), but the nadir value during the OGTT has differed in some (18, 19, 21) but not in other studies (5, 20).

We measured serum GH concentrations in treated and untreated acromegaly patients and in apparently healthy controls during OGTTs performed with 2 immunometric assays and compared patient classifications determined by the use of different decision limits, the consensus limit of 1 µg/L for diagnosis and remission of acromegaly, and tentative limits determined separately for the 2 assays on the basis of GH concentrations in male and female controls. Selected samples were also assayed by RIA.

### Materials and Methods

**PARTICIPANTS**

Serum samples were obtained from 78 apparently healthy controls (43 women and 35 men; median age, 32.5 years; range, 19.8–57.5 years) and from 71 treated acromegaly patients (44 women and 27 men; median age, 55.2 years; range, 32.2–79.8 years) and 7 patients with newly diagnosed untreated acromegaly (4 women and 3 men; median age, 54.6 years; range, 35.7–70.7 years; Table 1). In a limited substudy, we compared the Orion RIA with the immunometric assays, using serum samples (n = 40) obtained from 5 treated and 3 untreated acromegaly patients during OGTT.

The fertile healthy women were tested during the follicular phase of the menstrual cycle (days 7–11). None of the women received estrogens. Men and women had similar body mass index [median (range), 23.0 (19.7–30.5) kg/m² in women and 23.7 (20.9–29.4) kg/m² in men]. An OGTT with 75 g of glucose was performed on patients and controls after an overnight fast. A cannula was inserted into the cubital vein, and blood samples were taken at 0, 30, 60, 90, and 120 min. Serum samples were kept frozen for 4–6 weeks at −20 °C and then at −80 °C until assayed. The protocol was approved by the Ethics Committee of the Department of Medicine, Helsinki University Hospital, and written informed consent was obtained from all study participants.

### ASSAYS

The AutoDELFIA human GH (hGH) assay (Perkin-Elmer Wallac) is a solid-phase, 2-site immunofluorometric assay using 2 murine monoclonal antibodies that recognize various forms of M, 22 000, but not M, 20 000, variants. The Immulite 2000 assay (Diagnostic Products Corporation) is an immunochemiluminometric assay using a murine monoclonal antibody immobilized on a polystyrene bead and an alkaline phosphatase–labeled rabbit polyclonal antibody as tracer. The Orion Diagnostica Spectra RIA is a competitive-type RIA based on the use of goat anti-rabbit IgG–coated tubes, rabbit GH antiserum, and 125I-labeled hGH as a tracer. The AutoDELFIA and Immulite assays are calibrated against the hGH first IS 80/505, in which 1 mg = 2.6 IU of hGH, and the Orion RIA is calibrated against the first IRP 66/217, in which 1

### Table 1. Basal and nadir serum GH concentrations in controls and in treated and untreated acromegaly patients. a

<table>
<thead>
<tr>
<th>GH Concentration</th>
<th>Controls</th>
<th>Treated</th>
<th>Untreated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AutoDELFIA</strong></td>
<td><strong>Immulite 2000</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basal GH, µg/L</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>0.06 (0.02–3.43)</td>
<td>0.11 (0.03–4.81)</td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>0.99 (0.05–12.15)</td>
<td>1.96 (0.08–15.73)</td>
<td></td>
</tr>
<tr>
<td><strong>AutoDELFIA</strong></td>
<td><strong>Immulite 2000</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nadir GH, µg/L</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>0.02 (0.02–0.14)</td>
<td>0.05 (0.03–0.25)</td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>0.11 (0.02–0.71)</td>
<td>0.19 (0.03–1.30)</td>
<td></td>
</tr>
<tr>
<td><strong>AutoDELFIA</strong></td>
<td><strong>Immulite 2000</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basal GH, µg/L</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men (n = 27)</td>
<td>0.46 (0.08–3.50)</td>
<td>0.73 (0.09–5.92)</td>
<td></td>
</tr>
<tr>
<td>Women (n = 44)</td>
<td>1.24 (0.02–5.58)</td>
<td>1.86 (0.03–8.00)</td>
<td></td>
</tr>
<tr>
<td><strong>AutoDELFIA</strong></td>
<td><strong>Immulite 2000</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nadir GH, µg/L</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men (n = 3)</td>
<td>13.7 (8.04–44.6)</td>
<td>16.9 (9.62–62.1)</td>
<td></td>
</tr>
<tr>
<td>Women (n = 4)</td>
<td>19.2 (6.85–30.2)</td>
<td>26.4 (9.81–45.4)</td>
<td></td>
</tr>
</tbody>
</table>

*All results are expressed as the median (range).*
mg = 1.6 IU of hGH. Both standards are pituitary derived, consisting of a mixture of GH forms.

The intra- and interassay CVs for the Orion RIA at concentrations of 4.6, 13.6, and 32.5 μg/L were <6.7%, 2.4%, and 4.4%, respectively. The interassay CVs for the Immulite assay were 3.9% at 3.3 μg/L, 4.8% at 9.7 μg/L, and 3.2% at 21.4 μg/L. For the AutoDELFIA assay, the interassay CVs were 3.0% at 2.2 μg/L, 4.0% at 4.9 μg/L, and 4.8% at 12.5 μg/L.

Samples were measured in duplicate in the Orion assay, and in singlicate in the automatic AutoDELFIA and Immulite assays. According to the manufacturers, the detection limit of the Orion RIA is 0.38 μg/L (0.6 mIU/L), and for both the Immulite 2000 and AutoDELFIA, the detection limit is 0.01 μg/L (0.03 mIU/L). The lowest reported results are 0.63, 0.05, and 0.04 μg/L for the Orion RIA, Immulite, and AutoDELFIA assays, respectively.

**STATISTICAL ANALYSIS**

For comparison of different methods, all results are expressed in micrograms per liter. Values below the lowest reported result were assigned a value half of that. The nadir GH concentration was defined as the lowest value at any time point during the OGTT, and results are expressed as the median and range. We performed statistical analyses with Spearman correlation, Deming unweighted regression, and Wilcoxon signed-ranks test for comparing assay results, and Mann–Whitney and McNemar tests for comparing groups.

**Results**

**CORRELATION BETWEEN ASSAYS**

The results obtained by the AutoDELFIA and Immulite 2000 assays (n = 690) for controls and for treated and untreated acromegaly patients (Fig. 1) correlated strongly (r = 0.995; P < 0.001). The correlation equation based on Deming unweighted regression was:

\[ y \text{ (Immulite 2000)} = 1.37x \text{ (AutoDELFIA)} + 0.173 \text{ μg/L} \]

In a limited set of samples (n = 40), we compared results obtained by the Orion Spectria RIA with the results obtained by the AutoDELFIA and Immulite 2000. The AutoDELFIA assay consistently gave the lowest values and the RIA the highest. There was a strong correlation between all assays, but at concentrations <2 μg/L, the Orion assay showed no correlation with the other assays (Fig. 2). On the basis of slopes from the Deming regression analysis, the results obtained by the Orion assay for all samples were ~2.0-fold higher than those obtained by the AutoDELFIA assay and ~1.7-fold higher than those obtained by the Immulite 2000.

**GH CONCENTRATIONS IN CONTROLS**

We measured the basal and nadir GH concentrations in samples from 78 and 72 controls with the AutoDELFIA and samples from 66 and 60 controls with the Immulite, respectively. A total of 5 control samples, all from men [5 of the 66 (7.6%) tested with the AutoDELFIA and 4 of the 66 (6.1%) tested with the Immulite assay], had a basal concentration below the lowest reported result. Women controls had higher GH concentrations (P < 0.001) than men in both basal and nadir samples (Table 1). In control samples, the nadir GH was below the lowest reported result for 32% (19 of 60; 17 men and 2 women) with the AutoDELFIA and for 18% (11 of 60; 10 men and 1 woman)
with the Immulite assay. In the concentration range <2 μg/L, the measured concentrations obtained with the Immulite assay were ~2-fold higher than those obtained with the AutoDELFIA (Table 1). A nadir GH >1 μg/L was not found in any controls by the AutoDELFIA (Fig. 3) but was found in 2 female controls by the Immulite. In the age range studied (19.8–57.5 years), there was no significant age dependence in GH concentrations.

GH CONCENTRATIONS IN TREATED AND UNTREATED ACROMEGALY PATIENTS

To evaluate how different decision limits affect patient classification, we compared limits based on the highest assay- and sex-specific nadir GH values observed in controls during an OGTT and used the consensus decision limit of 1 μg/L to define remission in patients with treated acromegaly (2) (Table 2). With the consensus cutoff limit, more treated acromegaly patients were classified as being in remission by the AutoDELFIA (68% of women and 85% of men) than by the Immulite assay (55% and 82%, respectively). The difference was statistically significant in women (P = 0.031). When we used sex-specific nadir values from controls to define decision limits, the cutoff decreased substantially for men, and this reduced the proportion of acromegaly patients who would have been classified as cured. With AutoDELFIA results, the “remission or cure rate” in men decreased from 85% to 48% (P = 0.002), and with the Immulite results, it decreased from 82% to 52% (P = 0.008). In women the impact of sex-specific decision limits was small; with AutoDELFIA, the proportion decreased from 68% to 57% (P = 0.063), and with the Immulite, it increased from 55% to 61% (P = 0.250). When sex- and assay-specific cutoff values were used for the Immulite and AutoDELFIA methods, there were no statistically significant differences in the proportion of patients classified as cured.

For all methods, all untreated acromegaly patients had increased nadir GH values (>3 μg/L) in the OGTT. Thus, the decision limit is not critical for the diagnosis of new cases.

**Table 2. Classification of treated acromegaly patients (n = 71) considered to be in remission on the basis of the consensus statement nadir value (1 μg/L) and the assay-specific cutoff values based on the highest nadir values in the control groups of women and men, respectively.**

<table>
<thead>
<tr>
<th>Assay and sex-specific cutoffs</th>
<th>Women</th>
<th>Men</th>
</tr>
</thead>
<tbody>
<tr>
<td>AutoDELFIA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cutoff</td>
<td>≤1 μg/L</td>
<td>≤1 μg/L</td>
</tr>
<tr>
<td>n (%)</td>
<td>30 (68.2)</td>
<td>23 (85.2)</td>
</tr>
<tr>
<td>Immulite 2000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cutoff</td>
<td>≤1 μg/L</td>
<td>≤1 μg/L</td>
</tr>
<tr>
<td>n (%)</td>
<td>24 (54.5)</td>
<td>22 (81.5)</td>
</tr>
</tbody>
</table>

![Fig. 3. Nadir GH concentration during OGTT determined by the AutoDELFIA assay in 38 women (○) and 34 men (△) of various age.](image-url)

**Discussion**

Our results show that the correlation between the GH assays used was good, but the measured concentrations were different. The Immulite 2000 gave 37% higher and the Orion assay 109% higher results than the AutoDELFIA. The differences in calibration of the 3 assays can be explained in part by differences in recognition of the M<sub>4</sub> 20 000 form and calibration against different standards. The Immulite and AutoDELFIA assays use the same standard, but the lower results obtained with AutoDELFIA can be explained only partly by lack of recognition of the M<sub>4</sub> 20 000 isof orm; differences in recognition of complexes with the GH-binding protein may also have affected the results (9). The higher values obtained by the Orion Spectria RIA can be explained mainly by the use of IRP 66/217 and the different factors used to convert the values for this standard from mass units (mg/L) to international units (IU/L), 1.6 for IRP 66/217 vs 2.6 for IS 88/505, which is used in the 2 other methods. Thus the results for the Orion assay at concentrations >2 μg/L would have been similar to those obtained by the Immulite assay if the values had been expressed in IU/L rather than the more commonly used mass units.

In addition to the differences in calibration, there were larger differences in GH values at low concentrations. The Orion RIA did not measure concentrations <1 to 2 μg/L; thus, it was not useful for measurement of nadir GH values during an OGTT. The difference in values measured by the 2 immunometric assays also increased at low values. At these concentrations, the difference was ~2-fold, and the cutoff values calculated by use of the slope derived by regression analysis as a conversion factor...
would not have been valid for nadir values during an OGTT. Thus, it is necessary to establish assay-specific cutoff limits based on actual values measured by each method.

In agreement with earlier studies, our results show that healthy women have significantly higher GH concentrations than men, both in the basal state and during OGTT (16–18, 21). Chapman et al. (18) and Costa et al. (21) reported a sex-dependent difference in nadir values, but this was not observed by Freda and coworkers (5, 20). Gullu et al. (19) observed a sex difference only in the third decade of life. These studies used immunometric assays with reported detection limits of 0.002 to 0.06 µg/L, similar to those of our immunometric assays.

The nadir value during OGTT is commonly used to establish the diagnosis and to evaluate the response to therapy in acromegaly. According to the consensus statement, the nadir GH concentration measured by immunometric assays with a detection limit of at least 0.5 µg/L should be <1 µg/L for acromegaly to be excluded or to be considered cured (2). It has also been suggested that the cutoff value should be lowered further, to 0.14–0.30 µg/L (5, 15, 20, 22, 23). Our results indicate that these cutoffs would be too low for women, but a lower decision limit is justified for men. Furthermore, the limit should be assay dependent. To date, different decision limits have not been used for women and men or for different assays. Our results indicated that many male acromegaly patients may be classified erroneously as cured according to the nadir value in the OGTT with a cutoff of 1 µg/L. Depending on the assays used, 82%–85% of the treated acromegalic men were classified as being in remission with this cutoff value, compared with 48%–52% with assay- and sex-specific cutoff values. The corresponding differences for women were much smaller because the assay-specific cutoff values were close to 1 µg/L, i.e., 0.71 and 1.30 µg/L, respectively. It should be noted that these cutoffs are provisional; the number of controls was too small to establish reliable decision limits, but it is obvious that both basal and nadir values during OGTT are significantly lower in men than in women. Furthermore, it has been shown that GH concentrations decrease with age (16, 24). Our controls were younger (median age, 33 years) than the treated acromegalic patients (median age, 55 years). Thus it is possible that cutoff values should be lower in older than in younger patients. We did not find a significant age dependence in GH values, but our control population was too small with respect to the older age groups for evaluation of this aspect. Most of our female controls were of fertile age, and estrogens have been shown to increase serum GH (16), but menopausal status and hormone replacement have not been shown to affect nadir GH concentrations (20).

The prognosis of acromegaly patients is related to the GH concentration after therapy: those with higher concentrations having a shorter overall life expectancy and disease-specific survival (25–28). However, because sensitive GH assays have become widely used only recently, there are not yet sufficient data available from long-term follow-up studies to define the optimal cutoff determined by sensitive assays. The number of controls subjected to OGTT in our study was limited, and with respect to older age groups in particular, the provisional reference values derived from this study need to be confirmed in larger studies.

Establishment of reference values for the OGTT is quite demanding, and collaboration between different centers may be necessary. Furthermore, it remains to be investigated whether reference limits determined in healthy controls are optimal decision limits for evaluation of treatment response in acromegaly patients. Our findings demonstrate the need to establish decision limits according to the characteristics of the assay used (2).

In conclusion, we found that various GH assays give different results, particularly at low concentrations, and that GH concentrations are lower in men than in women. Establishing assay- and sex-specific decision limits for serum GH would therefore be useful to guide treatment of acromegaly patients. Improved standardization of GH assays would facilitate establishment of method-independent decision limits, but because of the heterogeneity of GH in the circulation, perfect standardization may be hard to achieve.

This study was supported by grants from the Finska Läkaresällskapet. We thank Perkin-Elmer Wallac for supplying the GH reagents used in this study.

References

9. Hansen TK, Fisker S, Hansen B, Sørensen HH, Christiansen JS,