Research Article

What is the Most Effective Method for Predicting Adult Height in Boys with Constitutional Delay of Growth and Puberty?

Akın Kağızmanlı G et al. Predicting Adult Height in Boys with Constitutional Delay of Growth and Puberty

Gözde Akın Kağızmanlı[1], Deniz Özalp Kızılay[2], Reyhan Deveci Sevim[3], Kübra Yüksek Acinikli[1], Fulya Mete Kalaycı[2], Ayşegül Tekneci[3], Korcan Demir[1], Ece Böber[1], Ahmet Anık[3], Samim Özen[2], Ayhan Abacı[1]
[1] Dokuz Eylül University Faculty of Medicine, Department of Pediatric Endocrinology, İzmir, Türkiye
[2] Ege University Faculty of Medicine, Department of Pediatric Endocrinology, İzmir, Türkiye
[3] Aydın Adnan Menderes University Faculty of Medicine, Department of Pediatric Endocrinology, Aydın, Türkiye

What is already known on this topic?
Predicted adult height can be calculated using methods like Bayley-Pinneau, Roche-Wainer-Thissen, and BoneXpert, which rely on bone age assessment. However, as these methods were originally developed for healthy children, their predictive accuracy has shown variability when applied to different patient groups.

What this study adds?
This study demonstrated that the Roche-Wainer-Thissen (RWT) method is more effective than other methods for estimating adult height in boys with delayed bone age, irrespective of whether the delay is ≤ 2 years or > 2 years.

Abstract
Background: Predicted adult height (PAH) can be calculated using methods such as Bayley-Pinneau (BP), Roche-Wainer-Thissen (RWT), and BoneXpert based on bone age (BA) assessment. Since these methods were developed for healthy children, varying results have been reported regarding their efficacy across different patient groups.
Objective: This study aimed to determine the most effective method for PAH by comparing the BP, RWT, and BoneXpert methods in boys with constitutional delay of growth and puberty (CDGP).
Subjects and Methods: Sixty-two male patients with CDGP who had reached their final height (FH) were included in the study. Two experienced clinicians reassessed left-hand and wrist radiographs taken at the time of diagnosis using the Greulich-Pyle (GP) atlas to manually determine BA. Among the methods used for PAH, the GP atlas was used for BP and RWT, while we used its GP-based electronic software for the BoneXpert method.
Results: The mean age at diagnosis of the patients was 14.2 ± 0.8 years, with 58.1% (n=36) having a similar family history. The mean height standard deviation (SD) score was -2.1 ± 0.9, and 24.2% (n=15) of patients received low-dose testosterone induction therapy. The median BAs of the patients were 12.5 (11.5–13.0) years using the GP atlas and 12.6 (11.8–13.4) years with BoneXpert (p<0.001). Boys who were treated with or not treated with testosterone therapy had similar mean height SD scores, median testicular volumes, and median BAs assessed by both methods. The mean target height and FH SD scores were -0.6 ± 0.6 and -0.6 ± 0.9, respectively (p=0.8). Almost all patients (n=60, 97%) achieved adult height within the target range, with no significant difference in the FH SD score between boys who received testosterone and those who did not (p=0.1). There was no significant difference between the FH and PAH when estimated by the BP and RWT methods (p=0.2 and p=0.6, respectively), while the BoneXpert method underestimated the FH (p<0.001). The BP and RWT methods provided better predictions in patients with BA ≤2 years compared to BoneXpert (p=0.3 and p=0.4 vs. p<0.001, respectively). Conversely, RWT and BoneXpert methods were more accurate in PAH in boys with delayed BA >2 years (p=0.1 and p=0.1, respectively), while the BP method resulted in overestimation (p=0.003).
Conclusion: The RWT method was found to be a better predictor of FH compared to the BP or BoneXpert methods in boys with delayed BA ≤ 2 years and >2 years.
Keywords: Adult height estimation; final height; constitutional delay of growth and puberty

Prof. Dr. Ayhan Abacı, Division of Pediatric Endocrinology, Faculty of Medicine, Dokuz Eylul University, İzmir, Türkiye
ayhanabaci@gmail.com
0000-0002-1812-0321

Introduction
Constitutional delay of growth and puberty (CDGP) is classified as the most prevalent cause of short stature and delayed puberty, primarily in boys. This temporary condition is considered a normal growth spectrum, characterized by a slowdown in linear growth, retarded bone maturation, and delayed onset of puberty and the pubertal growth spurt. The exact etiology of CDGP remains unclear, but about 50–75% of patients have a family history of delayed puberty, often following an autosomal dominant inheritance pattern. Typically, these children are expected to reach normal adult height after experiencing delayed but otherwise normal puberty (1–3).
In the management of boys with CDGP, reassurance and watchful waiting are generally adequate. However, for adolescents experiencing psychological challenges and low self-esteem, medical therapy with low-dose testosterone and psychological counseling may be necessary (1–4). Clinicians often use predicted adult height (PAH) based on bone age (BA) assessments to inform patients and their parents about future growth potential. In these patients, PAH can be calculated using methods such as Bayley-Pinneau (BP), Roche-Wainer-Thissen (RWT), and BoneXpert, which were originally developed for healthy children (5–9). Additionally, delayed bone age is a known factor that can limit the accuracy of adult height prediction. So far, only a few studies have evaluated prediction methods in patients with CDGP, and there is limited and conflicting information about their accuracy (5,7–10).
In the present study, we aimed to assess the accuracy of the BP, RWT, and BoneXpert methods in estimating adult height for boys with CDGP to provide them with realistic information about their future height potential.

Subjects and Methods
Patients
In this study, we included boys diagnosed with CDGP who were referred to pediatric endocrinology units for evaluation of short stature or delayed puberty between 2010 and 2018 and who had achieved their final height (FH). To increase the sample size, extended criteria of

delayed puberty were applied (11,12). The inclusion criteria were as follows: (i) boys aged 13 years or older with a testicular volume less than 6 mL, as measured by a Prader orchidometer, who exhibited spontaneous pubertal development before the age of 18 or after pubertal induction with low-dose testosterone; (ii) absence of any endocrine or chronic medical condition; and/or (iii) a familial history of pubertal delay. Male subjects who were born small for gestational age, had received any medication, or had systemic diseases, dysmorphic syndromes, skeletal abnormalities, or pituitary hormone deficiencies were excluded.

Final height was defined based on the following criteria: (i) fused epiphyses; (ii) a growth velocity of less than 1.0 cm in the preceding year; and (iii) completed secondary sexual characteristics. Boys were considered to have achieved their target height if their FH was within the 1.5 standard deviation (SD) score of the target height.

Clinical Assessment

We retrospectively gathered data on age, anthropometric measurements, physical examinations, and parental height, as well as laboratory and radiological findings from patient medical records. Height was measured to the nearest millimeter using a Harpenden stadiometer, and weight was measured with a SECA scale (Hamburg, Germany) to an accuracy of 0.1 kg, with patients wearing only underwear and no shoes. Standard deviation (SD) scores for height, weight, and body mass index (BMI) were calculated using an online calculator (Child Metrics) based on Turkish standards published by Neyzi et al (13). The pubertal stage was assessed using Tanner's standards (14). The genetic target height was calculated by adding 6.5 cm to the average parental height, following the Tanner formula (15). We recorded the levels of luteinizing hormone (LH), follicle-stimulating hormone (FSH), serum total testosterone (ng/mL) obtained at 8:00 am, and the gonadotropin-releasing hormone-stimulated LH levels from the patient files.

Bone Age Assessment and Adult Height Prediction Methods

The BAs of boys with CDGP were reassessed using both manual and automated methods. Initially, BAs were independently re-evaluated by two experienced clinicians (AA and GAK) using left-hand and wrist radiographs, according to the Greulich-Pyle (GP) atlas (16). The manual BA for each patient was subsequently calculated by averaging these independent assessments. Radiographs were stored on a PACS workstation, and all images were uploaded in DICOM format before being analyzed by BoneXpert software. Subsequently, the same radiographs were evaluated using the automated BA assessment method provided by BoneXpert Standalone, based on the GP atlas (Visiana, Holte, Denmark, www.boneXpert.com) (17).

Adult height predictions were made using the BP, RWT, and BoneXpert methods. The GP atlas was utilized for both the BP and RWT methods. For the BoneXpert method, two approaches were employed: (1) PAH was calculated using the BA automatically detected by the BoneXpert software, which is based on the GP atlas, after uploading left-hand x-rays, and (2) PAH was calculated by entering the BA manually assessed using the GP atlas into the BoneXpert's web page. Manually assessing BA using the GP atlas and applying it in the BoneXpert method for PAH calculation is routine practice in our clinics.

The estimated adult height for the BP and RWT methods was calculated using the online calculator (Child Metrics, www.childmetrics.org) (18). For the BP method, this calculator utilizes tables mentioned in the study by Post and Richman (19), which provide decimal fractions indicating the proportion of adult height attained at various BAs. These fractions are categorized based on whether the BA is average (within one year), delayed, or advanced relative to chronological age. PAH is determined by dividing the current height by the corresponding decimal fraction for the patient's BA.

The RWT method calculates PAH based on the child's height (standing height was used instead of recumbent length due to the retrospective nature of data collection in this study), weight, and bone age, incorporating mid-parental height and using sex- and age-specific coefficients (20). While recumbent length is generally recommended for younger children, standing height is more practical and is commonly measured in clinical practice for children older than two years. According to the World Health Organization Child Growth Standards, recumbent length is, on average, 0.7 cm greater than standing height (21). Although this small difference could theoretically affect PAH calculations, given the retrospective design of our study and the minimal variation between standing height and recumbent length, the impact on RWT prediction accuracy is expected to be negligible.

The BoneXpert method, available as a free online calculator at http://www.boneXpert.com/adult-height-predictor, is based on BA, chronological age, gender, height, father's height, mother's height, and ethnicity. For the ethnicity parameter, we selected the Caucasian European South population, as it most closely matches the Turkish population (22,23). In Türkiye, the average height is reported as 163.1 cm for females and 176.2 cm for males, whereas in the Caucasian European South population, the corresponding averages are 162 cm for females and 175 cm for males (13,24).

Ethics

This study was approved by the local Ethics Committee (Approval number: 2024/05-21) and performed in line with the principles of the Declaration of Helsinki. An informed written consent form was not obtained due to the retrospective nature of the study.

Statistical analysis

Statistical analyses were conducted using SPSS for Windows, version 24.0 (IBM Co., Armonk, NY, USA). The normality of the data was assessed with the Kolmogorov-Smirnov and Shapiro-Wilk tests. Clinical data were reported as numbers (%) for categorical variables, mean ± standard deviation for continuous variables with a normal distribution, and medians with the respective 25-75th percentile values for non-normally distributed variables. Comparisons between categorical variables were performed using the Pearson Chi-Square test or Fisher's Exact test, as appropriate. For continuous variables, the Student's t-test was used to compare normally distributed data between the two groups, while the Mann–Whitney U test was applied for non-normally distributed data.

The interclass correlation coefficient (ICC) was used to evaluate interobserver agreement, with ICC values interpreted as follows: excellent (greater than 0.9), good (0.75 to 0.9), moderate (0.5 to 0.75), and poor (less than 0.5). The Wilcoxon signed-rank test was employed to compare two related samples, with a p-value of <0.05 considered statistically significant.

Friedman's two-way analysis of variance was employed to compare differences between predicted heights and FHs. If a significant difference was found, the Bonferroni post hoc test was conducted for pairwise comparisons, with a p-value of <0.0167 considered significant, calculated using the formula p = 0.05 × 2/k(k - 1), where k is the number of comparisons.

Results

Baseline characteristics of study subjects

This study included sixty-two patients, with a mean age at diagnosis of 14.2 ± 0.8 years. Fifty-six subjects (90.3%) were born with normal birth weights, while the remaining subjects were born large for gestational age. Thirty-six patients (58.1%) had a family history of CDGP. The main reasons for presenting to pediatric endocrinology clinics were short stature ($n = 56$, 90.3%) and delayed puberty ($n = 5$, 8.1%). The baseline characteristics of the study subjects are presented in Table 1. The mean SD scores for height and BMI were -2.1 ± 0.9 [(-3.5) - (0.3)] and -0.8 ± 1.5 [(-4.3) - (2.6)], respectively. According to their age and sex, 36 subjects (58.1%) exhibited short stature, and 12 patients (19.4%) had a low BMI in the first evaluation.

Fifteen patients (24.2%) received low-dose testosterone therapy. Patients who underwent testosterone induction therapy were older, taller, and had a higher BMI SD score compared to those who did not ($p = 0.02$, $p = 0.04$, and $p = 0.01$, respectively). However, the treated and untreated groups were similar in terms of height SD score, and testicular volume at presentation ($p = 0.4$ and $p = 0.5$, respectively) (Table 1).

Bone age assessment

The median BAs evaluated using the GP atlas and the BoneXpert method were 12.5 (11.5 - 13.0) and 12.6 (11.8 - 13.4) years, respectively (p <0.001). For the GP atlas of determining BAs, the interobserver coefficient of variation was 0.964 (95% CI 0.941-0.979). The BA retardation

was 2.0 (1.3 – 2.6) years when BA was assessed by the GP atlas and 1.6 (1.0 – 2.3) years when determined by the BoneXpert method (p <0.001). The median BAs of boys, whether they were treated or not treated with testosterone, were found to be similar ($p = 0.09$ for the GP atlas and $p = 0.1$ for the BoneXpert method).

Evaluation at 12 months of follow-up

We had first-year data available for 36 patients. After a year, the mean age of these patients was 15.2 ± 0.9 years. Out of these 36 patients, 13 (36.1%) received low doses of testosterone treatment. The mean height SD score was -2.0 ± 1.1, while the height velocity was 7.3 ± 2.6 cm/year. It was observed that the height velocity was not significantly different in boys who received testosterone treatment and those who did not (8.1 ± 2.3 cm/year vs. 6.8 ± 2.7 cm/year, $p = 0.6$). The median testicular volume was 10 (8–12) mL for all the patients. The median testicular volume in the treated and untreated groups at 12 months of follow-up was also similar ($p = 0.4$).

Characteristics of boys at final height

The median age of the patients at their FH was 19.4 (18.5 to 20.3) years. They had mean target height and FH SD scores of -0.6 ± 0.6 [(-2.1) – (1.0)] and -0.6 ± 0.9 [(-2.5) – (1.4)], respectively ($p = 0.8$). The difference in SD scores between the target height and the FH was -0.04 ± 0.8 [(-2.2) - (2.0)]. Ninety-seven percent of the patients ($n = 60$) reached an adult height within the target height range. Adult height was less than 165 cm in only four boys (6.5 %). Furthermore, there was no significant difference in the FH SD score between boys who received testosterone and those who did not ($p = 0.1$). No significant difference in target height was found between the two groups ($p = 0.5$).

Comparison of adult height prediction methods

The median PAH SD scores calculated using the BP, RWT, and BoneXpert methods were -0.5 [(-1.3) – (-0.1)], -0.6 [(-1.0) – (0.0)], and -1.1 [(-1.6) – (-0.4)], respectively. Among the PAH methods, there was no significant difference between the FH and the PAH estimated by the BP and RWT methods ($p = 0.2$ and $p = 0.6$, respectively) (Table 2). Consequently, the BP and RWT methods provided more accurate predictions for boys with CDGP. Notably, the BoneXpert method underestimated the PAH in these patients (p <0.001) (Figure 1). For the BP, RWT, and BoneXpert methods, the differences between the SD scores of PAH and FH were 0.2 [(-0.4) – (0.7)], 0.03 [(-0.5) – (0.5)] and -0.4 [(-0.9) – (0.1)], respectively (Table 3). The median difference between PAH and FH for the BP and RWT methods was similar ($p = 0.2$). However, the median differences between PAH and FH for the BP and RWT methods were also significantly higher than the BoneXpert method (p <0.001 and p <0.001, respectively).

In the subgroup analysis of boys with delayed bone age ≤ 2 years or >2 years, the BP and RWT methods gave better predictions in patients with delayed bone age ≤ 2 years ($p = 0.3$ and $p = 0.4$, respectively). In this subgroup, the BoneXpert method underestimated the PAH. On the other hand, the RWT and BoneXpert methods were more accurate in predicting PAH in boys with delayed bone age >2 years ($p = 0.1$ and $p = 0.1$, respectively). However, the BP method resulted in overestimation in the same subgroup (Table 4). Consequently, the RWT method is the best predictor of final height among the three different methods in boys with delayed bone age ≤ 2 years and >2 years.

Discussion

In the literature, there is uncertainty about whether boys with CDGP can achieve their target height, regardless of whether receiving low-dose testosterone induction therapy. Additionally, studies evaluating adult height prediction methods for this population often show conflicting results, partly due to limited patient numbers. To address this issue, we collected data from a relatively large group of patients and compared the accuracy of the BP, RWT, and BoneXpert methods. Our findings indicated that the mean FH SD scores of boys with and without pubertal induction were similar, and the majority of patients were able to reach their target heights. Furthermore, this study demonstrated that the RWT method was more effective than other methods for estimating adult height in boys with delayed BA, regardless of whether the delay was ≤ 2 years or > 2 years.

Pharmacological induction of puberty to accelerate the pubertal growth spurt and enhance statural outcomes in boys with CDGP remains contentious. While some researchers suggest that this therapy negatively impacts FH (25), other studies indicate no significant effect of testosterone treatment on FH or PAH in boys with CDGP (8,25–27). For instance, Arrigo et al. (8) found no significant difference in FH between boys with CDGP who received low-dose testosterone therapy and those who did not, with similar height SD scores at diagnosis. Similarly, Kelly et al. (27) observed no significant difference in FH between the treated and untreated groups; however, they noted that the FHs of treated boys were closer to their genetic target heights compared to untreated boys, whose FHs were below their target heights. Consistent with these findings, our study also demonstrated no significant difference in FH SD scores between treated and untreated boys, suggesting that testosterone therapy neither improves nor impairs final height outcomes in boys with CDGP. Notably, even high-dose testosterone therapy has been reported to have no significant effect on the height-for-bone-age SD score, as shown by Büyükgebiz (28), further supporting the conclusion that testosterone therapy does not significantly influence final height outcomes in boys with CDGP.

Several studies demonstrate that patients with CDGP may not reach their genetic target height (7,29,30). Poyrazoğlu et al. reported that the FH of patients with CDGP was below their target height, with 46.3% of patients unable to attain their target height (7). Similarly, in a study of 15 boys, Rohani et al.(29) found that the FHs of the subjects were considerably less than their target heights (165.7 ± 2.89 cm vs. 171.8 ± 4.65 cm). In contrast, Kaickreuth et al. (6) noted that their patients achieved their genetic target height without needing growth-stimulating therapy. In another retrospective study evaluating boys with CDGP, it was reported that the FHs of the patients were similar to their target heights. Additionally, among those who did not undergo testosterone therapy, 3 out of 27 patients did not reach their target height, while only 1 out of 22 treated patients failed to reach their target height (8). In our study, consistent with previous research, we found no significant difference between the FH and the target height. Most of our patients reached an adult height within their expected target range. Only four boys had an adult height below 165 cm, reflecting familial height characteristics and influences.

There are data in the literature comparing adult height prediction methods for boys with CDGP (5–10,27). For instance, a study involving 14 male patients with CDGP found that height predictions using the BP method (173.9 ± 7.5 cm) were highly accurate when compared to the FH (171.3 ± 5.3 cm) (6). In a study by Arrigo et al. (8), no significant differences were found between the final adult height and the PAH calculated using the BP method in both testosterone-induced and non-induced groups. However, they observed discrepancies greater than ± 5 cm between FH and PAH in 33% of non-induced subjects and 23% of induced subjects. Kelly et al. suggested that the Tanner and Whitehouse RUS (TW2) method is useful and accurate. They observed that the FHs of boys with CDGP were closely related to the estimated heights, with only three patients having FHs below the predicted range (27). According to Poyrazoğlu et al. (7), the BP method provided a very reliable estimation of adult height compared to the TW method.

Our study indicated that the RWT method outperformed other methods in estimating adult height for boys with delayed BA, irrespective of whether the delay was ≤ 2 years or > 2 years. Consistent with this, Brämswig et al. concluded that the RWT method was the most accurate, while the BP method overestimated adult height in their cohort of 37 boys with untreated short stature and CDGP (9). Similarly, Reinehr et al. (10) reported that the BP method overestimated adult height, particularly in boys with a delayed BA of 2 years or more. To address this, they developed a new prediction model specifically for patients with CDGP, which they stated had a good predictive capability for subjects with retarded BA.

On the other hand, Unrath et al. (5) found that the BoneXpert method, which incorporates parents' heights, was more accurate in predicting FHs than the BP method in a cohort including boys with CDGP. Their study compared automated BA assessments using the BoneXpert software with manual BA assessments performed by the GP method. When the mean of BAs, blindly re-evaluated by three experienced pediatric endocrinologists, was considered the 'Reference' BA, it was found to be closer to the manual BA than the automated BA. The automated BA slightly overestimated the BA, while the manual BA values were generally lower than the reference BA. Additionally, using BA instead of automated BA in the BoneXpert adult height prediction calculator resulted in a slightly weaker, but still good, performance.

The studies mentioned above have shown varying results for height prediction models in boys with CDGP. In our study, the PAH estimated using both the BP and RWT methods was very closely aligned with and was not significantly different from the FH. In contrast, the BoneXpert method, whether using manual or automated BA assessments, underestimated the PAH in these patients. This underestimation may be attributed to several factors. First, BoneXpert relies on generalized growth models that may not fully account for the dynamic and individualized growth patterns of boys with CDGP, particularly those with delayed bone maturation and pubertal onset (5). Second, as observed in both our study and the findings of Unrath et al. (27), automated methods like BoneXpert tend to slightly overestimate BA compared to manual assessments, leading to discrepancies in adult height predictions. Furthermore, our study determined a significant difference between the median BAs obtained through manual and automated methods, with the automated BA consistently being more advanced. This discrepancy underscores that, despite their standardization and efficiency, automated systems like BoneXpert may have inherent margins of error, raising questions about their reliability. Future studies with larger sample sizes are needed to refine these models and enhance their predictive accuracy and clinical applicability.

Bone age retardation can result in inaccuracies in adult height predictions. Notably, in the aforementioned studies, the commonly used BP method tends to overestimate adult height in boys with CDGP (9,10). In our study, the BP and RWT methods were found to be more accurate for individuals with a delayed BA of less than two years. However, for boys with a BA delayed by more than two years, the BP method tended to overestimate, consistent with findings from previous studies. Conversely, the RWT and BoneXpert methods were found to be more reliable for these patients. This study demonstrated that the RWT method is the most accurate predictor of adult height, regardless of the duration of delayed BA in boys with CDGP. Its incorporation of multiple growth parameters, including height, weight, mid-parental height, and sex- and age-specific coefficients, likely accounts for its superior performance, particularly in patients with complex and variable growth patterns. In summary, these findings highlight the clinical utility of the RWT method as a reliable and precise tool for estimating adult height in boys with CDGP, even in the presence of delayed BA.

Study limitations

The current study has several limitations. First, the inclusion criteria for study participants were somewhat extended to increase the sample size. Additionally, the patients were recruited from different centers, resulting in a heterogeneous population; some patients received testosterone therapy, while others did not. Medical therapy was administered specifically to adolescents experiencing psychological challenges. Furthermore, there is inter- and intra-observer variability in the manual assessment of BA. Nonetheless, all radiographs were re-evaluated by two experienced pediatric endocrinologists, who demonstrated excellent agreement in BA determinations. Finally, the RWT method was originally designed for calculation using recumbent length, but due to the retrospective nature of the study, we used standing height measurements instead. Standing height is approximately 0.7 cm less than the recumbent length in children over two years old, which may affect RWT-based predictions.

Conclusion

In conclusion, our study showed that low-dose testosterone induction therapy did not negatively impact FH, and both treated and untreated boys attained heights in line with their genetic target heights. Furthermore, the RWT method appears to be more suitable for accurate height estimation, especially in conditions such as CDGP, which is characterized by delayed BA. Future research should focus on developing disease-specific prediction models that offer superior advantages over traditional methods for predicting adult height in boys with CDGP.

Declarations

Conflict of interest: The authors have no conflicts of interest to declare.

Ethical statement: This research complies with the guidelines for human studies and is conducted ethically following the World Medical Association Declaration of Helsinki.

Study approval statement: Institutional approval was granted by the Ethics Committee of Dokuz Eylül University Faculty of Medicine (Ethics approval number: 2024/05-21).

Informed consent: An informed written consent form was not obtained due to the retrospective nature of the study.

References

1. Sultan C, Gaspari L, Maimoun L, Kalfa N, Paris F. Disorders of puberty. Best Pract Res Clin Obstet Gynaecol. 2013;48:62-89.
2. Klein DA, Emerick JE, Sylvester JE, Vogt KS. Disorders of Puberty: An Approach to Diagnosis and Management. Am Fam Physician. 2017;96(9):590-599.
3. Gaudino R, De Filippo G, Bozzola E, Gasparri M, Bozzola M, Villani A, Radetti G. Current clinical management of constitutional delay of growth and puberty. Ital J Pediatr. 2022;48(1):45.
4. Raivio T, Miettinen PJ. Constitutional delay of puberty versus congenital hypogonadotropic hypogonadism: Genetics, management and updates. Best Pract Res Clin Endocrinol Metab. 2019;33(3):101316.
5. Unrath M, Thodberg HH, Schweizer R, Ranke MB, Binder G, Martin DD. Automation of bone age reading and a new prediction model improve adult height prediction in children with short stature. Horm Res Paediatr. 2012;78(5–6):312–9.
6. Kalckreuth G V., Haverkamp F, Kessler M, Rosskamp RH. Constitutional delay of growth and puberty: Do they really reach their target height? Horm Res Paediatr. 1991;35(6):222–5.
7. Poyrazoğlu Ş, Günöz H, Darendeliler F, Saka N, Bundak R, Baş F. Constitutional delay of growth and puberty: From presentation to final height. J Pediatr Endocrinol Metab. 2005;18(2):171–9.
8. Arrigo T, Cisternino M, Luca De F, Saggese G, Messina MF, Pasquino AM, De Sanctis V. Final height outcome in both untreated and testosterone-treated boys with constitutional delay of growth and puberty. J Pediatr Endocrinol Metab. 1996;9(5):511-7.
9. Brämsig JH, Fasse M, Holthoff ML, von Lengerke HJ, von Petrykowski W, Schellong G. Adult height in boys and girls with untreated short stature and constitutional delay of growth and puberty: Accuracy of five different methods of height prediction. J Pediatr. 1990;117(6):886–91.
10. Reinehr T, Hoffmann E, Rothermel J, Lehrian TJ, Brämsig J, Binder G. A New Model of Adult Height Prediction Validated in Boys with Constitutional Delay of Growth and Puberty. Horm Res Paediatr. 2019;91(3):186–94.
11. Sedlmeyer IL, Hirschhorn JN, Palmert MR. Pedigree analysis of constitutional delay of growth and maturation: Determination of familial aggregation and inheritance patterns. J Clin Endocrinol Metab. 2002;87(12):5581–6.
12. Harrington J, Palmert MR. Distinguishing constitutional delay of growth and puberty from isolated hypogonadotropic hypogonadism: Critical appraisal of available diagnostic tests. J Clin Endocrinol Metab. 2012;97(9):3056–67.
13. Neyzi O, Bundak R, Gökçay G, Günöz H, Furman A, Darendeliler F, Baş F. Reference Values for Weight, Height, Head Circumference, and Body Mass Index in Turkish Children. J Clin Res Pediatr Endocrinol. 2015;7(4):280-93.

14.   Tanner JM, Whitehouse RH. Clinical longitudinal standards for height, weight, height velocity, weight velocity, and stages of puberty. Arch Dis Child. 1976;51(3):170–9.

15.   Tanner JM. 1 Normal growth and techniques of growth assessment. Clin Endocrinol Metab. 1986;15(3):411–51.

16.   Greulich WW, Pyle SI. Radiographic atlas of skeletal development of the hand and wrist, 2nd edn. University Press, Stanford.1959.

17.   Thodberg HH, Kreiborg S, Juul A, Pedersen KD. The BoneXpert method for automated determination of skeletal maturity. IEEE Trans Med Imaging. 2009;28(1):52–66.

18.   Demir K, Özen S, Konakçı E, Aydın M, Darendeliler F. A comprehensive online calculator for pediatric endocrinologists: ÇEDD Çözüm/TPEDS metrics. J Clin Res Pediatr Endocrinol. 2017;9(2):182–4.

19.   Post EM, Richman RA. A condensed table for predicting adult stature. J Pediatr. 1981;98(3):440-2.

20.   Roche AF, Wainer H, Thissen D. The RWT method for the prediction of adult stature. Pediatrics. 1975;56(6):1027-33.

21.   World Health Organization 2021. Training course on the inpatient management of severe acute malnutrition: Module 2. Principles of care. Web Annex A: Weight-for-height reference card (WHO Child Growth Standards).

22.   Thodberg HH, Jenni OG, Caflisch J, Ranke MB, Martin DD. Prediction of adult height based on automated determination of bone age. J Clin Endocrinol Metab. 2009;94:4868–74.

23.   Thodberg HH, Neuhof J, Ranke MB, Jenni OG, Martin DD.Validation of bone age methods by their ability to predict adult height. Horm Res Paediatr. 2010;74:15–22.

24.   Visiana. BoneXpert Adult Height Predictor - User Manual. Version 3.0. Holte (Denmark): Visiana 2021.

25.   Martin MM, Martin AL, Mossman KL. Testosterone treatment of constitutional delay in growth and development: effect of dose on predicted versus definitive height. Acta Endocrinol Suppl (Copenh). 1986;279:147-52.

26.   Bergadá I, Bergadá C. Long Term Treatment with Low Dose Testosterone in Constitutional Delay of Growth and Puberty: Effect on Bone Age Maturation and Pubertal Progression. J Pediatr Endocrinol Metab. 1995;8(2):117–22.

27.   Kelly BP, Paterson WF, Donaldson MD. Final height outcome and value of height prediction in boys with constitutional delay in growth and adolescence treated with intramuscular testosterone 125 mg per month for 3 months. Clin Endocrinol (Oxf). 2003;58(3):267-72.

28.   Büyükgebiz A. Treatment of constitutional delayed puberty with a combination of testosterone esters. Horm Res. 1995;44 Suppl 3:32-4.

29.   Rohani F, Alai MR, Moradi S, Amirkashani D. Evaluation of near final height in boys with constitutional delay in growth and puberty. Endocr Connect. 2018;7(3):456–9.

30.   Wehkalampi K, Vangonen K, Laine T, Dunkel L. Progressive reduction of relative height in childhood predicts adult stature below target height in boys with constitutional delay of growth and puberty. Horm Res. 2007;68(2):99–104.

Table 1. The clinical and laboratory characteristics of the patients with CDGP at presentation.

Clinical features	All patients (*n*=62)	Patients who received testosterone induction therapy (*n*=15)	Patients who did not receive testosterone induction therapy (*n*=47)	*p*
Chronological age, years	14.2 ± 0.8	14.4 (14.1 – 15.2)	14.0 (13.7 – 14.6)	***0.02[a]***
Bone age by GP atlas, years	12.5 (11.5 – 13.0)	13.0 (11.5 – 13.5)	12.3 (11.5 – 13.0)	*0.09[a]*
Bone age by the BoneXpert method, years	12.6 (11.8 – 13.4)	13.4 (11.9 – 13.7)	12.4 (11.8 – 13.2)	*0.1[a]*
Height, cm	149.2 ± 7.0	152.5 ± 9.5	148.1 ± 5.8	***0.04[b]***
Height, SD score	-2.1 ± 0.9	-1.9 ± 1.2	-2.2 ± 0.7	*0.4[b]*
Body mass index, SD score	-0.8 ± 1.5	0.1 ± 1.7	-1.0 ± 1.4	***0.01[b]***
Target height, cm	172.3 ± 4.2	173.0 ± 4.0	172.1 ± 4.3	*0.5[b]*
Target height, SD score	-0.6 ± 0.6	-0.5 ± 0.6	-0.6 ± 0.7	*0.5[b]*
Age at final height, years	19.4 (18.5 – 20.3)	20.4 (19.5 – 22.0)	19.1 (18.3 – 19.8)	***0.001[a]***
Final height, cm	172.7 ± 5.8	174.6 ± 6.8	172.1 ± 5.3	*0.1[b]*
Final height, SD score	-0.6 ± 0.9	-0.3 ± 1.1	-0.7 ± 0.9	*0.1[b]*
Testicular volume, mL	4 (3 - 6)	4 (3 - 6)	4 (3 - 6)	*0.5[a]*
<4 mL, [n (%)]	34 (54.8 %)	5 (33.3 %)	24 (51.1 %)	*0.3[c]*
4-6 mL, [n (%)]	28 (45.2 %)	10 (66.7 %)	23 (48.9 %)	
Tanner stage (pubic hair)				
Stage 1 [n (%)]	24 (38.7 %)	5 (33.3 %)	19 (40.4 %)	
Stage 2 [n (%)]	37 (59.7 %)	10 (66.7 %)	27 (57.4 %)	
Stage 3 [n (%)]	1 (1.6 %)	0 (0 %)	1 (2.1 %)	*0.8[c]*
Stage 4 [n (%)]	0 (0 %)	0 (0 %)	0 (0 %)	
Stage 5 [n (%)]	0 (0 %)	0 (0 %)	0 (0 %)	
Laboratory				
FSH, mIU/mL	2.3 (1.6 – 2.9) **n=42**	1.9 (1.2 – 2.8) **n=15**	2.5 (1.7 – 3.3) **n=27**	-[d]
LH, mIU/mL	0.8 (0.6 - 1.0) **n=42**	0.7 (0.4 – 0.8) **n=15**	0.8 (0.7 - 1.5) **n=27**	-[d]
Total testosterone, ng/dL	19.4 (14.2 – 25.9) **n=42**	19.4 (15.0 – 31.0) **n=15**	19.2 (12.5 – 24.9) **n=27**	-[d]
Peak LH, mIU/mL	13.0 (9.7 – 21.4) **n=12**	14.8 (9.1 – 21.7) **n=10**	11.8* **n=2**	-[d]

Data are presented as mean ± standard deviation for normal distribution and median ($25^{th} – 75^{th}$ percentile) for those not distributed normally.
[a]Mann-Whitney U test, [b]Student's T-test, [c]Pearson chi-square test; $p<0.05$.
[d]Statistical comparisons could not be performed due to missing data.

For Peak LH, only two patients were included; therefore, only the mean value is presented.
Abbreviations: GP, Greulich-Pyle; SD score, standard deviation score; FSH; follicle-stimulating hormone, LH; luteinizing hormone.

Table 2. Comparison of final heights and predicted heights in boys with CDGP.

| | Final height | PAH (Bone age assessed by the GP atlas) | | | PAH (Bone age assessed by BoneXpert) |
		BP	**RWT**	**BoneXpert**	**BoneXpert**
Height, cm	172.0 (168.0 – 177.0)[a,b]	173.4 (168.5 -177.7)	172.4 (169.5 – 176.2)	170.7 (168.0 – 175.4)[a]	169.4 (166.6 – 174.0)[b]
Height, SD score	-0.7 [(-1.3) – (0.1)][c,d]	-0.5 [(-1.3) – (-0.1)]	-0.6 [(-1.0) – (0.0)]	-0.9 [(-1.3) – (-0.1)][c]	-1.1. [(-1.6) – (-0.4)][d]

Abbreviations: PAH, predicted adult height; GP, Greulich-Pyle; BP, Bayley-Pinneau; RWT, Roche-Wainer-Thissen; SD score, standard deviation score.
Data are given as median (25 - 75p)[a-d] values with the same letter designation were different in the Wilcoxon signed-rank test, at a *p*-value of <0.05.
Height, cm: Final height vs PAH (BoneXpert; bone age assessed by GP atlas), [a]*p*=0.01; Final height vs PAH (BoneXpert; bone age assessed by BoneXpert), [b]*p*<0.001.
Height, SD score: Final height vs PAH (BoneXpert; bone age assessed by GP atlas), [c]*p*=0.01; Final height vs PAH (BoneXpert; bone age assessed by BoneXpert), [d]*p*<0.001.

Table 3. Comparison of the difference between predicted heights and final heights.

| | Bone age assessed by the GP atlas | | | Bone age assessed by BoneXpert | |
	BP	**RWT**	**BoneXpert**	**BoneXpert**	*p**
PAH-FH difference, cm	1.3 [(-2.5) – (4.3)][a,b]	0.2 [(-2.9) – (2.6)][c,d]	-1.6 [(-4.4) – (2.0)][a,c,e]	-2.7 [(-5.7)- (0.6)][b,d,e]	***<0.001***
PAH-FH difference, SD score	0.2 [(-0.4) – (0.7)][f,g]	0.03 [(-0.5) – (0.5)][h,i]	-0.3 [(-0.7) – (0.3)][f,h,j]	-0.4 [(-0.9) – (0.1)][g,i,j]	***<0.001***

Abbreviations: FH, final height; PAH, predicted adult height; GP, Greulich-Pyle; BP, Bayley-Pinneau; RWT, Roche-Wainer-Thissen; SD score, standard deviation score.
Data are given as median (25 - 75p).*Friedman's Two-way analysis of variance.
[a-u] values with the same letter designation were different in the post hoc analysis of pairwise groups, at a p-value of <0.008. [a]*p*= 0.001, [b]*p*<0.001, [c]*p*= 0.001, [d-j]*p*<0.001.

Table 4. Comparison of predicted heights and final heights based on the bone age delay.
Abbreviations: PAH, predicted adult height; GP, Greulich-Pyle; BP, Bayley-Pinneau; RWT, Roche-Wainer-Thissen; SD score, standard

|	Bone age assessed by the GP atlas				Bone age assessed by BoneXpert	
	Final height (n = 31)	**BP (n = 31)**	**RWT (n = 31)**	**BoneXpert (n = 31)**	**Final height (n = 39)**	**BoneXpert (n = 39)**
Delayed bone age ≤ 2 years						
Height, cm	175.0 (168.5 – 175.7)[a]	172.4 (167.4 – 177.5)	172.7 (170.1 – 177.3)	171.0 (168.2 – 175.7)[a]	174.0 (168.5 – 178.5)[c]	169.8 (166.3 – 174.8)[c]
Height, SD score	-0.2 [(-1.3) – (-0.1)][b]	-0.6 [(-1.4) – (0.2)]	-0.6 [(-1.0) – (0.2)]	-0.8 [(-1.3) – (-0.1)][b]	-0.4 [(-1.3) – (0.4)][d]	-1.0 [(-1.6) – (-0.2)][d]
	Final height (n = 31)	**BP (n = 31)**	**RWT (n = 31)**	**BoneXpert (n = 31)**	**Final height (n =23)**	**BoneXpert (n =23)**
Delayed bone age > 2 years						
Height, cm	170.0 (167.6 – 175.0)[e]	173.5 (170.5 – 179.0)[e]	171.6 (169.0 – 175.2)	170.4 (167.5 – 174.2)	170.0 (167.1 – 174.0)	168.7 (166.8 – 173.9)
Height, SD score	-1.0 [(-1.4) – (-0.2)][f]	-0.4 [(-0.9) – (-0.5)][f]	-0.6 [(-1.1) – (-0.2)]	-0.9 [(-1.4) – (-0.3)]	-1.0 [(-1.5) – (-0.4)]	-1.2 [(-1.5) – (-0.4)]

deviation score.
Data are given as median (25 - 75p).
[a-d] values with the same letter designation were different in the Wilcoxon signed-rank test, at a *p*-value of <0.05.
[a]*p*=0.02, [b]*p*=0.02, [c]*p*<0.001, [d]*p*<0.001, [e]*p*=0.003, [f]*p*=0.003.

Figure 1. Comparison of predicted adult heights, target heights, and final heights in boys with CDGP.
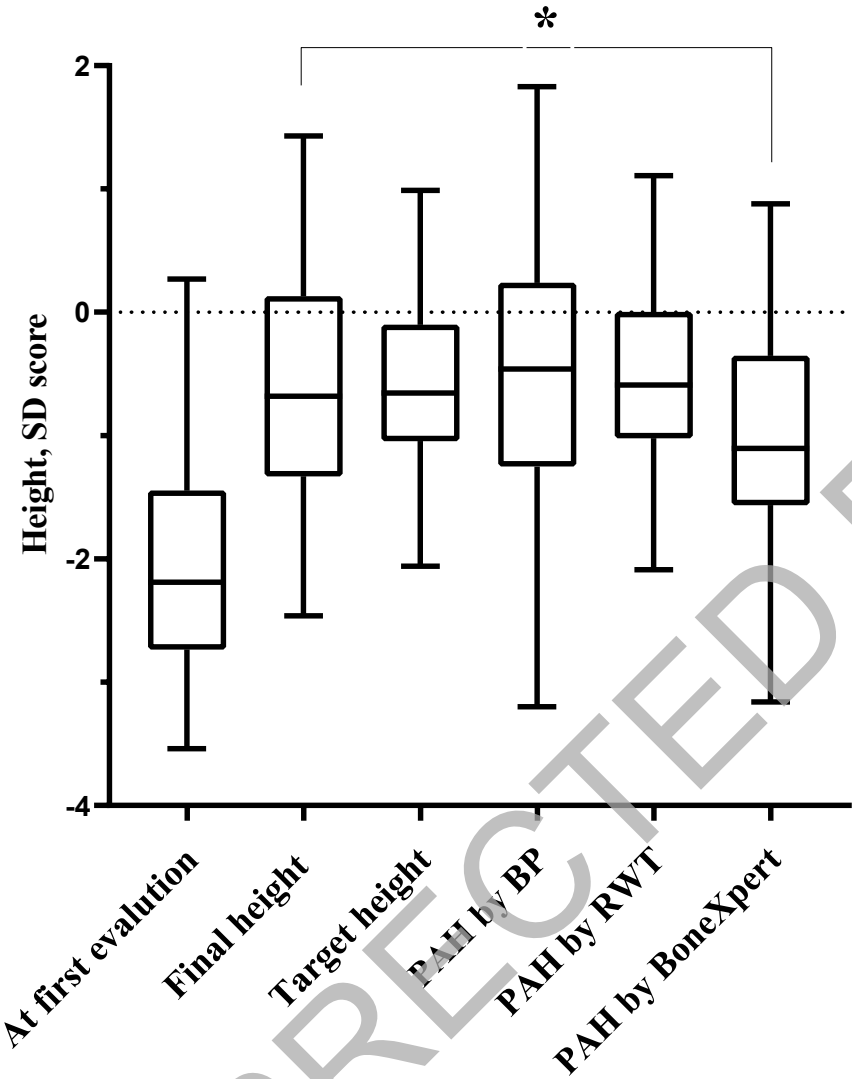
Abbreviations: PAH, predicted adult height; BP, Bayley-Pinneau; RWT, Roche-Wainer-Thissen; SD score, standard deviation score.
*Wilcoxon signed-rank test, $p < 0.05$