Which Patients Do Not Require a GH Stimulation Test for the Diagnosis of Adult GH Deficiency?

MARK L. HARTMAN, BRENDA J. CROWE, BEVERLY M. K. BILLER, KEN K. Y. HO, DAVID R. CLEMMONS, AND JOHN J. CHIPMAN, ON BEHALF OF THE HYPOCCS ADVISORY BOARD AND THE U.S. HYPOCCS STUDY GROUP

Lilly Research Laboratories (M.L.H., B.J.C., J.J.C.), Eli Lilly & Co., Indianapolis, Indiana 46285; Massachusetts General Hospital (B.M.K.B.), Boston, Massachusetts 02114; Garvan Institute of Medical Research, St. Vincent's Hospital of Sydney (K.K.Y.H.), Sydney, NSW 2064, Australia; and University of North Carolina (D.R.C.), Chapel Hill, North Carolina 27599

Adult GH deficiency (GHD) is currently diagnosed in patients with either a history of childhood-onset GHD or acquired hypothalamic-pituitary disease by GH stimulation testing. However, GH stimulation tests are invasive, time consuming, and associated with side effects. Based on preliminary analyses of patients enrolled in the U.S. Hypopituitary Control and Complications Study (HypoCCS), we proposed the presence of adult GHD could be predicted with 95% accuracy by the presence of three or more pituitary hormone deficiencies (PHDs) or a serum IGF-I concentration less than 84 μg/liter (11 nmol/liter). To validate the diagnostic utility of these criteria, we studied results obtained in 817 adult patients (mean [SD] age: 46.4 [15.7] yr, body mass index: 30.1 [7.2] kg/m^2) enrolled in HypoCCS who had serum GH concentrations from stimulation tests (11 different tests used, excluding clonidine) and serum IGF-I (competitive binding RIA) measured at the central laboratory (Esoterix Endocrinology, Calabasas Hills, CA). When patients were stratified into subgroups on the basis of the presence of zero, one, two, three, and four additional PHDs, median (25th, 75th percentile) peak GH levels (micrograms per liter) were 3.5 (0.85, 7.1), 0.73 (0.18, 4.2), 0.29 (0.05, 1.4), 0.06 (0.025, 0.295), and 0.025 (0.025, 0.07), respectively. The

mean log (peak GH) concentration was significantly different among the subgroups ($P < 0.05$). The proportion of patients in each group with severe GHD diagnosed by stimulation testing (peak GH < 2.5 μg/liter) was 41%, 67%, 83%, 96%, and 99% for patients with zero, one, two, three, and four PHDs, respectively. The positive predictive values (PPVs) for GHD of three PHDs, four PHDs, and serum IGF-I less than 84 μg/liter were 96%, 99%, and 96%, respectively. The PPV of these three diagnostic criteria was also 95% or more after excluding the data originally used to identify these potential predictors. Taken together, the presence of either three or four additional PHDs or IGF-I less than 84 μg/liter (55% of the patients met at least one of these criteria) reliably predicted GHD with a high PPV (95%), high specificity (89%), and moderate sensitivity (69%). We concluded that patients with an appropriate clinical history and either the presence of three or four additional PHDs or serum IGF-I less than 84 μg/liter (measured in the Esoterix assay) do not require GH stimulation testing for the diagnosis of adult GHD. In clinical practice, we suggest that other causes of low serum IGF-I should be excluded before applying these diagnostic criteria. (*J Clin Endocrinol Metab* 87: 477–485, 2002)

THE DIAGNOSIS OF adult GH deficiency (GHD) is established by clinical and biochemical criteria. For the diagnosis of adult-onset GHD, patients are selected on the basis of a history of acquired disease, surgery, radiation therapy, or trauma involving the hypothalamus or pituitary gland. Adults with a history of childhood GHD must have the diagnosis of GHD confirmed as an adult (1, 2). The current consensus of regulatory agencies (3, 4) and scientific societies (1, 2) is that these patients with appropriate clinical and historical characteristics should then have the diagnosis of adult GHD confirmed by a provocative test of GH secretion. The suggested peak GH cut-point for the diagnosis of adult GHD varies from either less than 3 or 5 μg/liter on an insulin tolerance test (ITT) (1, 2, 5) to less than 5 μg/liter, irrespective of what GH stimulation test is used (3, 4).

However, problems exist with the implementation of these diagnostic criteria in clinical practice (6). The recommended

"gold standard" GH stimulation test is the ITT, which is associated with uncomfortable side effects and can be dangerous in patients with a history of coronary artery disease or seizure disorders. The need to monitor the patient constantly during insulin-induced hypoglycemia makes this test resource intensive and difficult to perform in some office practice settings. Therefore, many other GH stimulation tests are employed in clinical practice without adjusting the peak GH criteria, despite the known variability of the GH response to these various provocative agents. Significant cost savings could be achieved if a simpler diagnostic approach could be identified, even for a subgroup of patients.

Two other approaches have been suggested for the diagnosis of adult GHD: the use of clinical history or serum IGF-I concentrations. In patients with organic hypothalamic-pituitary disease, the likelihood of GHD increases with increasing numbers of pituitary hormone deficiencies (PHDs) (1). Four studies have reported that the probability of GHD (peak GH criteria ranging from < 2.3 to < 5 μg/liter) in patients with three to four PHDs ranges from 91% to 100% (7–10). In addition, patients with very low values of serum IGF-I in the absence of other causes also have a high probability of GHD (1). However, no algorithms for either of these approaches to

the diagnosis of adult GHD have been validated prospectively in a large number of patients.

In the present study, we investigated the predictive value of three or more PHDs or a subnormal IGF-I concentration as diagnostic predictors of adult GHD in patients with a history of hypothalamic-pituitary disease or childhood-onset GHD. We sought to identify an IGF-I cut-point below which the probability of GHD would be 95% or more. Preliminary analyses of patients enrolled in the U.S. Hypopituitary Control and Complications Study (HypoCCS) indicated that adult GHD could be predicted with 95% accuracy by the presence of either three or more PHDs or a serum IGF-I concentration less than 84 μg/liter (11 nmol/liter), as measured by Esoterix Endocrinology (Calabasas Hills, CA). The purpose of the present study was to validate these potential predictors of adult GHD in a large series of patients with appropriate clinical histories.

Materials and Methods

Subjects and study design

Subjects included in this study were enrolled in HypoCCS, a postmarketing safety surveillance study of the long-term outcomes of GH (Humatrope [somatropin of rDNA origin], Eli Lilly & Co., Indianapolis, IN) replacement therapy in adults with GHD. The study was conducted in the offices of 151 endocrinologists in the United States. Data collected on case report forms were verified against the original source data by monitors performing on-site chart reviews. Additional data validation procedures were performed after data entry and queries were sent to the sites to resolve any remaining issues. Each study site obtained approval of the protocol by the local institutional review board. Each patient provided written informed consent before enrolling in the study. The patients recruited for the study had a history of either adult-onset hypothalamic-pituitary disease or childhood-onset GHD. All patients who entered this study had a clinical history obtained by the investigator and had a GH stimulation test performed using a standard pharmacological provocative agent if this had not already been done. To continue in HypoCCS, patients must have had a peak GH concentration of less than 5 μg/liter during the GH stimulation test. Other PHDs were diagnosed by the participating investigators as part of their clinical practice. Here we report the results of diagnostic testing from 817 adult patients (mean [SD] age: 46.4 [15.7] yr, body mass index [BMI]: 30.1 [7.2] kg/m^2) enrolled in HypoCCS who had serum IGF-I and GH concentrations measured at the central laboratory. Patients who were not eligible to continue in HypoCCS because of a peak GH more than 5 μg/liter were included in this analysis of diagnostic test results. These 817 patients were a subset of the 1550 patients with validated case report forms from the baseline visit received as of December 31, 1999. Patients were excluded from this analysis if the serum IGF-I and GH concentrations were measured in local laboratories because this would increase the variability of the results. Results from certain stimulation tests were also excluded, as described below.

Hormonal assays

All samples were analyzed in duplicate at the central laboratory (Esoterix Endocrinology). Serum GH concentrations were measured using an immunochemiluminometric assay (ICMA) specific for 22-kD human GH with a sensitivity of 0.05 μg/liter (11). The intra- and interassay coefficients of variation (CV) ranges were 3.8% to 9.1% and 8% to 10%, respectively, for a quality control range of 0.3–20 μg/liter. Samples higher than 20 μg/liter were repeated on dilution. This assay is calibrated against the World Health Organization 80/505 international GH standard (human pituitary-derived GH) but uses native-sequence recombinant human GH as standard (Eli Lilly & Co.). This method yields results that are on average half of those obtained with a polyclonal RIA (Stene M., Esoterix Endocrinology, personal communication). Serum IGF-I concentrations were measured in a highly specific competitive binding RIA after acid-ethanol extraction (12). The assay uses native-

sequence IGF-I (Bachem, Torrance, CA) as standard but is standardized 16% higher than mass correct IGF-I (Genentech, Inc., South San Francisco, CA). IGF-II is added to each assay tube to eliminate potential interference from residual low-molecular-weight IGF-binding proteins. The assay sensitivity was 12.9 μg/liter. The intra- and interassay CV ranges were 4.1% to 6.5% and 6.6% to 8.4% for a quality control range of 60 μg/liter to 500 μg/liter, respectively (Stene M., Esoterix Endocrinology, personal communication).

GH stimulation tests

GH stimulation tests were employed based on the individual investigator's choice. Results obtained from clonidine stimulation tests (n = 62) were excluded from this analysis because this test has been shown to be inadequate for stimulation testing in adults (13). Patients whose test type was not specified were also excluded (n = 24). The GH stimulation tests performed in the 817 patients included in this study are shown in Table 1. For the purposes of this analysis, adult GHD was defined as a peak GH response less than 2.5 μg/liter in response to a stimulation test. This strict definition was adopted for two reasons: 1) 11 different provocative tests were employed, which have varying potencies for stimulation of GH secretion (6); and 2) the criterion for adult GHD of a peak GH less than 5 μg/liter during a stimulation test was developed in older studies that employed polyclonal RIAs to measure serum GH concentrations (5). Adjustment of peak GH cut-off values for assay differences has been recommended by international consensus guidelines for the diagnosis of adult GHD (1). Because the GH ICMA used in this study yields serum GH concentrations that are on average approximately half of those obtained with a polyclonal RIA (Stene M., Esoterix Endocrinology, personal communication), a criterion of less than 2.5 μg/liter was adopted. This definition is consistent with the adult GHD indication for Humatrope approved by the U.S. Food and Drug Administration (FDA).

Statistical analyses

Three interim analyses of the HypoCCS study were performed to identify initially potential predictors of GHD and then to validate the diagnostic performance of these predictors on independent sets of data. A predictor of adult GHD was considered to be validated if the positive predictive value (PPV; see below for definition) was at least 95% on both an initial data set and on a second, independent set of data. These independent data sets were created by excluding the data that were initially used to identify the predictor in an earlier analysis. This criterion for validation was prospectively specified after the first analysis in 1998. The first analysis was performed in April 1998 when data from 162 patients were available. The second analysis was performed in January 1999 when data from 395 patients were available. The third analysis was performed in January 2000 with the final data set of 817 patients. The first analysis identified four PHDs and a serum IGF-I less than 84 μg/liter as potential predictors of GHD (defined by peak GH < 2.5 μg/liter), and these predictors were then validated with the second analysis. The second analysis identified three PHDs as a potential predictor of adult GHD, and this was validated with the third analysis. Results are reported for both the entire group of 817 patients available at the January

TABLE 1. GH stimulation tests performed in the 817 patients included in this study

GH stimulation test	No. of patients	% of total
Arginine	292	35.7
L-Dopa	207	25.3
L-Dopa/propranolol	104	12.7
ITT	93	11.4
Arginine/L-dopa	55	6.7
L-Dopa/clonidine	26	3.2
Insulin/GHRH	25	3.1

Tests employing two provocative agents are shown as agent 1/agent 2. Other tests performed with a frequency of 1% or less included glucagon, arginine/clonidine, arginine/glucagon, and GHRH.

2000 analysis as well as the smaller groups of patients used for validation of the predictors of GHD.

Values of peak GH concentrations that were less than 0.05 μg/liter were imputed to be 0.025 μg/liter and values of IGF-I that were below the quantifiable limit of the assay were imputed to be half of the limit for the statistical analyses. This approach was adopted because patients with pituitary disease have been reported to have nadir GH concentrations as low as 0.002 μg/liter and peak GH concentrations as low as 0.013 μg/liter in an ultrasensitive GH assay (14). Thus, half of the detection limit is a better approximation of the true concentration than the detection limit itself. Simple least-squares regression was used to test for relationships among peak GH values, serum IGF-I levels, age, and BMI. The analysis of these relationships was restricted to data collected at the same visit. Peak GH concentrations were log transformed before analysis to stabilize the variance. Median and 25th and 75th percentiles of peak GH concentrations were calculated for each subgroup of "number of additional PHDs." Because there were substantial differences in the variances of the raw peak GH values between these subgroups, an ANOVA was performed on log-transformed peak GH values to test for differences among subgroups; Tukey's test was used to keep the family error for all-pairwise comparisons at a rate of 0.05. For all correlations of variables with IGF-I, an outlying IGF-I value of 740 μg/liter was deleted before calculating the correlation coefficient and associated P value.

Sensitivity, specificity, and PPV were calculated for the identified predictors of adult GHD using the numbers of patients with true positive (TP), true negative (TN), false positive (FP), and false negative results for each predictor (15). Where appropriate, missing values were deleted before these were calculated. Sensitivity was defined as the percentage of patients with GHD (peak GH < 2.5 μg/liter) who were detected by the predictor (calculated as TP/[TP + false negative]). Specificity was defined as the percentage of non-GHD patients (peak GH ≥ 2.5 μg/liter) who were correctly identified by the predictor as non-GHD (calculated as TN/[TN + FP]). PPV was defined as the likelihood that a patient with a positive test (presence of a predictor) had a peak GH less than 2.5 μg/liter (calculated as TP/[TP + FP]).

Results

Initial analysis to identify predictors of adult GHD

A preliminary analysis of 162 patients enrolled in HypoCCS was performed in 1998 to identify potential predictors of adult GHD. The four PHDs considered in this and all subsequent analyses were: 1) TSH deficiency; 2) ACTH deficiency; 3) gonadotropin deficiency (LH and/or FSH deficiency were counted as one deficiency); and 4) arginine vasopressin (AVP) deficiency (central diabetes insipidus). PRL deficiency was not included in the analysis. Among 55 patients with three PHDs, 53 (96%) had a peak GH less than 2.5 μg/liter. All of the 25 patients with four PHDs had a peak GH less than 2.5 μg/liter. Analysis of the serum IGF-I results revealed that 77 of 78 patients (99%) with a serum IGF-I of less than 84 μg/liter had a peak GH less than 2.5 μg/liter. These three potential predictors of adult GHD were then validated prospectively in a larger number of patients as more patients were enrolled in HypoCCS; these results are now reported in detail.

Relationships among peak GH, IGF-I, age, and BMI

Figure 1 demonstrates that the log of peak GH was highly correlated with serum IGF-I in the entire cohort of 817 patients ($r = 0.60$, $P < 0.0001$). However, in patients who demonstrated no response to a GH stimulation test (peak GH < 0.05 μg/liter), a broad range of serum IGF-I concentrations was observed, ranging from below the limit of assay sensitivity to 177 μg/liter. The relationships between the log of
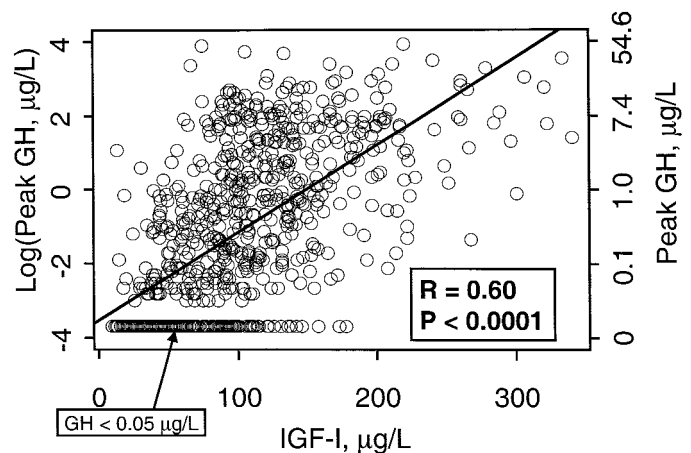

FIG. 1. Simple least-squares regression of the relationship between the log of peak GH concentration on standard GH stimulation tests and serum IGF-I concentrations in the entire cohort of 817 patients. To convert serum IGF-I from μg/liter to nmol/liter, multiply by 0.13.
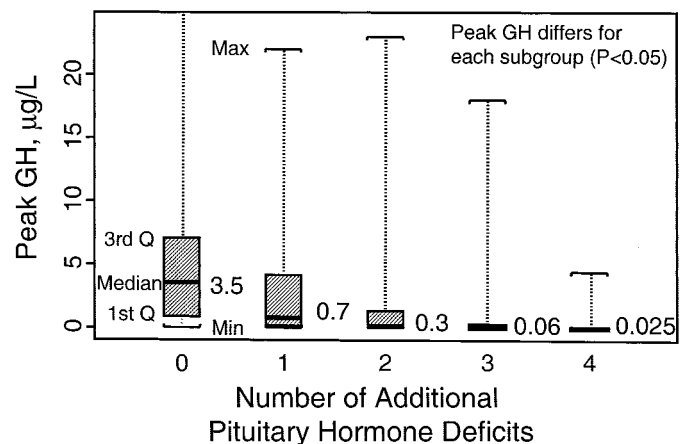

FIG. 2. Median, 25th (first quartile, first Q) and 75th (third quartile, third Q) percentiles of peak GH concentrations on standard GH stimulation tests for each subgroup of number of additional pituitary hormone deficiencies. The minimum (Min) and maximum (Max) values are also plotted, but not all values are visible on this scale. The peak GH level for each group differed significantly from that observed in the other groups ($P < 0.05$).

peak GH and age ($r = -0.21$, $P = 0.0001$) and the log of peak GH and BMI ($r = -0.24$, $P = 0.0001$) were weaker, although statistically significant. Similarly, the relationships between serum IGF-I concentrations and age ($r = -0.33$, $P = 0.0001$) and serum IGF-I and BMI ($r = -0.10$, $P = 0.007$) were weak, although statistically significant.

Relationship between peak GH and the number of PHDs

The distribution of peak GH concentrations according to the number of additional PHDs is shown in Fig. 2. The median value for peak GH concentration decreased with each additional PHD, being 3.5, 0.7, 0.3, 0.06, and 0.025 μg/liter for zero, one, two, three, and four PHDs, respectively. The peak GH level for each group differed significantly from that observed in the other groups ($P < 0.05$). Figure 3 illustrates the number of patients within each subgroup of number of PHDs, divided by whether the peak GH concentration
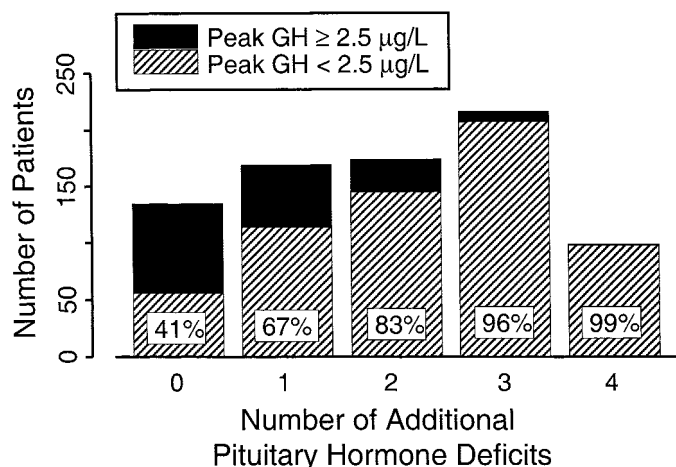
FIG. 3. The number of patients within each subgroup of number of pituitary hormone deficiencies divided by whether the peak GH concentration was less than 2.5 μg/liter or 2.5 μg/liter or greater on a standard GH stimulation test. The percentage of patients within each subgroup with a peak GH level less than 2.5 μg/liter is shown.

FIG. 4. Serum IGF-I concentrations for individual patients plotted against the number of additional pituitary hormone deficiencies. In the *left panel* of the figure, results are shown for those patients who had peak GH less than 2.5 μg/liter on GH stimulation testing. In the *right panel*, results are shown for those patients who had peak GH 2.5 μg/liter or greater. To convert serum IGF-I from μg/liter to nmol/liter, multiply by 0.13.

was less than 2.5 μg/liter or 2.5 μg/liter or more. The percentage of patients with peak GH less than 2.5 μg/liter was 41%, 67%, 83%, 96%, and 99% for patients with zero, one, two, three, and four additional PHDs, respectively.

Relationships among IGF-I, peak GH, and number of additional PHDs

In Fig. 4, serum IGF-I concentrations for individual patients are plotted against the number of additional PHDs, with results divided by whether the peak GH concentration was less than 2.5 μg/liter or 2.5 μg/liter or more. There was a wide spectrum of serum IGF-I levels for patients meeting the strict criterion for adult GHD. The vast majority of patients (94%) with peak GH 2.5 μg/liter or more had serum IGF-I greater than 84 μg/liter (11 nmol/liter), which was the diagnostic cut-point established in our initial preliminary analysis of 162 patients in 1998. An even greater percentage of patients (96%) with IGF-I values below the cut-point of 84 μg/liter had peak GH less than 2.5 μg/liter. A total of 11 patients had a serum IGF-I less than 84 μg/liter and a peak GH 2.5 μg/liter or more; three of these patients are not illustrated in the figure because the number of additional PHDs was not provided by the investigator. These latter three patients had peak GH greater than 5 μg/liter and hence did not qualify for the protocol. This is the most likely explanation for the missing clinical data.

Prediction of adult GHD by number of PHDs and serum IGF-I

Among the 793 patients for whom clinical data about the number of PHDs were available, 216 (27%) had three PHDs and 99 (13%) had four PHDs. The presence of three or four PHDs had PPVs of 96% and 99%, respectively, for adult GHD, defined by a peak GH less than 2.5 μg/liter. The PPV for both predictors of adult GHD was also 95% or greater after excluding the data originally used to identify these potential predictors. The specificity of these predictors was high (95% and 99%, respectively), indicating that patients
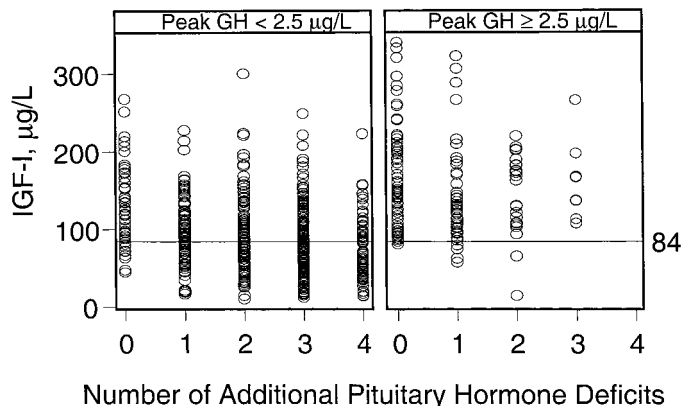
who were not GHD would be classified incorrectly as GHD at most 5% of the time. In contrast, the sensitivity of these predictors was very low (33% and 16% for three and four PHDs, respectively) because many patients with fewer than three PHDs are GH deficient. Table 2 provides the clinical characteristics of the 10 patients who had three or four PHDs but had peak GH 2.5 μg/liter or more during a GH stimulation test. In contrast, 305 of 315 (97%) patients who had three or four PHDs had peak GH less than 2.5 μg/liter.

Among the 785 patients for whom a serum IGF-I concentration was available at baseline, 40% had a serum IGF-I less than 84 μg/liter (11 nmol/liter). A serum IGF-I of less than 84 μg/liter had a PPV of 96% for adult GHD, defined by a peak GH less than 2.5 μg/liter. The PPV for this predictor of adult GHD was 97% after excluding the data originally used to identify 84 μg/liter as a potentially useful cut-point for the diagnosis of adult GHD. The specificity of this predictor was also high at 94%, but the sensitivity was low at 50%. Thus, many patients with a serum IGF-I level above this cut-point had GHD. Table 3 provides the clinical characteristics of the 11 patients who had serum IGF-I less than 84 μg/liter but who had a peak GH 2.5 μg/liter or more during a GH stimulation test. By way of comparison, 300 of 311 (96%) patients with serum IGF-I less than 84 μg/liter had peak GH less than 2.5 μg/liter, and 172 of 474 (36%) patients with serum IGF-I greater than 84 μg/liter had peak GH 2.5 μg/liter or more.

Taken together, the presence of either three or four PHDs or a serum IGF-I less than 84 μg/liter (11 nmol/liter) yielded a PPV for adult GHD of 95%. This diagnostic approach had a specificity of 89% and a sensitivity of 69%. Thus, many patients who did not meet these criteria were GH deficient on GH stimulation testing. However, 55% of the patients in the present analysis met at least one of these diagnostic criteria. The PPV for this combination rule was also 95% after excluding the data originally used to define this diagnostic rule.

We considered other combinations of predictors for adult GHD. The PPV and specificity for adult GHD may be in-

TABLE 2. Ten patients who had three or more PHDs but who had a peak GH of 2.5 μg/liter or greater during a GH stimulation test

Sex	Age (yr)	Onset type	BMI (kg/m^2)	Clinical diagnosis	No. of PHD	Estrogen therapy (women)	Serum IGF-I (μg/liter^a)	GH stimulation test	Peak GH (μg/liter)
F	42.8	AO	22.9	Idiopathic	3	Yes (td)	113	LD/CL	10.0
F	52.8	AO	24.3	Pituitary adenoma	3	Yes (po)	MD	LD/PR	6.7
F	55.3	AO	27.8	Idiopathic	3	Yes (po)	197	ARG/LD	2.9
M	39.8	AO	29.6	Malignant hyperthermia	3	NA	136	ARG	5.2
M	43.2	AO	32.3	Pituitary adenoma	3	NA	166	ARG	7.8
M	31.2	AO	29.6	Pituitary adenoma	3	NA	167	LD/PR	6.4
M	43.7	AO	22.5	Pituitary adenoma	3	NA	107	ITT	3.2
M	45.7	AO	24.2	Idiopathic	3	NA	137	ARG	18.0
M	17.7	CO	17.8	Radiation, chemotherapy (leukemia)	3	NA	265	ARG	15.0
M	34.6	CO	N/A	Craniopharyngioma	4	NA	MD	ITT	4.5

AO, Adult-onset; CO, childhood-onset; td, transdermal; po, oral; LD, L-dopa; CL, clonidine; PR, propranolol; ARG, arginine; NA, not applicable; MD, missing data.

^{a} To convert serum IGF-I from μg/liter to nmol/liter, multiply by 0.13.

TABLE 3. Eleven patients who had serum IGF-I less than 84 μg/liter but who had a peak GH at least 2.5 μg/liter during a GH stimulation test

Sex	Age (yr)	Onset type	BMI (kg/m^2)	Clinical diagnosis	No. of PHD	Estrogen therapy (women)	Serum IGF-I (μg/liter^a)	GH stimulation test	Peak GH (μg/liter)
F	52.8	AO	28.6	Idiopathic	1	No	79	ARG	5.0
F	49.7	AO	36.0	Idiopathic	1	Yes (po)	79	LD	5.9
F	51.5	AO	39.7	Pituitary adenoma	1	No	57	ARG/LD	2.9
F	32.3	N/A	25.9	Pituitary adenoma	DNR	Yes (po)	77	ARG/LD	6.5
F	36.4	AO	33.1	Temporal lobe tumor	1	Yes (po)	74	ARG	11.0
F	47.7	DNR	18.2	DNR	DNR	Yes (po)	67	ARG	29.0
F	58.2	DNR	16.1	DNR	DNR	Yes (po)	74	ARG	49.0
F	76.0	AO	29.6	Idiopathic	1	Yes (po)	62	ARG	6.8
F	47.4	AO	22.9	Idiopathic	2	Yes (po)	65	ARG	5.8
F	49.1	AO	45.4	Empty sella	0	Yes (po)	80	LD	4.3
M	27.7	CO	20.0	Cranial radiation	2	NA	14	ARG	2.9

AO, Adult-onset; CO, childhood-onset; po, oral; ARG, arginine; LD, L-dopa; DNR, data not recorded (early in the trial this information was not required for patients that were not GH deficient by stimulation testing); NA, not applicable.

^{a} To convert serum IGF-I from μg/liter to nmol/liter, multiply by 0.13.

creased to 100% if one requires the combination of three or four PHDs and a serum IGF-I less than 84 μg/liter. However, the sensitivity for this combination was very low at 30%; only 23% of the patients in the present analysis met these criteria. Similarly, the combination of two or more PHDs and a serum IGF-I less than 84 μg/liter yielded a PPV of 99% and a specificity of 99%, but the sensitivity was again low at 41%; only 33% of the patients in the present analysis met these criteria. The sensitivity for adult GHD can be increased to 81% with the combination of either two or more PHDs or a serum IGF-I less than 84 μg/liter. However, the PPV and specificity for this combination were only 91% and 76%, respectively. In the present analysis, 68% of the patients met these latter criteria.

We also explored the impact of changing the diagnostic criteria to a peak GH less than 5 μg/liter. In this case, the PPV of three PHDs, four PHDs, and serum IGF-I less than 84 μg/liter were 97%, 100%, and 97%, respectively. The combination of either three or four PHDs or serum IGF-I less than 84 μg/liter had a PPV of 97%, a specificity of 89%, and a sensitivity of 64%.

Predictive value of specific types of PHD

Among patients with three PHDs (n = 216), the most common combinations of PHDs were: 1) ACTH deficiency, TSH deficiency, and gonadotropin deficiency (n = 189); and 2) TSH deficiency, gonadotropin deficiency, and AVP defi-

ciency (n = 21). The PPV for prediction of a peak GH value less than 2.5 μg/liter for these two combinations of three PHDs were 96% and 95%, respectively. There were only six patients with the remaining possible combinations of three PHDs. Thus, the specific combination of types of PHDs did not affect the diagnostic utility of the presence of three PHDs. Among patients with two PHDs (n = 174), the most common combinations of PHDs were: 1) TSH deficiency and gonadotropin deficiency (n = 102); 2) TSH deficiency and ACTH deficiency (n = 38); and 3) ACTH deficiency and gonadotropin deficiency (n = 22). The PPV for prediction of a peak GH value less than 2.5 μg/liter for these three combinations of two PHDs were 81%, 89%, and 95%, respectively. The PPV for each of the remaining combinations of two PHDs (total of 12 patients) was less than 70%. Although the combination of ACTH deficiency and gonadotropin deficiency had a PPV greater than 95%, the reliability of this combination for prediction of GHD cannot be considered validated because of the small number of patients. Among patients with one PHD (n = 169), the PPV for prediction of a peak GH value less than 2.5 μg/liter was 70%, 69%, 68%, and 29% for gonadotropin deficiency (n = 74), ACTH deficiency (n = 16), TSH deficiency (n = 72), and AVP deficiency (n = 7), respectively.

Discussion

In this analysis of the clinical characteristics and biochemical testing results of 817 patients with a history of either

adult-onset hypothalamic or pituitary disease or childhood-onset GHD, we report that adult GHD could be predicted with 95% accuracy by the presence of either three or four PHDs or a serum IGF-I concentration less than 84 μg/liter (11 nmol/liter). These predictors of adult GHD also had a PPV of 95% or greater after excluding the data originally used to identify these criteria. Therefore, we consider these predictors to be prospectively validated. A strict definition of adult GHD (peak GH < 2.5 μg/liter) was used in light of the fact that 11 different GH stimulation tests were used in the clinical practice of the investigators participating in HypoCCS, a postmarketing safety surveillance study of GH replacement therapy in adults with GHD. This is important because the potency of the various provocative agents used for stimulation of GH secretion varies considerably (6). However, a peak GH cut-point of less than 5 μg/liter, irrespective of the provocative agent used, is in wide use in clinical practice in the United States because of the recommendations of the American Association of Clinical Endocrinologists and the approved labeling of some of the recombinant human GH products prescribed in the United States (2–4). Although the Growth Hormone Research Society has recommended a peak GH cut-point of less than 3 μg/liter specifically with an ITT (1), a peak GH less than 3 μg/liter is widely used in clinical practice as the cut-point for a variety of GH stimulation tests in both Europe and the United States. Thus, the diagnostic criteria evaluated in the present study will predict adult GHD more conservatively than is presently the case in clinical practice.

The finding that the presence of three or four PHDs had a high predictive value for adult GHD is in agreement with several previous studies. Toogood *et al.* (7) reported that in 190 adults with hypothalamic or pituitary disease, GHD was present in 24%, 55%, and 91% of patients with zero, one, or two to three PHDs, respectively. In that single-center study, GHD was defined by a peak GH value less than 5 mU/liter (~2.3 μg/liter) on an ITT in all patients and the results of patients with two or three PHDs were combined because the median peak GH values did not differ significantly between those two groups (7). Weissberger *et al.* (8) reported that among 103 adults with adult-onset pituitary disease, 51% had three or four PHDs. In these patients with three or four PHDs the peak GH response to an ITT was less than 1 μg/liter in 89% and less than 5 μg/liter in 100%. Similarly, Bates *et al.* (9) reported that patients with three PHDs had peak GH responses (ITT or glucagon test) less than 2 mU/liter (~0.9 μg/liter) and less than 10 mU/liter (~4.6 μg/liter) in 89% and 100% of the cases, respectively. The largest study before the present investigation was a French multicenter study conducted by Sassolas *et al.* (10). They reported that among 549 patients with hypothalamic or pituitary disease, GHD was present in 20%, 46%, 70%, and 93% of patients with zero, one, two, and three PHDs, respectively. In that study, GHD was defined by a peak GH cut-point of less than 3 μg/liter using a variety of GH stimulation tests. Partial GHD was defined as a peak GH between 3 and 5 μg/liter and was present in 11%, 13%, 11%, and 5% of patients with zero, one, two, and three PHDs, respectively. The most commonly used test was the ITT, which was performed in 75% of the patients. Other tests that were commonly used included GHRH, com-

bined glucagon-propranolol, ornithine and L-dopa. In the present study, GHD was defined by a stricter peak GH cut-point (< 2.5 μg/liter) than was used by Sassolas *et al.* (10), but fewer patients received ITTs (Table 1). The percentage of patients with three PHDs reported to be GH deficient in our study was comparable to that observed by Sassolas *et al.* (10) (96% *vs.* 93%, respectively). Most of the previous studies have not reported the prevalence of GHD in patients with panhypopituitarism and diabetes insipidus (7, 9, 10). In our study, 99% of these patients had GHD; the only patient with four PHDs who was not classified as GH deficient had a peak GH value of 4.5 μg/liter.

A careful review of the 10 patients with three to four PHDs who had a peak GH result 2.5 μg/liter or more reveals that the use of this prediction rule will yield only rare instances of inappropriate GH treatment (Table 2). Three of these patients had a peak GH value less than 5 μg/liter, which meets the criteria for adult GHD approved by the Food and Drug Administration and American Association of Clinical Endocrinologists (2–4). Four patients had borderline normal results with peak GH between 5 and 8 μg/liter. Although these patients do not meet criteria for treatment with GH, they may have partial GHD. Three patients were clearly normal with peak GH values 10 μg/liter or more. This contrasts with 305 patients who had three to four PHDs but had peak GH results less than 2.5 μg/liter. Thus, if GH stimulation tests had not been performed and GH replacement therapy had been offered to all 315 patients with three or four PHDs, only 2.2% would have received GH therapy inappropriately.

The use of three or four PHDs to diagnose adult GHD compares very favorably with GH stimulation testing. Although the ITT has been considered the "gold standard" test (1), it is not a perfect test. Hoeck *et al.* (16) reported that the reproducibility of the GH response to an ITT was poor in normal subjects, particularly in women. Of the eight normal women who received an ITT on two separate days, four had a peak GH less than 5 μg/liter on one occasion but not on the other (16). Although Hoffman *et al.* (5) reported a perfect separation of hypopituitary patients and normal subjects using a peak GH cut-point of less than 5 μg/liter with an ITT, this was not the case in the study of Fisker *et al.* (17). The GH response to L-arginine also does not consistently exceed 5 μg/liter in normal control subjects (17, 18). Furthermore, external factors such as physical activity and compliance with fasting may alter the GH response to provocative testing under less standardized conditions as commonly occurs in clinical practice (17). Thus, the diagnostic accuracy of the clinical criteria of three or four PHDs appears to be at least comparable with GH stimulation testing. We propose that GH stimulation testing for patients with three or four PHDs is unnecessary.

Although measurement of serum IGF-I has been very useful in the diagnosis of childhood GHD (19, 20), its diagnostic utility in adults has been questioned because of the large amount of overlap between normal subjects and adults with pituitary disease and GHD proven by GH stimulation testing. Hypopituitary adults with GHD have been reported to have normal serum IGF-I levels in 37–70% of patients in various studies (5, 9, 18, 21, 22). This is owing in part to the fact that multiple factors regulate serum IGF-I concentrations

including nutritional status; hepatic and renal function; and circulating concentrations of thyroid hormone, androgens, and estrogens (20, 23). In addition, changes in concentrations of IGF-binding proteins (IGFBPs) influence the total concentration of IGF-I in plasma. In patients with GHD, the decrease in serum levels of IGFBP-3 is partially compensated for by a rise in IGFBP-2 concentrations (24). Thus, the total serum IGF-I levels may not decrease below the normal range in many patients with GHD. Current methods for measurement of free IGF-I also do not reliably discriminate GHD patients from normal subjects (22). The predictive value of serum IGF-I may be improved by the use of locally determined normal ranges adjusted for age and sex (25). Therefore, the present consensus is that serum IGF-I alone should not be used to make the diagnosis of adult GHD (1, 2). Some authors have advocated using serum IGF-I as a screening test. Using this approach, patients with a serum IGF-I at least 1 SD below the age-adjusted mean are selected for GH stimulation testing (26). However, some GHD patients will be missed using this approach. Among patients with an IGF-I SD score above −1 in the present study, 46% had a peak GH less than 2.5 μg/liter and 67% had a peak GH less than 5 μg/liter.

In the present study, we sought to identify an IGF-I cut-point below which the probability of GHD would be 95% or greater, and thus potentially eliminate the need for GH stimulation testing. The cut-point of 84 μg/liter (11 nmol/liter) had a PPV of 96% for adult GHD, defined by a peak GH less than 2.5 μg/liter. This diagnostic prediction rule was identified in a preliminary analysis of 162 patients and was then validated prospectively in independent data sets (excluding the original 162 patients) as the number of patients in the HypoCCS database increased. It is important to note that this cut-point is valid for only the IGF-I assay employed in this study (Esoterix Endocrinology competitive binding RIA) because significant differences exist among results obtained with different IGF-I assays. Using a similar approach but a different IGF-I assay (RIA, Nichols Institute Diagnostics, San Juan Capistrano, CA), Baum *et al.* (18) reported that 20 of 23 middle-aged GHD adults had values below 144 μg/liter whereas all 17 normal controls were above this limit. Span *et al.* (27) reported that a serum IGF-I value of 15 nmol/liter (~115 μg/liter), measured by an in-house RIA, had a PPV for adult GHD of 93% in patients aged 40 yr or less. However, in patients over age 40 yr, the PPV was only 59%. In the present study, a serum IGF-I value less than 84 μg/liter was highly predictive of GHD in all age groups. The use of IGF-I SD scores may help to generalize our findings beyond the assay used in this study. Using regression equations (developed by Blum WF, personal communication), we determined that a serum IGF-I of 84 μg/liter in the assay employed corresponded to an IGF-I SD scores of −3 for patients over the age of 28 yr. For patients under the age of 28 yr, an even lower IGF-I SD scores was needed. An IGF-I SD scores of less than −2 (corresponding to the lower limit of the age-adjusted normal range) had a PPV of only 83%, which is insufficient to eliminate the need for GH stimulation testing. Similarly, Thissen *et al.* (28) reported that the sensitivity of an IGF-I SD scores of −2 with the Nichols RIA was 52% on extracted plasma and 68% on unextracted plasma. Thus, the serum IGF-I must be substantially below the lower limit of the

normal range to have diagnostic utility. Further comparative studies are needed at both a national and international level to determine the impact of assay differences on this diagnostic criterion of serum IGF-I less than 84 μg/liter. Until those studies are completed, measurement of serum IGF-I should not replace GH stimulation testing for the diagnosis of adult GHD unless the same IGF-I assay employed in this study is used or unless a useable conversion factor is available.

A careful review of the 11 patients with a serum IGF-I less than 84 μg/liter who had a peak GH result 2.5 μg/liter or greater in this study reveals that the use of this prediction rule (with this IGF-I assay) will yield only rare instances of inappropriate GH treatment (Table 3). Of these 11 patients, three had peak GH less than 5 μg/liter, five had peak GH between 5 and 7 μg/liter (partial GHD), and three were clearly not GH deficient because the peak GH exceeded 10 μg/liter. This contrasts with 300 patients who had serum IGF-I less than 84 μg/liter but had peak GH results less than 2.5 μg/liter. Thus, if GH stimulation tests had not been performed and GH replacement therapy had been offered to all 311 patients with serum IGF-I less than 84 μg/liter, only 2.6% would have received GH therapy who did not have a peak GH less than 5 μg/liter. The 11 patients with serum IGF-I less than 84 μg/liter and peak GH result 2.5 μg/liter or greater likely had less severe pituitary disease than the other patients with serum IGF-I less than 84 μg/liter in this study. Although the number of PHD was not known for three of the patients, none of the remaining eight patients had three or more PHDs. Furthermore, four of the nine patients for whom clinical diagnoses were provided had idiopathic hypothalamic-pituitary disease.

Other causes of low serum IGF-I besides GHD should be kept in mind. Two of the 11 patients had BMIs of 18.2 and 16.1 kg/m^2 and peak GH concentrations of 29 and 49 μg/liter, respectively. These patients may have been malnourished. Malnutrition and prolonged fasting are associated with low serum IGF-I concentrations and increased GH secretion (23, 29, 30). In addition to malnutrition, poorly controlled type I diabetes mellitus, hypothyroidism, and hepatic insufficiency may all be associated with low IGF-I concentrations (20, 23). Thus, these disorders should be excluded before using a low serum IGF-I as a marker of GHD (1). Ten of the 11 patients were women; 8 of these 10 women were taking oral estrogen preparations. Oral estrogen administration is known to decrease serum IGF-I concentrations and increase 24-h GH release. This effect does not occur with conventional doses of transdermal estrogen but is observed with high-dose transdermal estrogen (31, 32). Thus, caution should be used with the use of serum IGF-I for the diagnosis of adult GHD in women receiving oral estrogen, particularly if the patient has less than three PHDs. The presence of three to four PHDs appears to be a more reliable predictor of adult GHD for women receiving oral estrogen because only two such women had peak GH greater than 2.5 μg/liter (actual peaks of 2.9 and 6.7 μg/liter).

The present study is the first to evaluate various combinations of PHD and low serum IGF-I concentrations for the prediction of adult GHD. In this study, the presence of either three or four PHDs or a serum IGF-I less than 84 μg/liter had

a PPV of 95% for adult GHD with a specificity of 89% and a sensitivity of 69%. This combination rule offers the best combination of high specificity and moderate sensitivity, potentially allowing for more than half of patients meeting clinical criteria to be diagnosed without a GH stimulation test. Although a PPV of 95% was the criterion we chose for an acceptable prediction rule, some may wish to use only a prediction paradigm that provided 100% PPV and 100% specificity. This can be done by requiring the combination of three or four PHDs and a serum IGF-I less than 84 μg/liter. However, this approach yields a very low sensitivity (30%) and hence more patients will continue to need GH stimulation testing. Previous studies have also not evaluated the impact of specific types of PHDs on the predictive value for adult GHD. In this study, the specific combinations of PHDs did not change the predictive value of three PHDs. Although there were some differences in the PPV for specific combinations of two PHDs, none of these combinations can be considered valid predictors of adult GHD because of the relatively small numbers of patients with these specific combinations.

In summary, adult GHD can be predicted with 95% accuracy by the presence of either three or four PHDs or a serum IGF-I concentration less than 84 μg/liter (11 nmol/liter) in the Esoterix Endocrinology assay. We propose that adult patients with three or four PHDs (three or four of the following deficiencies: TSH, ACTH, gonadotropins [LH and/or FSH], and AVP [central diabetes insipidus]) do not require a GH stimulation test to make the diagnosis of adult GHD. These clinical predictors are at least as accurate as GH stimulation tests performed in routine clinical practice. The diagnostic utility of a serum IGF-I less than 84 μg/liter to predict adult GHD is limited to the particular IGF-I assay employed in this study. Until further comparative studies of other commercially available IGF-I assays are performed, patients whose IGF-I levels are measured with other assays should have a GH stimulation test to confirm the diagnosis of adult GHD. Other causes of low serum IGF-I should be excluded before using IGF-I as a marker of GHD.

Stonesifer, Federal Way; West Virginia: Stephen R. Grubb; Wisconsin: Thad Hagen, Milwaukee; Yoram Shenker, Madison.

References

1. **Growth Hormone Research Society** 1998 Consensus guidelines for the diagnosis and treatment of adults with growth hormone deficiency: summary statement of the Growth Hormone Research Society workshop on adult growth hormone deficiency. J Clin Endocrinol Metab 83:379–381
2. **American Association of Clinical Endocrinologists** 1998 AACE clinical practice guidelines for growth hormone use in adults and children. Endocr Pract 4:165–173
3. **Physician's desk reference** 2001a 5th ed. Montvale, NJ: Medical Economics; Humatrope; p. 1733–1736
4. **Physician's desk reference** 2001b 5th ed. Montvale, NJ: Medical Economics; Nutropin; p. 1304–1307
5. **Hoffman DM, O'Sullivan AJ, Baxter RC, Ho KY** 1994 Diagnosis of growth hormone deficiency in adults. Lancet 343:1064–1068
6. **Shalet SM, Toogood A, Rahim A, Brennan BMD** 1998 The diagnosis of growth hormone deficiency in children and adults. Endocr Rev 19:203–223
7. **Toogood AA, Beardwell CG, Shalet SM** 1994 The severity of growth hormone deficiency in adults with pituitary disease is related to the degree of hypopituitarism. Clin Endocrinol (Oxf) 41:511–516
8. **Weissberger AJ, Verikiou K, Sönksen PH** 1994 Diagnosis of growth hormone deficiency in adults. Lancet 344:483
9. **Bates AS, Evans AJ, Jones P, Clayton RN** 1995 Assessment of GH status in adults with GH deficiency using serum growth hormone, serum insulin-like growth factor-I and urinary growth hormone excretion. Clin Endocrinol (Oxf) 42:425–430
10. **Sassolas G, Borson Chazot F, Jaquet P, et al.** 1999 GH deficiency in adults: an epidemiological approach. Eur J Endocrinol 141:595–600
11. **Stene MA, Moxness MS, Sinha YN** 1998 High sensitivity immunoassay for growth hormone: adult normal ranges and clinical utility for predicting acromegaly with single sample basal level testing. Proc 80th Meeting of the Endocrine Society, New Orleans, LA, 1998, pp. 290–291 (Abstract P2-181)
12. **Blum WF, Brier BH** 1994 Radioimmunoassays for IGFs and IGFBPs. Growth Regul 4 [Suppl 1]:11–19
13. **Rahim A, Toogood AA, Shalet SM** 1996 The assessment of growth hormone status in normal young adult males using a variety of provocative agents. Clin Endocrinol (Oxf) 45:557–562
14. **Toogood AA, Nass PM, Pezzoli SS, O'Neill PA, Thorner MO, Shalet SM** 1997 Preservation of growth hormone pulsatility despite pituitary pathology, surgery, and irradiation. J Clin Endocrinol Metab 82:2215–2221
15. **Mausner JS, Bahn AK** 1974 Epidemiology. Philadelphia: WB Saunders; 237–263
16. **Hoeck HC, Vestergaard P, Jakobsen PE, Laurberg P** 1995 Test of growth hormone secretion in adults: poor reproducibility of the insulin tolerance test. Eur J Endocrinol 133:305–312
17. **Fisker S, Jørgensen JOL, Ørskov H, Christiansen JS** 1998 L-Arginine and insulin-tolerance tests in the diagnosis of adult growth hormone deficiency: influence of confounding factors. Clin Endocrinol (Oxf) 48:109–115
18. **Baum HB, Biller BM, Katznelson L, Oppenheim DS, Clemmons DR, Cannistraro KB, Schoenfeld DA, Best SA, Klibanski A** 1996 Assessment of growth hormone (GH) secretion in men with adult-onset GH deficiency compared with that in normal men—a clinical research center study. J Clin Endocrinol Metab 81:84–92
19. **Lee PDK, Wilson DM, Rountree L, Hintz RL, Rosenfeld RG** 1990 Efficacy of insulin-like growth factor I levels in predicting the response to provocative growth hormone testing. Pediatr Res 27:45–51
20. **Blum WF** 2000 Insulin-like growth factors and IGF-binding proteins: their use for diagnosis of growth hormone deficiency. In: Juul A, Jørgensen JOL, eds. Growth hormone in adults. 2nd ed. Cambridge: Cambridge University Press; 54–86
21. **Svensson J, Johannsson G, Bengtsson BÅ** 1997 Insulin-like growth factor-I in growth hormone-deficient adults: relationship to population-based normal values, body composition and insulin tolerance test. Clin Endocrinol (Oxf) 46:579–586
22. **Musolino NRC, Da Cunha Neto MB, Marino R, Giannella-Neto D, Bronstein MD** 1999 Evaluation of free insulin-like growth factor-I measurement on the diagnosis and follow-up treatment of growth hormone-deficient adult patients. Clin Endocrinol (Oxf) 50:441–449
23. **Clemmons DR, Underwood LE** 1991 Nutritional regulation of IGF-I and IGF binding proteins. Ann Rev Nutr 11:393–412
24. **Smith WJ, Nam TJ, Underwood LE, Busby WH, Celnicker A, Clemmons DR** 1993 Use of insulin-like growth factor-binding protein-2 (IGFBP-2), IGFBP-3, and IGF-I for assessing growth hormone status in short children. J Clin Endocrinol Metab 77:1294–1299
25. **Granada ML, Murillo J, Lucas A, et al.** 2000 Diagnostic efficiency of serum IGF-I, IGF-binding protein-3 (IGFBP-3), IGF-I/IGFBP-3 molar ratio and urinary GH measurements in the diagnosis of adult GH deficiency: importance of an appropriate reference population. Eur J Endocrinol 142:243–253
26. **Klibanski A, Clemmons DR, Christiansen J, Underwood L** 1999 Growth hormone replacement therapy for growth hormone deficiency in adults: the changing role of growth hormone therapy. Endocr Pract 5:88–96
27. **Span JPT, Pieters GFFM, Sweep CGJ, Swinkels LMJW, Smals AGH** 1999 Plasma IGF-I is a useful marker of growth hormone deficiency in adults. J Endocrinol Invest 22:446–450
28. **Thissen JP, Ketelslegers JM, Maiter D** 1996 Use of insulin-like growth factor-I (IGF-I) and IGF-binding protein-3 in the diagnosis of acromegaly and growth hormone deficiency in adults. Growth Regul 6:222–229
29. **Ho KY, Veldhuis JD, Johnson ML, Furlanetto R, Evans WS, Alberti KG, Thorner MO** 1988 Fasting enhances growth hormone secretion and amplifies the complex rhythms of growth hormone secretion in man. J Clin Invest 81:968–975
30. **Hartman ML, Veldhuis JD, Johnson ML, Lee MM, Alberti KG, Samojlik E, Thorner MO** 1992 Augmented growth hormone (GH) secretory burst frequency and amplitude mediate enhanced GH secretion during a two day fast in normal men. J Clin Endocrinol Metab 74:757–765
31. **Weissberger AJ, Ho KKY, Lazarus L** 1991 Contrasting effects of oral and transdermal routes of estrogen replacement therapy on 24-hour growth hormone (GH) secretion, insulin-like growth factor I, and GH-binding protein in postmenopausal women. J Clin Endocrinol Metab 72:374–381
32. **Friend KE, Hartman ML, Pezzoli SS, Clasey JL, Thorner MO** 1996 Both oral and transdermal estrogen increase growth hormone release in postmenopausal women—a clinical research center study. J Clin Endocrinol Metab 81:2250–2256