Analysis of Bone Age Data From National Cooperative Growth Study Substudy VII

Stephen F. Kemp, MD, PhD*, and Judy P. Sy, PhD‡

ABSTRACT. National Cooperative Growth Study sub-study VII was conducted 1) to compare standardized hand–wrist and knee bone age determinations in pubertal children treated with growth hormone (GH); 2) to compare local determinations of bone ages with centrally determined bone ages; 3) to relate the response to GH therapy to the bone age determinations; and 4) to ascertain the predictive value of each type of bone age determination. Eligible subjects were those in the National Cooperative Growth Study who were at Tanner pubertal maturation. Eligible subjects were those in the National Cooperative Growth Study who were at Tanner pubertal maturation. The strongest correlation was between the hand–wrist and knee values, while the remaining correlation coefficients were less strong. There were 990 patients in the study; in 925 (677 boys), there were both hand–wrist and knee bone age determinations from the baseline pubertal radiographs. There was only one radiographic assessment in 496 patients, two in 205 patients, and three to eight in the remaining patients. The strongest correlation was between the hand–wrist bone age and the hand–wrist plus knee bone age (r = .95). Also strongly correlated were knee with hand–wrist (r = .872) and knee with hand–wrist plus knee (r = .914). For none of these bone age methods was any statistically significant difference found between the methods. The locally determined bone ages correlated strongly with the centrally determined bone ages for knee (r = .850), hand–wrist (r = .928), and hand–wrist plus knee (r = .930); however, the locally determined knee and hand–wrist values were less (by ~0.3 year) than the centrally determined values. These differences, however, do not appear to be clinically significant. Pediatrics 1999;104:1031–1036; bone age determinations, recombinant human growth hormone, hand–wrist radiographs, knee radiographs.

ABBREVIATIONS. GH, growth hormone; NCGS, National Cooperative Growth Study; SDS, standard deviation score(s).
An integral part of growth hormone (GH) therapy is using bone age determinations to assess the patient’s growth potential and to monitor the effect of therapy on his or her predicted adult height. Bone ages are estimated by the methods of Greulich and Pyle, Tanner-Whitehouse, or Fels for the hand–wrist1 and the method of Roche, Wainer, and Thissen for the knee.2 Each of these methods was developed from studies in children who did not have GH deficiency and were not being treated with GH. The Greulich and Pyle standards, which date from 1950 and are the methods used most commonly for estimating bone age, were developed from a sample of white children who were above average in economic and educational status.3 These standards do not reflect accurately the present population of normally growing children and do not represent the current population of children with growth disorders.

The period of greatest uncertainty in assessing bone ages is the last stage of puberty, ie, in children maturing at an average rate, the ages of ~15 to 17 years (boys) and 14 to 16 years (girls), when skeletal maturation is proceeding rapidly under the influence of sex steroids and GH. The hand and wrist epiphyses frequently fuse before adult height is achieved. Garn et al4 have reported that boys may grow as much as 1.6 cm and girls as much as 2.3 cm after complete hand–wrist epiphyseal fusion. The time of fusion of the hand–wrist and of the knee may differ by as much as 2 years.3

The National Cooperative Growth Study (NCGS) substudy VII was conducted1 to compare standardized hand–wrist and knee bone age determinations in pubertal children who were being treated with GH,2 to compare locally determined bone ages with centrally determined hand–wrist and knee bone ages,3 to relate the response to GH therapy to bone age determinations, and4 to ascertain the predictive value of each type of bone age determination.

METHODS

Subjects

All subjects were being treated with GH and were enrolled in the NCGS; a postmarketing surveillance study established in 1985 to monitor the use, safety, and effectiveness of recombinant human GH manufactured by Genentech, Inc (South San Francisco, CA). There were >30,000 children enrolled in the NCGS at the time of this analysis.

Any patient in the NCGS who was at Tanner pubertal stage 2 or greater for breasts (girls) or genitals (boys) was eligible for inclusion in the NCGS substudy VII. The subjects were examined routinely by their physicians at intervals of 3 to 4 months (not to exceed 6 months) until they reached adult or near-adult height. If a subject discontinued GH therapy, then height measurements and bone age determinations were encouraged until he or she attained adult height.

Radiographs of the knee for determining bone ages were taken at least at the beginning of the study and, if there was more than one radiograph, near the last visit. Hand–wrist radiographs were required at least annually. All radiographs were submitted to the Fels Institute (Yellow Springs, OH). Bone ages for each subject were determined separately from the hand–wrist and knee radiographs. A combined hand–wrist and knee bone age determination was made. These were determined by recording the grades of the indicators (landmarks on specific bones) using published descriptions.1,2 The estimates of the combined hand–wrist and knee bone age were based on all the data available in the pair of radiographs and thus are not necessarily equal to the mean of the separate bone ages.

For purposes of identifying subjects who reached near-adult height, near-adult height was defined as the last available height of a patient if both the chronologic age and the bone age were at least 14 years (girls) or 16 years (boys). If a bone age was not available at the time of the last height, the bone age was extrapolated by adding the time difference between the date of the last height measurement and the date of the last radiograph (up to 3 years) to the last bone age available. If the chronologic age at the last height was 18 years (women) or 20 years (men), then the last height was defined as near-adult height regardless of the bone age at the last height.

Statistical Analysis

Analyses were based on two subject subsets: 1) subjects with baseline pubertal knee and hand–wrist radiographs taken during the study, and 2) subjects in the first subset who had pubertal treatment follow-up data and had reached near-adult height. Standard deviation scores (SDS) for height were calculated as height SDS = (achieved height – mean height of normal children of the same age and sex)/SD for height of normal children of the same age and sex. The height standards for normal children were those reported by the National Center for Health Statistics.

Descriptive statistics are given as frequencies and percentages for categoric data and as means and SD for continuous variables. Pearson correlation coefficients and the corresponding P values were obtained to assess the linear relation between the bone age methods. The differences in bone age determinations between methods within each patient were obtained, and significance tests were performed by using the paired t test.

The endpoint for pubertal growth response was defined as the cumulative change in height SDS from NCGS substudy VII baseline to near-adult height. This was computed as the last height SDS, standardized based on age at the last height, minus the height SDS at NCGS substudy VII baseline. A multiple regression analysis was performed to determine the variables that predict pubertal growth response. The predictor variables included demographics, patient characteristics at NCGS substudy VII baseline, treatment variables, and bone age variables. The following predictor variables were considered in the analysis: sex; midparental height SDS; log of maximum stimulated GH level at enrollment in the NCGS; pubertal status at NCGS substudy VII baseline; age at NCGS substudy VII baseline; height SDS at NCGS substudy VII baseline; body mass index SDS at NCGS substudy VII baseline; NCGS substudy VII pubertal GH treatment time; posttreatment follow-up time; log of GH dose (mg/kg/wk) at NCGS substudy VII baseline; number of injections per week at NCGS substudy VII baseline; previous GH treatment time (from enrollment in the NCGS to bone age delay at NCGS substudy VII baseline); and bone age delay at NCGS substudy VII baseline (based on knee bone age, hand–wrist bone age, and locally determined bone age). Various models were fit to determine the following: 1) whether the knee, hand–wrist, and locally determined bone ages each were predictive of pubertal growth in response to GH treatment, given the other predictors; 2) predictive value of knee bone age, given hand–wrist bone age; 3) predictive value of knee bone age, given locally determined bone age; and 4) predictive value of hand–wrist bone age, given locally determined bone age. The importance of each predictor was assessed by using the P values obtained from the F tests of significance of the corresponding regression coefficients. The coefficient of multiple determinations (R2), which measures the proportion of variance of the response that is explained by the predictors in the model, is reported and was used in comparing models. P values < .05 were considered statistically significant.

RESULTS

Patients with Pubertal Radiographs

In the NCGS core study, there were 990 patients who were considered for substudy VII. A total of 1948 sets of hand–wrist and/or knee radiographs were made in these patients during the study. To be eligible for the study, patients had to be pubertal (Tanner stage 2 or greater) at the time of enrollment. We identified 932 of these patients as being
pubertal at the time of a radiographic assessment (either at enrollment or at a later time). The remaining 58 patients consisted of those who were prepubertal at the time of all radiographic assessments (40 patients) or in whom there were insufficient data for ascertaining whether they were pubertal (18 patients).

In these 932 patients, we identified the first radiographic assessment at which the patient was pubertal (the NCGS substudy VII baseline pubertal radiograph). A height measurement in the NCGS core study was matched to an NCGS substudy VII radiographic assessment if the radiograph had been made within 90 days of the height measurement (the nearest visit). Data at this visit were necessary for computing the Bayley-Pinneau predicted adult height at the time of the radiograph. In these 932 patients, 1869 radiographs were made, 1829 of which were made during puberty.

In 925 of these patients, there was a complete pair of hand–wrist and knee bone age determinations at the baseline pubertal radiographic assessment. There were a total of 1814 pubertal radiographs in these 925 patients. In 496 subjects, there was a single pubertal radiographic assessment, in 205 there were two, and in 224 there were three to eight. In most of the patients with follow-up radiographic assessments, there was a complete pair of hand–wrist and knee radiographs.

The correlations and differences between the various methods of determining bone age at baseline are shown in Table 1. The correlation between the hand–wrist bone age and the knee bone age was strong, with few values >2 SD from the mean, as shown in Fig 1. The distribution of the differences between the knee and the hand–wrist bone ages is shown in Fig 2. The combined hand–wrist and knee bone age determination also correlated strongly with the hand–wrist bone age \(r = .995, P < .0001\) as well as with the knee bone age \(r = .914, P < .0001\). The different centrally determined bone ages were not statistically different from each other. In 684 subjects, it was possible to compare the centrally determined baseline hand–wrist bone ages with the locally determined bone ages. These determinations are shown in Fig 3. There was a strong linear correlation \(r = .928\) between the centrally determined and the locally determined bone ages. The centrally determined bone ages were statistically greater than the locally determined bone age by \(-0.3\) year \((P < .0001)\), but the differences did not appear to be clinically significant.

Patients were enrolled in NCGS substudy VII at an average of 3.0 ± 1.9 years after enrollment in the NCGS. They then continued treatment with GH for an average of 1.5 ± 1.2 years. In many patients, the first pubertal radiograph was obtained at their last reported visit during GH treatment or several years after that, which did not permit analysis of growth in relation to changes in the bone age during puberty in these patients.

### Table 1. Correlations of and Differences in BA Determinations Using the Baseline Pubertal Radiographs

<table>
<thead>
<tr>
<th>n</th>
<th>Correlation</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(r)</td>
<td>(P)</td>
</tr>
<tr>
<td>Knee and hand–wrist</td>
<td>925</td>
<td>.872</td>
</tr>
<tr>
<td>Knee and hand–wrist + knee</td>
<td>925</td>
<td>.914</td>
</tr>
<tr>
<td>Hand–wrist and hand–wrist + knee</td>
<td>925</td>
<td>.995</td>
</tr>
<tr>
<td>Knee and locally determined</td>
<td>684</td>
<td>.850</td>
</tr>
<tr>
<td>Hand–wrist and locally determined</td>
<td>684</td>
<td>.928</td>
</tr>
<tr>
<td>Hand–wrist + knee and locally determined</td>
<td>684</td>
<td>.930</td>
</tr>
</tbody>
</table>

---

**Fig 1.** Correlation of knee bone age with hand–wrist bone age at enrollment in NCGS substudy VII (n = 925). The solid line is the best linear fit of the data \((r = .872, P < .0001)\). The dashed line is the line of identity. The dotted lines are the line of identity ± 0.7 year (±1 SD from the mean of repeated bone age determinations from the same radiograph).

**Fig 2.** Distribution of the difference between the knee and the hand–wrist bone age determinations.
Patients With Near-adult Heights

There were 315 subjects with NCGS substudy VII baseline height and radiographic bone age (hand–wrist and knee) data and pubertal treatment follow-up data who had reached near-adult height. We analyzed the pubertal growth during GH therapy in these patients as defined by the change in height SDS from the NCGS substudy VII baseline to near-adult height. Of these 315 patients, 253 (80.3%) had discontinued GH treatment; the remaining subjects were either still being treated with GH or had become inactive in the study. There were posttreatment follow-up height data for 155 subjects (49.2%).

Complete data on the predictor variables selected for analysis that could be examined in terms of models that predicted the response to GH therapy were available for 245 subjects. Of these, 218 subjects also had locally determined bone ages. The following models were considered:

- Model A—All variables except bone age delay, which include sex; midparental height SDS; log of maximum stimulated GH level; pubertal stage at NCGS substudy VII baseline; height SDS at NCGS substudy VII baseline; body mass index at NCGS substudy VII baseline; NCGS substudy VII pubertal treatment time; posttreatment follow-up time; log of GH dose at NCGS substudy VII baseline; number of injections of GH per week at NCGS substudy VII baseline; and previous GH treatment time (basic variables) ($R^2 = .477$)
- Model B—Basic variables, knee bone age delay ($R^2 = .502$)
- Model C—Basic variables, hand–wrist bone age delay ($R^2 = .501$)
- Model D—Basic variables, knee bone age delay, hand–wrist bone age delay ($R^2 = .508$)
- Model E—Basic variables, locally determined bone age delay ($R^2 = .499$)
- Model F—Basic variables, locally determined bone age delay, knee bone age delay ($R^2 = .510$)
- Model G—Basic variables, locally determined bone age delay, hand–wrist bone age delay ($R^2 = .507$)

In none of these models was the dose of GH, maximum stimulated GH level, body mass index SDS, height SDS, or posttreatment follow-up time statistically significant. Duration of GH treatment, age at enrollment in NCGS substudy VII, midparental height SDS, and number of injections of GH per week all had positive effects on pubertal growth, and duration of previous GH treatment had a negative effect. A quadratic function for duration of previous treatment provided a better fit than did a linear function, which is consistent with the observation that the height SDS begins to level off after 4 to 5 years of treatment in those who had been treated previously with GH. The pubertal stage at enrollment in NCGS substudy VII contributed only minimally to pubertal growth (models A, E, F, and G); those at an earlier pubertal stage appeared to grow more than did those at a later stage. Sex was statistically significant in the model without bone age delay (model A), with the growth response being less in boys than in girls, but was not statistically significant in any of the models with bone age delay.

In model B, knee bone age delay was a significant predictor ($P = .0008$), given the other variables in the model. Its positive effect indicates that the increase in height SDS was greater in those in whom the bone age delay was greater. However, the increase in $R^2$ was small (.025—from .477 to .502). The results for hand–wrist bone age delay and locally determined bone age delay followed a pattern similar to that of knee bone age delay. In model C, hand–wrist bone age delay also was a significant predictor ($P = .001$), given the other basic variables in the model, with a positive effect on the endpoint. But the increase in $R^2$ was also small (.024—from .477 to .501). Likewise, in model E, locally determined bone age delay was a significant positive predictor ($P = .0037$), given the other basic variables in the model. But again, the increase in $R^2$ was not impressive (.022—from .477 to .499).

In model D, we find that having both knee bone age delay and hand–wrist bone age delay results in redundancy. Given hand–wrist bone age delay, we do not need knee bone age delay ($P = .083$). Likewise, given knee bone age delay, we do not need hand–wrist bone age delay ($P = .101$). In models F and G, we assess whether having either of the centrally determined bone ages adds significant information, given that we have the locally determined bone age. The results show that hand–wrist bone age delay adds some, but not significant, information, given locally determined bone age delay (model G: $P = .076$, $R^2$ increases by .008—from .499 to .507). On the other hand, knee bone age delay, although adding statistically significant information to locally de-
determined bone age delay, does not increase the $R^2$ to a great extent (model F: $P = .039$, $R^2$ increases by .011—from .499 to .510).

The two treatment time variables were consistently the most significant variables in all the models examined.

From the near-adult height data, we examined the relation between GH treatment time and the change in height SDS during the course of the study. The change in height SDS in these patients according to the duration of GH therapy during puberty is shown in Fig 4. There was a positive effect of treatment time during puberty, indicating that GH therapy during puberty continues to result in increases in the height SDS.

The change in height SDS according to the duration of GH therapy before enrollment in NCGS sub-study VII is shown in Fig 5, indicating that the increase in height SDS during puberty is greater when the preenrollment duration of therapy is shorter. The contribution of pubertal growth to the difference in height SDS decreases with longer preenrollment treatment, from 1 SDS with 0 year to no effect with 5 years of prepubertal therapy.

**CONCLUSIONS**

The NCGS sub-study VII was a multicenter study that examined the role of bone age determinations in children who were treated with GH during puberty. There were 1948 sets of centrally read radiographs (hand–wrist and/or knee) from 990 patients. All these patients were enrolled in the NCGS, which made it possible to compare the centrally estimated bone ages with the locally estimated bone ages and height data collected during the study.

There was a strong correlation between the bone ages determined from the hand–wrist and those determined from the knee radiographs, and the differences between these two determinations were not statistically significant. The linear fit of these data deviates from the line of identity, such that at older chronologic ages, the hand–wrist bone age tends to be greater than the knee bone age, and at younger chronologic ages, it tends to be less. It is more difficult to compare the combined hand–wrist and knee bone age with either of the separate bone ages alone, because the combined (hand–wrist + knee) determination included many of the same skeletal indicators used in the separate determinations. Thus, the extremely strong correlation of the combined bone age with the hand–wrist bone age ($r$ is close to 1) indicates that the hand–wrist plus knee method is dominated by the same indicators that are used in the hand–wrist method.

The locally determined bone ages (those in the NCGS core study database) correlated well with the centrally determined bone ages. It is apparent from Fig 3 that the locally determined bone ages represent a discrete variable (likely resulting from comparisons with the examples in the Greulich and Pyle atlas8), and the Fels method yields a continuous variable. The stronger correlation of the locally determined bone age with the centrally determined hand–wrist or combined hand–wrist plus knee bone age is likely because almost all the locally determined bone ages are from hand–wrist radiographs (in most cases, the same radiograph that was used for the central determination). Centrally determined bone ages may be preferred when exact values are required, as in clinical studies, but for most clinical purposes, locally determined bone ages are sufficient.

There were sufficient data in 315 patients for examining the effect of GH therapy on growth during puberty. The data indicate clearly that treatment with GH during puberty results in an increase in the height SDS. Consistent with previous observations in prepubertal children,9 the duration of GH therapy is a major factor in the extent of the response. The increase in height SDS also depends on the duration of GH therapy before puberty; the longer the duration of GH therapy before puberty, the less the increase in height SDS during puberty. One possible explanation for this is that because catch-up growth occurs early in the course of GH therapy, those who
enter puberty with little exposure to GH therapy may be experiencing catch-up growth after the onset of puberty. Continued use of GH in those with greater exposure before puberty serves only to maintain the gain in height SDS that was achieved before puberty.

The different models all predicted the response to GH therapy during puberty with $R^2$ values near .5. Without bone age delay as a variable, the $R^2$ was .477; adding bone age data, whether from locally determined bone ages, centrally determined bone ages, or from hand–wrist or knee radiographs, increased the $R^2$ value to ~.51. Except for age at enrollment in NCGS substudy VII, the effects of other significant variables are intuitive and consistent with our expectations. The positive effect of age at enrollment on the change in height SDS appeared to be counterintuitive. On further analysis, it was found that the age at the last visit during GH treatment was positively correlated with our defined growth response endpoint ($r = .389$), which suggests that the older a patient was at the last visit, the greater was the response to treatment. Because the age at NCGS substudy VII baseline and the age at the last visit on treatment were positively correlated, the effect of the first age variable was confounded with the effect of the second age variable.

Data from NCGS substudy VII indicate that bone age determinations are an important predictor of the response to GH therapy, even during puberty. It appears that in most patients, knee radiographs do not provide sufficient additional information to justify their being used when the hand–wrist bone age does not reflect the clinical situation, and a knee bone age determination is warranted. Locally determined bone ages appear to provide adequate clinical information, and centrally determined bone ages are unnecessary except in studies in which precise bone age determinations are important.

**ACKNOWLEDGMENTS**

Supported by an educational grant from Genentech, Inc, South San Francisco, CA.

We thank Ken Dana for preparing the dataset and program support, and Sandra L. Blethen, MD, PhD, for helpful suggestions.

**REFERENCES**

2. Roche AF, Wainer H, Thissen D. Skeletal Maturity: The Knee Joint as a Biological Indicator. New York, NY: Plenum Medical Book Co; 1975

**Regular Monitoring of Bone Age Is Not Useful in Children Treated With Growth Hormone**

Darrell M. Wilson, MD

**ABSTRACT.** Although bone age estimates are traditionally used to monitor children receiving growth hormone therapy, few data support this practice. Bone age determination is fraught with technical difficulties, resulting in high interobserver differences. Longitudinal studies show that an individual’s bone age can change erratically over time. The resulting errors in predicted adult heights based on these bone age determinations are large. Moreover, growth hormone therapy appears to accelerate bone maturation. The radiographic evidence of this acceleration can be delayed. In this setting, improvements in predicted adult heights can be artifactual large. Routine monitoring of bone age during GH therapy is unnecessary. *Pediatrics* 1999;104:1036–1039; *Bayley and Pinneau, bone age determination, Greulich and Pyle, predicted height, radiography, Tanner and Whitehouse.*

**ABBREVIATION.** GH, growth hormone.

Monitoring is useful only if it guides the hand of therapy. In this article, I review the role of repeated bone age determinations in individual patients during growth hormone (GH) therapy. For monitoring, as for diagnosis, the tests used should be accurate and reproducible. Moreover, they should help predict some clinically important out-