0031-3998/90/2701-0045$02.00/0
PEDIATRIC RESEARCH

Vol. 27, No. 1, 1990
Printed in U.S.A.

Efficacy of Insulin-Like Growth Factor I Levels in Predicting the Response to Provocative Growth Hormone Testing

PHILLIP D. K. LEE, DARRELL M. WILSON, LOIS ROUNTREE, RAYMOND L. HINTZ, AND RON G. ROSENFELD

The Children's Hospital, Denver, Colorado 80218; and Department of Pediatrics, Stanford University, Stanford, California 94305

ABSTRACT. Clinical testing of growth hormone (GH) sufficiency is a controversial area in endocrinology. Due to the episodic nature of endogenous GH secretion, diagnosis of GH deficiency has been defined as a failure to achieve normal GH levels in response to at least two stimuli. This testing is associated with significant patient morbidity and cost. We analyzed our experience over a 4-y period to determine whether clinical or biochemical variables could be used to predict the results of a specific GH testing procedure. Of 180 cases analyzed (67% male, mean age 8.89 ± 4.39 y, range neonate-16 y), eight cases had incomplete GH testing results. Of the remaining 172, 19 were GH deficient (GH level <7 ng/mL). Younger age, higher body mass index and a greater degree of bone age delay were characteristic of the GH-deficient population; however, none of these variables alone was of diagnostic utility. Serum IGF-I level was below the normal range for 81% of the GH deficient and 47% of the GH-sufficient children; and was the only single variable that provided a reasonable between-group distinction. Discriminant analysis resulted in development of a new variable, based on IGF-I z scores, chronologic age, degree of bone age delay, and body mass index, which would have allowed exclusion of GH deficiency without provocative testing for 58% of the GH sufficient population, whereas permitting the diagnosis of GH deficiency for all GH-deficient subjects. Our data are dependent on the IGF-I assay method and the clinical definition for GH deficiency; therefore, the calculated predictive values are not applicable to all clinical populations. However, our data provide a new perspective on the integration of IGF-I levels and clinical information in predicting GH sufficiency. (*Pediatr Res* 27:45–51, 1990)

Abbreviations

BA, bone age
BMI, body mass index
BW, body weight
CA, chronologic age
GH, growth hormone

Growth retardation is one of the most common problems encountered in pediatric endocrinology. Initial evaluation usually includes exclusion of hypothyroidism and non-endocrine organic causes, and consideration of familial and psychosocial factors (1, 2). Determination of GH sufficiency is frequently included as part of this evaluation.

Because GH secretion is episodic, determination of GH sufficiency generally involves administration of a stimulus to enhance pituitary GH secretion, followed by multiple venous blood samplings for determination of GH levels. Commonly used stimuli include exercise, sleep, and pharmacologic agents such as clonidine, L-DOPA, arginine, ornithine, and insulin (3, 4). By convention, failure to attain a normal circulating GH level after two provocative stimuli defines classical GH deficiency (5). Provocative GH testing involves morbidity and discomfort for children due to the side-effects of the pharmacologic agents and the prolonged period of blood sampling. Furthermore, patient cost and time commitment for health personnel are significant.

In view of these considerations, more efficient screening procedures for GH sufficiency have been sought. Single GH levels obtained after exercise have given variable results (6). Recently, urinary GH levels have shown promise, although the assay itself is somewhat cumbersome and not readily available (7, 8).

IGF-I, also known as somatomedin-C, is a 7.5-kD protein that mediates the cellular growth-promoting actions of GH (9). Blood IGF-I levels are dependent on GH adequacy, with low levels in hypopituitarism and elevated levels in acromegaly (10). Unlike GH, blood IGF-I concentrations are fairly constant through the day, due to the presence of specific binding proteins (11). Average IGF-I concentrations are dependent on age, sex, and nutritional status (12).

IGF-I levels have been advocated as a screening procedure for GH deficiency (13, 14), although there are limited data to support this use in a clinical setting. We reviewed our experience using IGF-I levels and other clinical variables in the evaluation of GH sufficiency. Our data indicate that IGF-I levels may be useful in predicting response to provocative GH testing, particularly when considered together with other clinical variables.

MATERIALS AND METHODS

Study population. Data were collected over a 4-y period from 177 consecutive children evaluated for possible GH deficiency in the Pediatric Endocrinology clinic at the Children's Hospital at Stanford. Three children were evaluated on two separate occasions due to changing clinical status (*e.g.* cranial irradiation subsequent to initial GH evaluation), giving a total of 180 data sets. Criteria for evaluation of GH status varied by individual case, but generally included the following: ht <5th centile, abnormal growth rate, delayed bone age, predicted adult ht <145 cm for girls and <160 cm for boys, and exclusion of other causes for growth retardation, including hypothyroidism. This popula-

Received April 18, 1989; accepted September 12, 1989.
Correspondence and reprints Phillip D.K. Lee, M.D., Diabetes Research Center, Baylor College of Medicine, 8080 North Stadium Drive, Houston, TX 77054.
Supported by a Basic Research Support grant from The University of Colorado Health Sciences Center (P.D.K.L.) and a grant from The Children's Hospital Kempe Research Center, Denver, CO (P.D.K.L.).

tion represents ~50% of the total population referred for evaluation of growth retardation during this period.

Testing protocol. Our routine protocol for evaluation of GH adequacy was as follows: children were admitted to the outpatient clinic after an overnight 6- to 8-h fast for the initial screening procedure. A heparin-flushed intravenous catheter was placed for blood sampling. L-DOPA or clonidine HCl was then orally administered, and blood samples were obtained for GH levels predose, and at 60 and 90 min postdose. The L-DOPA dosage was 125 mg for up to 30 lb BW, 250 mg for >30 up to 60 lb BW, and 500 mg for >60 lb BW. The clonidine HCl dosage was 5 μg/kg BW rounded to the nearest 50 μg, maximum dose = 250 μg. In eight cases, an exercise test was performed, involving 15–20 min of unmonitored exertion on an exercise bicycle, with GH levels measured at baseline, immediately postexercise, and 20 and 40 min postexercise.

A normal response was defined as a peak GH level of ≥7 μg/L. Children who failed to achieve this level during the screening procedure were then admitted to the hospital for an arginine-insulin infusion test. After an overnight fast, 0.5 g/kg of arginine HCl (max 30 g) was infused intravenously over 30 min. Premarin (conjugated estrogens) 2.5 mg was administered orally with water at bedtime in the evening before the test and immediately before the test. GH levels were measured at baseline and 15, 30, 45, and 60 min after the start of the arginine infusion, and a cortisol level was obtained at baseline. Regular insulin, 0.1 U/kg, was then given intravenously, and GH levels were obtained at 15, 30, 45, 60, 90, 150, and 210 min postinsulin infusion. Cortisol levels were measured at 30, 60, and 90 min after insulin infusion. An adequate hypoglycemic response was a serum glucose decline of ≥50% from baseline or an absolute nadir of <2.8 mM. A peak GH response of ≤7 μg/L after these procedures defined GH deficiency. A normal cortisol response was considered to be a rise of 10 μg/mL over baseline or a level of ≥552 nM (20 μg/dL) for any measurement. There were 19 children who had the arginine-insulin infusion test without a prior screening procedure. These included children with conditions which strongly predispose to GH deficiency, including craniopharyngioma and cranial irradiation. The test procedure was modified by elimination of the insulin-infusion test for seven children who were believed to be at risk for severe hypoglycemia. In these cases, failure to achieve a peak GH level in response to at least two other stimuli defined GH deficiency.

Assay methods. Serum GH and IGF-I levels were measured by radioimmunoassay at Endocrine Sciences Laboratories, Tarzana, CA. For the IGF-I assay, samples were prepared by acidification and ethanol precipitation (15) and a final second-antibody precipitation step is used to minimize interference by endogenous IGF binding proteins. Cortisol levels were determined by RIA at Stanford University Hospital.

Data collection and analysis. The following data were collected at the time of the screening procedure: age, ht, wt, ancillary diagnoses, and baseline and peak GH levels. Heights (mean of three consecutive determinations) were measured on a calibrated Harpendon stadiometer by a single observer (L.R.). IGF-I levels within 6 mo of the screening procedure were included. Cortisol levels were tabulated for children tested with the insulin infusion test. Growth rate data were included only if >3 mo of ht measurements in our clinic were available. Since this information was available for <50% of the patients, growth rate was not included in the final data analyses. Body mass index was calculated as wt/ht^2.

Data were analyzed using SAS-PC (16). Descriptive data are expressed as the mean ± SD. Between-group comparisons were analyzed using the unpaired t test assuming equal variances, unless the variances were demonstrated to be unequal by the folded form of the F-statistic, in which case the t test was done using the Satterwaite approximation. Correlations between separate variables were calculated using the Pearson method. p values of <0.05 were considered significant. Predictive values were calculated using discriminant analysis (16).

Z scores for IGF-I levels were determined from the age- and sex-related normal values provided by Endocrine Sciences Laboratory, Tarzana, CA. Z scores for height were determined using the National Health Examination Survey data (17). BA was determined from hand and wrist radiographs and BA z scores were calculated using the expected BA and SD (18).

RESULTS

Study population. CA at evaluation ranged from newborn to 16 y, with a mean of 8.89 ± 4.39 y for the 180 data sets. There were 118 males (67%) and 59 females (33%). The three children who were tested on two separate occasions included two boys and one girl. All patients were euthyroid at the time of GH testing.

Diagnoses assigned prior to GH testing included Turner syndrome (n = 8), other chromosomal or dysmorphic syndromes (n = 17), failure to thrive (n = 7), idiopathic or iatrogenic hypopituitarism (n = 26), constitutional delay of growth and development (n = 20), normal variant short stature (n = 82), diabetes mellitus (n = 2), or other conditions, such as steroid-dependent illnesses (n = 14). Four children were thought to be probably normal, but were tested for GH sufficiency after consideration of growth rate and genetic potential. A total of 166 of 179 (92%, 1 missing data) cases had a height z score less than −1.65 (~5th centile) at the time of evaluation.

GH testing. A total of 160 children had an initial outpatient GH screening procedure performed, for a total of 161 total tests; 108 passed this initial test and required no further GH testing. Of the 53 patients who failed (33%) the initial screen, 45 (85%) had a second test performed and eight were lost to follow-up. Fourteen of the 45 (31%) children receiving a second test failed both the GH screening and second tests.

Of the 161 screening tests, 114 used clonidine-stimulation, 39 used L-DOPA stimulation, and eight were exercise tests. Eighty-three of 114 (73%) children tested with clonidine had a GH level ≥7 μg/L, whereas 20 of 39 (52%) tested with L-DOPA attained this level ($p < 0.02$ by χ^2 for clonidine *versus* L-DOPA). Ten of 31 (32%) of the clonidine test failures, and four of 19 (21%) of the L-DOPA test failures failed a second test and were diagnosed as GH deficient (NS by χ^2 for clonidine *versus* L-DOPA). The clonidine test appears to be a more specific screening test for GH deficiency than the L-DOPA test. The number of children tested with exercise-stimulation was too small to allow accurate outcome comparisons.

For the 101 children with a GH level ≥7 μg/L after L-DOPA or clonidine stimulation, nine (9%) had a peak GH level at baseline, 56 (55%) had a peak GH level at 60 min, and 35 (35%) had a peak at 90 min. One had a peak at 120 min after clonidine (not a routine time of sampling), and the time of sampling was not recorded for two children. Eighteen of 114 (16%) clonidine and four of 33 L-DOPA tests (10%, six with unrecorded baseline GH level) had a GH ≥ 7 μg/L at baseline, and theoretically would not have needed the stimulation procedure. Of the five children who passed an exercise test, one had a peak at baseline, three immediately after exercise, and one at 40 min after exercise.

An additional 19 children did not have an initial outpatient GH screening procedure and five (26%) of these were diagnosed as GH deficient on the basis of an insulin-arginine test. Therefore, a total of 19 children (11 males, eight females), representing 11% of the complete GH testing data sets, was diagnosed with GH deficiency. Figure 1 depicts the process of GH evaluation for the 180 cases, leading to a final diagnosis of GH deficiency for 19 patients.

GH testing results in relation to pre- and posttesting diagnoses were examined. Only one patient thought to have "normal variant short stature" before GH testing failed to achieve a GH level ≥7 μg/L on multiple tests, whereas 8 of 18 "hypopituitary" children passed the GH testing. However, 18 of the 19 (95%) of the GH-deficient children were thought to have hypopituitarism before GH testing. The diagnostic yield in the other groups was

nil. This apparently high predictive value of clinical acumen is probably related to other factors, such as referral bias (including testing previously done elsewhere), and predisposing conditions (irradiation, surgery).

Of the 18 children with a pre-GH testing diagnosis of hypopituitarism, five had a final diagnosis of idiopathic isolated GH deficiency (including one with a history of premature birth), two had GH and thyrotropin deficiency, two had hereditary GH deficiency, six had intracranial tumors (two craniopharyngiomas, one nonfunctional adenoma, one Cushing's disease, two other), two had septooptic dysplasia, and one had histiocytosis X. The child with a pre-GH test diagnosis of normal variant short stature had a final diagnosis of idiopathic isolated GH deficiency. Fifteen of the 19 GH-deficient children were subsequently treated with exogenous GH in our clinic, and 14 of 15 had a marked increase in growth rate after 1 y of therapy.

Twelve of 14 patients with ht z scores above −1.65 SD (5th centile) were found to have normal GH levels. The two who were GH deficient were just at −1.65 SD for ht, and both had received cranial irradiation. Interestingly, all eight children who failed the initial outpatient screening test and were subsequently lost to follow-up were male.

IGF-I levels. IGF-I levels were below the age- and sex-related absolute range for 13 of 16 (81%) GH-deficient children, and 62 of 133 (47%) non-GH-deficient children ($p < 0.005$ by χ^2). Three of 19 GH deficient (16%) and 20 of 153 GH-sufficient children (13%) did not have IGF-I levels available within 6 mo of the GH tests. Eight children had incomplete GH data sets (failed the outpatient screening test, then lost to follow-up); and of this group, IGF-I levels were low for three, normal for four, and missing for one.

If a 2 SD (rather than absolute) range is used, the percentages falling below the 2 SD curve are 62.5% (10/16) for GH deficient and 27% (36/133) for non-GH deficient ($p < 0.005$ by χ^2). The percentages using a 1 SD range are 94% (15/16) and 84% (112/133), respectively ($p < 0.005$ by χ^2).

IGF-I z scores were −2.06 ± 0.77 SD ($n = 16$) for the GH-deficient group, −1.69 ± 0.62 SD for the group with incomplete results ($n = 7$), and −1.57 ± 0.61 SD for the GH-sufficient group (Fig. 2). The GH-deficient and sufficient groups differed significantly ($p < 0.01$).

Other variables. Table 1 compares several other variables with regard to distinguishing GH-sufficient and GH-deficient children, in the group as a whole and after separation by BA ≥ 7 y. As mentioned above, in the overall population, only IGF-I z scores were significantly different between the GH-deficient and GH-sufficient groups. When separated according to BA ≥ 7 y, the discriminant value of IGF-I was preserved only for the younger group. Overall, there was no single variable that allowed clinically adequate discrimination of GH-sufficient and GH-deficient groups.

Although pre-GH test diagnosis was evidently highly predictive of GH sufficiency in our study population, this variable is largely subjective, and was not included in our predictive value analysis.

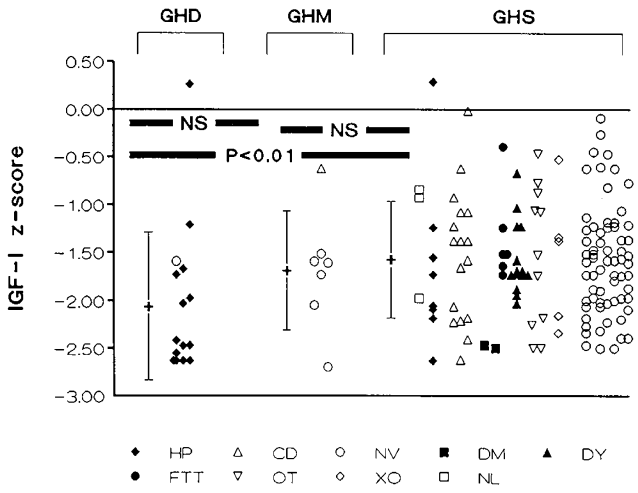

Fig. 2. IGF-I z scores *versus* final diagnosis. IGF-I z scores, means and SD are depicted according to the final diagnosis, which is indicated at the top. Abbreviations: GHD, GH deficient; GHS, GH-sufficient; GHM, missing complete GH testing data; NS, nonsignificant ($p > 0.05$) by repeated t test. The patients are identified according to pre-GH test clinical diagnosis by symbols as indicated at the bottom of the figure. HP, hypopitutitarism; CD, constitutional delay of growth and development; NV, normal variant short stature; DM, diabetes mellitus; DY, chromosomal or dysmorphic syndrome; FTT, failure to thrive; OT, other conditions, including steroid-dependent illnesses; XO, Turner syndrome; NL, normal stature.

Table 1. *Comparisons by BA**

Group	Variable	GH sufficient	GH deficient	p
All patients	n	153	19	
	CA (y)	9.08 ± 4.41 (153)	7.56 ± 3.90 (19)	NS
	ht z score	−2.81 ± 1.10 (152)	−3.06 ± 1.24 (19)	NS
	BA z score	−2.12 ± 1.37 (143)	−2.57 ± 1.78 (18)	NS
	IGF-I z score	−1.57 ± 0.61 (133)	−2.06 ± 0.77 (16)	0.004
	BMI	16.37 ± 2.68 (145)	17.79 ± 3.94 (19)	NS
BA < 7 y	n	60	12	
	CA (y)	5.46 ± 2.85 (60)	5.25 ± 2.00 (12)	NS
	ht z score	−2.78 ± 1.00 (60)	−3.58 ± 1.17 (12)	NS
	BA z score	−2.26 ± 1.62 (60)	−3.17 ± 1.69 (12)	NS
	IGF-I z score	−1.59 ± 0.53 (52)	−2.21 ± 0.52 (10)	0.001
	BMI	15.22 ± 1.78 (52)	15.85 ± 1.25 (12)	NS
BA ≥ 7 y	n	83	7	
	CA (y)	12.45 ± 2.04 (83)	11.51 ± 3.08 (7)	NS
	ht z score	−2.69 ± 0.84 (83)	−2.19 ± 0.80 (7)	NS
	BA z score	−2.03 ± 1.17 (83)	−1.39 ± 1.38 (6)	NS
	IGF-I z score	−1.59 ± 0.68 (73)	−1.81 ± 1.08 (6)	NS
	BMI	17.37 ± 2.77 (83)	21.12 ± 4.80 (7)	NS

* GH-sufficient and GH-deficient groups were separated and analyzed for the six variables indicated, for the group as a whole (omitting the eight incomplete data sets) and according to BA < 7 y or ≥7 y. Only IGF-I z score differed between the diagnostic groups, and this was preserved only for the younger BA. A $p < 0.05$ was considered significant. An assumption of unequal variances was used for the analyses involving BMI, as described in "Materials and Methods." Numbers in parentheses indicate the number of cases for which data were available. NS, not significant.

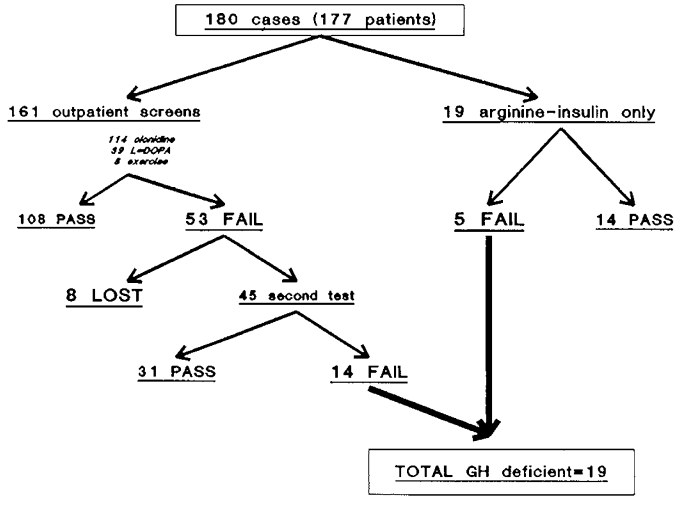

Fig. 1. Flow diagram of clinical GH evaluations. GH testing process for the 180 cases evaluated; leading to a final diagnosis of GH deficiency for 19 cases. See text for explanation.

Predictive values. Predictive values were calculated by discriminant analysis (16) for the 180 data sets (19 GH deficient, 153 GH sufficient, eight missing). Forty sets were found to be missing data for one or more variables. Ht z score, BMI, IGF-I z score, CA, and BA z score were initially entered, and the STEPDISC procedure was used with forward selection (16). BMI, IGF-I z score, CA and BA z score were found to independently vary with final diagnosis, whereas ht z score was deleted. The CANDISC procedure was then used to derive canonical variables. This resulted in a new variable, CAN1, incorporating the relative predictive values of four independent variables:

$$\text{CAN1} = 0.24[\text{CA}]-0.33[\text{BMI}]+0.98[\text{IGF-I z score}] +0.26[\text{BA z score}]+5.4$$

Figure 3 shows the distribution of the CAN1 scores for the GH-sufficient and GH-deficient groups. The mean CAN1 score was 0.16 ± 1.03 for the GH-sufficient group and −1.21 ± 0.73 for the GH-deficient group. With CAN1 = 0.034 chosen as a boundary providing 100% sensitivity, 58% (68/117) of the GH-sufficient children for whom a score can be calculated had a CAN1 above this value and would not have required provocative GH testing to exclude the diagnosis of GH deficiency. All 15 GH-deficient children for whom a score can be calculated had a CAN1 value ≤0.03385. A score could not be calculated for two of eight children missing complete diagnostic data sets, four of 19 (21%) of the GH-deficient group, and 36 of 153 (21.4%) of the GH-sufficient group. The CAN1 predictive value is presented only as an indicator of the relative values of individual variables in this particular study, and is not intended to be used as a universal diagnostic measure. Therefore, error limits for CAN1, which would depend on the boundary selected, are not presented.

Excluding the 23 cases that would have passed the GH screening procedure without stimulation (*i.e.* baseline GH ≥ 7 ng/mL), 113 cases went on to a final diagnosis of GH sufficient, and 92 of these had data that allowed calculation of CAN1. Of these, 56 (61%) had CAN 1 above 0.034 and 36 (39%) were below 0.034.

It should be noted that of the four independent variables, only IGF-I z score ($p = 0.004$) differed significantly between groups; whereas BMI ($p = 0.14$), CA ($p = 0.15$), and BA z score ($p = 0.20$) did not. However, the constructed variable, CAN1, suggests that the diagnosis of GH deficiency is more likely for younger children with a higher BMI, low IGF-I for age, and a greater degree BA delay.

Sex differences. The 2:1 male to female ratio in the study population raises the possibility that there may be a bias toward evaluating less severely growth-retarded boys. We compared the male and female groups overall and by final diagnosis (GH sufficient or deficient). Variables chosen were those found to be predictive for GH deficiency (see previous section), including CA, ht z score, IGF-I z score, BA z score, and BMI. As shown in Table 2, the girls as a group tended to be evaluated at a greater degree of ht deficit. After assignment into diagnostic groups, there were no differences between boys and girls for ht z score, but the BA z score was significantly lower for girls with GH deficiency.

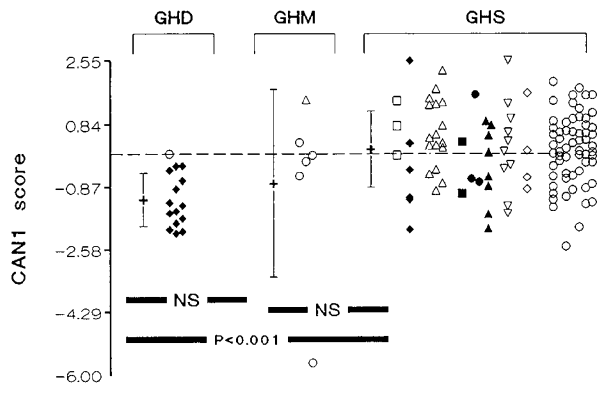

Fig. 3. Calculated predictive variable *versus* final diagnosis. Calculated values for the predictive variable, CAN1, means, and SD are depicted according the final diagnosis. The dotted line is drawn through CAN1 = 0.034. See text for further explanation. Abbreviations and figure symbols see Figure 2.

The eight children missing complete diagnostic data sets were all boys. Analysis of variance for the five variables examined showed no significant differences between this group and the GH-sufficient or GH-deficient male groups (data not shown).

Other correlates. We examined other factors that might influence either the GH test response or IGF-I levels in the GH sufficient (peak GH ≥ 7 ng/mL) population. For these correlations, the incomplete and GH deficient data sets were deleted. As expected, CA showed a strong positive correlation with IGF-I levels ($r = 0.66$, $p = 0.0001$, $n = 133$) within the CA range studied.

As shown in Table 3, the screening test baseline GH level varied negatively with age ($r = -0.30$, $p = 0.0003$, $n = 136$), BMI ($r = -0.26$, $p = 0.0026$, $n = 131$), and ht z score ($r = -0.25$, $p = 0.004$, $n = 136$). Peak GH levels on either the screening procedure or the second procedure did not correlate with age, BMI, or ht z score. Peak GH levels on the second procedure correlated negatively with BA z score ($r = -0.45$, $p = 0.003$, $n = 42$), whereas peak GH level for the screening procedure did not ($r = -0.11$, $p = 0.22$, $n = 132$). BA z score also showed a negative correlation with CA ($r = -0.20$, $p = 0.002$, $n = 143$).

IGF-I z scores correlated weakly with the peak GH level on screening test ($r = 0.21$, $p = 0.02$, $n = 123$) and negatively with BA z score ($r = -0.18$, $p = 0.05$, $n = 125$); but not with the baseline GH levels ($r = 0.14$, $p = 0.13$, $n = 121$), second-test GH peak ($r = -0.06$, $p = 0.74$, $n = 36$), or BMI ($r = 0.13$, $p = 0.15$, $n = 125$).

Cortisol levels were examined for those patients who underwent an insulin-infusion test. Neither the cortisol nor the GH peak levels correlated with the level of the glucose nadir. However, maximal cortisol levels did show a negative correlation with chronologic age ($r = -0.48$, $p = 0.004$, $n = 35$) and positive correlations with peak GH level during the second GH-test procedure ($r = 0.37$, $p = 0.03$, $n = 35$). There were no relationships between the cortisol peak and IGF-I z score, BA z score, or BMI.

DISCUSSION

The diagnosis of GH deficiency is a controversial area in pediatric endocrinology. By convention, failure to achieve a "normal" GH level in response to two known stimuli of GH secretion defines classical GH deficiency (5). The normal/abnormal GH response boundary is somewhat arbitrary, generally ranging between 5 and 10 μg/L; and is complicated by variability in GH levels as measured by different GH assay techniques (19, 20), possible false-low responses in prepuberty (21, 22) [although discrepant results have been reported (23)], and apparently normal GH responses in a large percentage of individuals previously diagnosed as GH deficient (24). In addition, correlations between pharmacologically provoked GH levels and measures of physiologic secretion (*e.g.* continuous sampling) have been variable (25–28). Despite these problems, however, provocative testing of GH secretion is generally regarded as the standard for diagnosis of "classical" GH deficiency.

Using this criterion, with peak GH level <7 μg/L as the diagnostic boundary, we analyzed our clinical experience over a 4-y period to determine whether particular clinical and biochemical indicators may predict those children who are most likely to fail provocative GH testing. Nineteen of 180 children screened over this 4-y period were thus diagnosed as GH deficient, 13 on the basis of three provocative tests (screening test + arginine + insulin) and six on the basis of two tests (arginine + insulin). We

Table 2. *Comparisons by sex**

Group	Variable	Male	Female	*p*
All patients	*n*	120	60	
	CA (y)	9.26 ± 4.51 (120)	8.25 ± 4.07 (60)	NS
	ht z score	−2.73 ± 0.99 (112)	−3.06 ± 1.30 (59)	0.02
	BA z score	−1.99 ± 1.38 (108)	−2.56 ± 1.44 (53)	NS
	IGF-I z score	−1.67 ± 0.62 (100)	−1.55 ± 0.69 (49)	NS
	BMI	16.64 ± 2.38 (105)	16.30 ± 3.60 (59)	NS
GH deficient	*n*/total	11/120	8/60	NS
	CA (y)	7.57 ± 3.81 (11)	7.54 ± 4.30 (8)	NS
	ht z score	−2.74 ± 1.15 (11)	−3.51 ± 1.28 (8)	NS
	BA z score	−1.77 ± 1.09 (11)	−3.83 ± 1.99 (7)	0.01
	IGF-I z score	−1.87 ± 0.86 (11)	−2.48 ± 0.26 (5)	NS
	BMI	17.70 ± 2.82 (11)	17.90 ± 5.34 (8)	NS
GH sufficient	*n*/total	101/120	52/60	NS
	CA (y)	9.44 ± 4.56 (101)	8.36 ± 4.07 (52)	NS
	HT z score	−2.73 ± 0.98 (101)	−2.99 ± 1.31 (51)	NS
	BA z score	−2.01 ± 1.42 (97)	−2.37 ± 1.26 (46)	NS
	IGF-I z score	−1.64 ± 0.58 (89)	−1.44 ± 0.65 (44)	NS
	BMI	16.52 ± 2.31 (94)	16.10 ± 3.26 (51)	NS

* Male and female subjects were separated and compared for the six variables indicated, for the group as a whole and according to final diagnosis. Note that eight males failed the outpatient screening test but were lost to follow-up, and are not included in these tables. The girls had a significantly lower ht z score for the group as a whole, and a lower BA z score in the GH-deficient subgroup. A $p < 0.05$ was considered significant. Numbers in parentheses indicate the number of cases for which data were available. NS, not significant.

Table 3. *Cross-correlations with GH levels**

GH level		CA	ht-z	BA-z	IGFI-z	BMI
Screening test						
Baseline	*r*	−0.30	−0.25	−0.06	−0.14	−0.26
	p	0.0003	0.004	NS	NS	0.003
	n	136	136	129	121	131
Peak	*r*	0.06	0.02	0.11	0.21	−0.12
	p	NS	NS	NS	0.02	NS
	n	139	139	132	123	134
Arginine-insulin test						
Peak	*r*	−0.14	−0.21	−0.45	−0.06	−0.23
	p	NS	NS	0.003	NS	NS
	n	45	44	42	36	41

* Baseline and peak screening-test, and peak second-test GH levels are cross-correlated with the five variables indicated for the GH-sufficient group only. The correlation coefficient (*r*), *p* value, and number of cases (*n*) are shown for each correlation. Numbers in parentheses indicate the number of cases for which data was available. See text for further discussion. NS, not significant.

then approached the issue of whether provocative GH-testing could have been avoided from a practical standpoint using variables which are generally available during a routine clinical evaluation.

Growth parameters were of limited diagnostic use. Although 12 of 14 children with ht z score above −1.65 (5th centile) were found to be GH sufficient, the clinical significance of this is uncertain. Many children with acquired GH deficiency (*e.g.* due to surgery or irradiation) may initially have a ht >5th centile, although they are often not evaluated for GH sufficiency until their ht falls below the 5th centile. The two GH-deficient children in our series with ht z score > −1.65 had received cranial irradiation.

The fact that only a limited proportion of our subjects had accurate growth data that would permit determination of pre-testing growth velocity is not surprising from a practical standpoint. Many of the children received GH testing within 3 mo after initial referral. Before evaluation, most of the children had been followed only intermittently elsewhere, with growth measurements that are not directly comparable with our own. At least 10 mo of growth data are recommended for calculation of growth velocity (29). It has been our clinical impression that growth velocities are highly variable for individual short children, and this is supported in the literature (30–34). Furthermore, although growth velocity has been advocated as an indicator of GH status, we found that statural growth velocity was <5th centile for age for only four of 10 children recently diagnosed with GH deficiency (Lee PDK, unpublished data). Therefore, although statural growth velocity is undoubtedly decreased in GH-deficient children (35), the availability and utility of this observation in an individual patient may be limited.

The degree of BA delay, as measured by BA z score, tended to be greater for the GH-deficient group, but this did not attain clinical significance. This is probably due to the large population of children with "constitutional delay" who, by definition, have delayed skeletal maturation (36).

IGF-I levels have been advocated as a screening procedure for GH deficiency on the basis of studies showing decreased levels of IGF-I in GH-deficient individuals (10, 13, 14). However, the diagnostic utility of this observation is uncertain. Using a commercial assay, Reiter and Lovinger (37) found low IGF-I levels in seven of 25 (28%) GH-sufficient short children and 12 of 16 (75%) GH-deficient children. Moore *et al.* (38) obtained IGF-I levels for 143 short children, and performed GH testing only for the 78 who had an IGF-I level <0.5 U/mL. Of this group, nine were found to be GH deficient. GH levels were not measured for the 65 children with IGF-I levels >0.5 U/mL, therefore the diagnostic sensitivity and specificity of IGF-I levels could not be calculated. Furthermore, both of these studies relied on an assay that does not incorporate a procedure for removal of endogenous IGF-binding proteins that may interfere with the IGF-I assay (15, 39, 40).

IGF-I levels were found to be low for 13 GH-deficient patients reported from Brazil as compared to nine short GH-sufficient children (41). Rayner *et al.* (42) studied IGF-I levels for 32 GH-deficient and 27 GH-sufficient short children and found that IGF-I levels were low for 62.5% of the GH-deficient children. The sensitivity was higher for children with BA < 8 y (89%), and lower for the 14 children with BA > 8 y (29%). This was confirmed in our group using a BA boundary of 7 y. However,

the predictive value of this observation in a given clinical situation is limited.

In a previous study of 68 GH-deficient children, we reported that 82% had an IGF-I level below the 95th centile confidence limit constructed from 197 children with ht between the 5th and 95th centile for age (43). However, 32% of the 44 GH-sufficient children with ht < 5th centile also had a low IGF-I level, resulting in poor discrimination between populations. IGF-I levels for this previous study were measured after acidification and column-chromatography of the samples for removal of IGF-binding proteins.

In our report, we found that age-adjusted levels of IGF-I measured using a commercial assay, with removal of endogenous IGF-binding protein by acid/ethanol treatment, may have diagnostic value for GH deficiency. However, 100% sensitivity is not achieved even with use of a 1 SD normal range, whereas specificity at this level is extremely poor. The best discrimination is found if the absolute age-related norms provided by the laboratory are used, although three of 16 cases of GH-deficiency (19%) would still have been missed if IGF-I levels alone had been used as a screening test.

Discriminant analysis resulted in a unique linear combination of variables that would have allowed GH-deficiency to be excluded without provocative testing for 58% of the GH-sufficient population, whereas allowing eventual diagnosis of all cases of GH deficiency. Of the variables included, IGF-I levels were the best single predictor of eventual outcome for provocative GH testing; and the sensitivity and specificity of the IGF-I level may be increased by consideration of other easily obtained clinical data. However, although the constructed predictive variable is indicative of the relative value of individual clinical measures, the precise formula should not be generalized to other clinic populations, especially if a different method for IGF-I determination is used.

The 2:1 male to female ratio in our clinic population evaluated for GH deficiency is typical of the referral patterns reported by others (35, 44). This ratio raised a concern that girls with GH deficiency may be underdiagnosed, or that the evaluation of girls for GH deficiency may be delayed. Although we did detect a tendency in the group as a whole for girls to have a greater deficit in ht at the time of GH testing, these differences were not significant when the population was grouped according to GH sufficiency. For the GH-deficient subgroup, girls tended to be evaluated at a greater degree of BA delay. Therefore, the evaluation of girls for GH deficiency may be delayed relative to boys, but there is no evidence that this compromises our ability to diagnose GH deficiency in girls.

We also examined the GH-sufficient population to determine the factors that may influence the degree of GH response on provocative testing. The negative correlation of basal fasting GH with age may be related to the previously reported prepubertal decrease in GH secretion in short children (21, 22, 45, 46); although the peak GH responses did not show this correlation. The negative correlation of second-test peak GH levels with the degree of bone age delay may also be related to this factor. We found no correlation of IGF-I z scores and basal GH levels, which is not surprising given the pulsatile nature of GH secretion; although IGF-I z scores did vary with peak screening test GH levels. The decline of peak cortisol levels with age has not been previously reported and may deserve further investigation. Overall, when considered from a clinical standpoint, no particular variable was useful in predicting the actual degree of GH response on provocative testing.

IGF-I levels are commonly obtained during routine evaluation of children with growth retardation, although data relating to the clinical interpretation of IGF-I levels is limited. Our data, derived from a routine pediatric endocrinology clinic, indicate that IGF-I levels may have clinical use in predicting response to provocative GH testing, particularly if considered concurrently with other clinical variables. In this sense, IGF-I levels may help to guide the evaluation of an individual short child. Although application of our constructed predictive variable is limited to our specific study population, our findings provide a unique perspective on the utility of IGF-I assays and provide direction for further study.

REFERENCES

1. Rimoin DL, Horton WA 1978 Short stature. Part II. J Pediatr 92:697–704
2. Frasier SD 1981 Short stature in children. Pediatr Rev 3:171–179
3. Frasier SD 1974 A review of growth hormone stimulation tests in children. Pediatrics 53:929–936
4. Alsever RN, Gotlin RW 1978 Handbook of Endocrine Tests in Adults and Children, 2nd ed. Year Book Medical Publishers, Chicago, pp 5–14
5. Ad Hoc Committee on Growth Hormone Usage, The Lawson Wilkins Pediatric Endocrine Society, and Committee on Drugs, American Academy of Pediatrics 1983 Growth hormone in the treatment of children with short stature. Pediatrics 72:891–894
6. Parkin JM 1986 Exercise as a test of growth hormone secretion. Acta Endocrinol [Suppl] 279:47–50
7. Sukegawa I, Hizuka N, Takano K, Asakawa K, Horikawa R, Hashida S, Ishikawa E, Mohri ZI, Murakami Y, Shizume K 1988 Urinary growth hormone (GH) measurements are useful for evaluating endogenous GH secretion. J Clin Endocrinol Metab 66:1119–1123
8. Albini CH, Quattrin T, Vandlen RL, MacGillivray MH 1988 Quantitation of urinary growth hormone in children with normal and abnormal growth. Pediatr Res 23:89–92
9. Hintz RL 1981 The somatomedins. Adv Pediatr 28:293–317
10. Zapf J, Walter H, Froesch ER 1981 Radioimmunological determination of insulinlike growth factors I and II in normal subjects and in patients with growth disorders and extrapancreatic tumor hypoglycemia. J Clin Invest 68:1321–1330
11. Minuto F, Underwood LE, Grimaldi P, Furlanetto RW, Van Wyk JJ 1981 Decreased serum somatomedin C concentrations during sleep. Temporal relationship to the nocturnal surges of growth hormone and prolactin. J Clin Endocrinol Metab 52:399–403
12. Clemmons DR, Van Wyk JJ 1984 Factors controlling blood concentrations of somatomedin-C. Clin Endocrinol Metab 13:113–143
13. Kao PC, Abboud CF, Zimmerman D 1986 Somatomedin C: an index of growth hormone activity. Mayo Clin Proc 61:908–909
14. Underwood LE, D'Ercole AJ, Van Wyk JJ 1980 Somatomedin-C and the assessment of growth. Pediatr Clin North Am 27:771–782
15. Daughaday WH, Kapadia M, Mariz I 1987 Serum somatomedin binding proteins: physiologic significance and interference in radioligand assay. J Lab Clin Med 109:355–363
16. SAS User's Guide: Statistics 1982 SAS Institute, Cary, NC
17. Hamill PVV, Drizd TA, Johnson CL, Reed RR, Roche AF 1977 NCHS growth curves for children from birth to 18 years. Vital Health Stat 165:1–74
18. Greulich WW, Pyle SI 1959 Radiographic Atlas of Skeletal Development of the Hand and Wrist. Stanford University Press, Stanford, CA
19. Levin PA, Chalew SA, Martin L, Kowarski AA 1987 Comparison of assays for growth hormone using monoclonal or polyclonal antibodies for diagnosis of growth disorders. J Lab Clin Med 109:85–88
20. Celniker AC, Chen AB, Wert Jr RM, Sherman BM 1989 Variability in the quantitation of circulating growth hormone using commercial immunoassays. J Clin Endocrinol Metab 68:469–476
21. Gourmelen M, Pham-Huu-Trung MT, Girad F 1979 Transient partial hGH deficiency in prepubertal children with delay of growth. Pediatr Res 13:221–224
22. Bierich JR 1983 Treatment of constitutional delay of growth and adolescence with human growth hormone. Klin Padiatr 195:309–316
23. Lanes R, Bohoquez L, Leal V, Hernandez G, Borges M, Hurtado E, Moncada G 1986 Growth hormone secretion in patients with constitutional delay of growth and pubertal development. J Pediatr 109:781–783
24. Clayton PE, Price DA, Shalet SM 1987 Growth hormone state after completion of treatment with growth hormone. Arch Dis Child 62:222–226
25. Sanz MT, Rochiccioli P 1985 Comparison of spontaneous growth hormone secretion during daytime and sleep in children with short stature. Horm Res 22:17–23
26. Ward PS, Savage DCL 1985 Growth hormone responses to sleep, insulin hypoglycemia and arginine infusion. Hormone Res 22:7–11
27. Tauber MT, Uboldi F, Pienkowksi C, Chatelain P, Rochiccioli P 1987 Comparison of 24 hour growth hormone secretion and pharmacological tests in 38 children with short stature. J Pediatr Endocrinol 2:101–113
28. Rose SR, Ross JL, Uriarte M, Barnes KM, Cassorla FG, Cutler GB 1988 The advantage of measuring stimulated as compared with spontaneous growth hormone levels in the diagnosis of growth hormone deficiency. N Engl J Med 319:201–207
29. Tanner JM, Davies PW 1985 Clinical longitudinal standards for height and height velocity for North American children. J Pediatr 107:317–329
30. Horner JM, Thorsson AV, Hintz RL 1978 Growth deceleration patterns in children with constitutional short stature: an aid to diagnosis. Pediatrics 62:529–534
31. Rimoin DL, Horton WA 1978 Short stature. Part I. J Pediatr 92:523–528
32. Tanner JM 1986 Physical development. Br Med Bull 42:131–138
33. Polychronakos C, Abu-Srair H, Guyda HJ 1988 Transient growth deceleration in normal short children. A potential source of bias in growth studies. Eur J

Pediatr 147:582–583
34. Marshall WA 1971 Evaluation of growth rate in height over periods of less than one year. Arch Dis Child 46:414–420
35. Hindmarsh P, Smith PJ, Brook CGD, Matthews DR 1987 The relationship between height velocity and growth hormone secretion in short prepubertal children. Clin Endocrinol 27:581–591
36. Rosenfeld RG 1982 Evaluation of growth and maturation in adolescence. Pediatr Rev 4:175–183
37. Reiter EO, Lovinger RD 1981 The use of a commerically available somatomedin-C radioimmunoassay in patients with disorders of growth. J Pediatr 99:720–724
38. Moore DC, Ruvalcaba RHA, Smith EK, Kelley VC 1982 Plasma somatomedin-C as a screening test for growth hormone deficiency in children and adolescents. Hormone Res 16:49–55
39. Lee PDK, Rosenfeld RG 1987 Clinical utility of insulin-like growth factor assays. Pediatrician 14:154–161
40. Powell DR, Rosenfeld RG, Baker BK, Liu F, Hintz RL 1986 Serum somatomedin levels in adults with chronic renal failure: the importance of measuring insulin-like growth factor I (IGF-I) and IGF-II in acid-chromatographed uremic serum. J Clin Endocrinol Metab 63:1186–1192
41. Gianella-Neto D, Wajchenberg BL, Mendonca BB, Liberman B, Nery M, Macchione M, Spencer EM 1988 Somatomedin levels in the diagnosis and therapy of growth hormone deficiency. Horm Metab Res 20:175–181
42. Rayner PHW, Rudd BT, Thomas PH, Williams JW 1988 Growth hormone deficiency and the measurement of somatomedin C/IGF-I: the influence of sexual maturation. Clin Endocrinol 28:361–371
43. Rosenfeld RG, Wilson DM, Lee PDK, Hintz RL 1986 Insulin-like growth factors I and II in evaluation of growth retardation. J Pediatr 109:428–433
44. Aynsley Green A, MacFarlane JA 1983 Method for earlier recognition of abnormal stature. Arch Dis Child 58:535–537
45. Greene SA, Torresani T, Prader A 1987 Growth hormone response to a standardised exercise test in relation to puberty and stature. Arch Dis Child 62:53–56
46. Josefsberg Z, Kauli R, Keret R, Brown M, Bialik O, Greenberg D, Laron Z 1983 Tests for hGH secretion in childhood. Comparison of response of growth hormone to insulin hypoglycemia and to arginine in children with constitutional short stature in different pubertal stages. Pediatr Adolesc Endocrinol 12:66–74