AJO-DO

# Evaluating the agreement of skeletal age assessment based on hand-wrist and cervical vertebrae radiography

**Philipp Beit,[a] Timo Peltomäki,[b] Marc Schätzle,[c] Luca Signorelli,[d] and Raphael Patcas[c]**
*Zurich and Wil, Switzerland, and Tampere, Finland*

**Introduction:** The aim of this study was to examine the agreement of skeletal age assessment based on hand-wrist radiographs with cephalogram-based cervical vertebrae evaluation. To circumvent bias and loss of information from staging, a quantitative approach was applied to determine morphologic changes. **Methods:** We analyzed 730 sets of radiographs (cephalogram and hand-wrist) of untreated subjects (352 boys, 378 girls; age range, 6-18 years) from a growth study, each sex as a separate sample. Skeletal age was determined on the hand-wrist radiographs according to the method of Greulich and Pyle. Morphometric changes of the vertebral bodies C2 through C4 were measured (concavity, anterior height, and angle) and tested for correlations with the method of Greulich and Pyle. All correlating variables were included in a multiple linear regression to generate a calculated skeletal age. To establish the agreement between the method of Greulich and Pyle and calculated skeletal age, Bland-Altman plots were made, limits of agreement were identified, and cross-tables (before and after peak height velocity) were computed. Similarly, the agreement between the method of Greulich and Pyle and each subject's chronologic age was estimated for comparison. **Results:** Concavity of C2, C3, and C4; anterior height of C3 and C4; and the angle of C3 correlated with skeletal age highly significantly ($P < 0.0001$) in both sexes, and calculated skeletal age was established based on a linear regression. The agreement between the method of Greulich and Pyle and calculated skeletal age was modest (limits of agreement: boys, ±3.5 years; girls, ±3.3 years) and substantially weaker than the agreement between the method of Greulich and Pyle and chronologic age (limits of agreement: boys, +2.1 to −1.7 years; girls, +2.2 to −1.2 years). Similarly, calculated skeletal age resulted in considerably more false predictions of peak height velocity (boys, 18.9%; girls, 12.9%) than did chronologic age (boys, 7.1%; girls, 7.4%). **Conclusions:** Morphometric assessment of age-dependent changes in the cervical spine offers no advantage over chronologic age, in either assessing skeletal age or predicting the pubertal growth spurt. (Am J Orthod Dentofacial Orthop 2013;144:838-47)

The evaluation of skeletal age is essential in many orthodontic treatment approaches, especially regarding the correction of skeletal imbalance.[1] In functional orthopedics, which aims to exploit mandibular growth, success is intimately linked to growth potential. But growth of the mandible is not linear throughout development.[2,3] Chronologic age has been deemed an inadequate indicator to identify stages of growth because of individual variations in timing, velocity, and duration of growth.

The periods of acceleration and deceleration during growth are based on the complex endocrine regulation of craniofacial growth.[4] Although a novel approach with insulin-like growth factor 1 as an indicator for the pubertal growth spurt has been described in scientific literature, this method has not yet reached clinical applicability.[5] Thus, the assessment of biologic age remains restricted to 2 approaches: clinical evaluation based on various secondary indicators (sexual maturation, longitudinal records of body height,[3,6] menarche in girls, or voice change in boys[7]) and the appraisal of skeletal maturity based on radiographs (most

[a]Research fellow, Department for Orthodontics and Pediatric Dentistry, Center of Dental Medicine, University of Zurich, Zurich, Switzerland.
[b]Head, Oral and Maxillofacial Unit, Department of Ear and Oral Diseases and Department of Otolaryngology, University of Tampere, Tampere, Finland.
[c]Senior lecturer, Department for Orthodontics and Pediatric Dentistry, Center of Dental Medicine, University of Zurich, Zurich, Switzerland.
[d]Private practice, Wil, Switzerland.
All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest, and none were reported.
Address correspondence to: Raphael Patcas, Department for Orthodontics and Pediatric Dentistry, Center of Dental Medicine, University of Zurich, Plattenstrasse 11, 8032 Zurich, Switzerland; e-mail, raphael.patcas@zzm.uzh.ch.
Submitted, May 2013; revised and accepted, July 2013.
0889-5406/$36.00

http://dx.doi.org/10.1016/j.ajodo.2013.07.015

**Table I.** Descriptive analysis: sample size and distribution according to age

| | *Boys* | | | *Girls* | | |
|---|---|---|---|---|---|---|
| *Age group (y)* | *Mean age (y)* | *Age range (y)* | *Subjects (n)* | *Mean age (y)* | *Age range (y)* | *Subjects (n)* |
| 6 | 6.0 | 5.9-6.1 | 13 | 6.0 | 6.0-6.1 | 9 |
| 7 | 7.0 | 7.0-7.2 | 10 | 7.0 | 7.0-7.2 | 7 |
| 8 | 8.0 | 8.0-8.2 | 14 | 8.0 | 8.0-8.0 | 15 |
| 9 | 9.0 | 8.9-9.2 | 35 | 9.0 | 9.0-9.2 | 41 |
| 10 | 10.0 | 10.0-10.1 | 16 | 10.0 | 9.9-10.1 | 23 |
| 11 | 11.0 | 10.9-11.1 | 12 | 11.0 | 10.9-11.3 | 31 |
| 12 | 12.0 | 12.0-12.2 | 49 | 12.0 | 12.0-12.2 | 47 |
| 13 | 13.0 | 13.0-13.1 | 13 | 13.0 | 13.0-13.2 | 35 |
| 14 | 14.0 | 14.0-14.3 | 49 | 14.0 | 13.9-14.3 | 63 |
| 15 | 15.0 | 14.9-15.2 | 66 | 15.0 | 14.9-15.2 | 57 |
| 16 | 16.0 | 15.9-16.2 | 46 | 16.0 | 16.0-16.1 | 32 |
| 17 | 17.0 | 17.0-17.1 | 25 | 17.0 | 17.0-17.2 | 15 |
| 18 | 18.0 | 18.0-18.1 | 4 | 18.0 | 18.0-18.1 | 3 |
| Total | | | 352 | | | 378 |

commonly, hand-wrist radiographs[8-11] or cervical vertebral maturation[6,12-16]).

Skeletal maturation assessed with a hand-wrist radiograph is considered to be the gold standard.[17,18] Its alleged high reliability is due to the fact that a multitude of hand and wrist bones can be depicted on 1 radiograph, and their various stages of ossification allow a precise correlation to overall growth.[19] Several methods were described to assess hand-wrist radiography, and for endocrinologists, the concordant result of different methods is considered most reliable.[20]

Lamparsky[16] demonstrated as early as 1975 that, alternatively, morphologic changes in cervical vertebrae could also be used to evaluate skeletal maturity. Many subsequent authors confirmed a correlation between the maturation of the hand-wrist bones and the cervical vertebrae, as well as a correlation between the cervical vertebrae and facial growth.[10] Based on these findings and since cephalograms are routinely taken in orthodontics, it would indeed appear that a hand-wrist radiographic examination becomes superfluous.

Thus, the recommendation to replace hand-wrist radiographic examination with cervical vertebrae assessment has been widely expressed and has been the subject of various studies. Although the substitution has been vastly supported in previous years, recently more studies have reported poor reproducibility and reliability of the cervical vertebrae maturation assessment.[21-25] This verdict was based mainly on the difficulty of staging and classifying the vertebral bodies of C3 and C4 as trapezoidal, rectangular horizontal, square, or rectangular vertical.[22] The difficulty in staging the cervical vertebrae morphology has rendered its use questionable as a strict clinical guideline for timing of orthodontic treatment.[22,23]

The aim of this study was therefore to circumvent the shortcomings of staging and reexamine the agreement of skeletal age assessment based on hand-wrist radiographs with cephalogram-based cervical vertebrae evaluation. To overcome the drawbacks of staging in common cervical spine assessments, a quantitative approach of morphologic changes was applied. To verify the accuracy of the calculated cervical age, the results were compared with the agreement between skeletal age assessments based on hand-wrist radiographs and chronologic age.

## MATERIAL AND METHODS

The material of this cross-sectional study consisted of radiographic records (lateral cephalograms and hand-wrist radiographs) from the Zurich Craniofacial Growth Study performed from 1981 to 1984 at the Department of Orthodontics and Pediatric Dentistry of the University of Zurich in Switzerland. In the original study, healthy schoolchildren of white ethnicity from local public schools with no history of orthodontic treatment were randomly selected for radiographic records (cephalogram and hand-wrist radiograph). All radiographic records (n = 1372) of this study were taken close to each subject's birthday (range, 6-18 years; Table I).

Legal approval for releasing the data was obtained by the Federal Commission of Experts for Professional Secrecy in Medical Research.[26] All cephalograms included in this study had to be of sufficient quality, and the cervical spine (C2-C4) had to be clearly visible. Thus, the final sample consisted of 730 subjects (378 girls, 352 boys).

The lateral cephalograms were taken with the head stabilized by ear rods and nasal support. The Frankfort

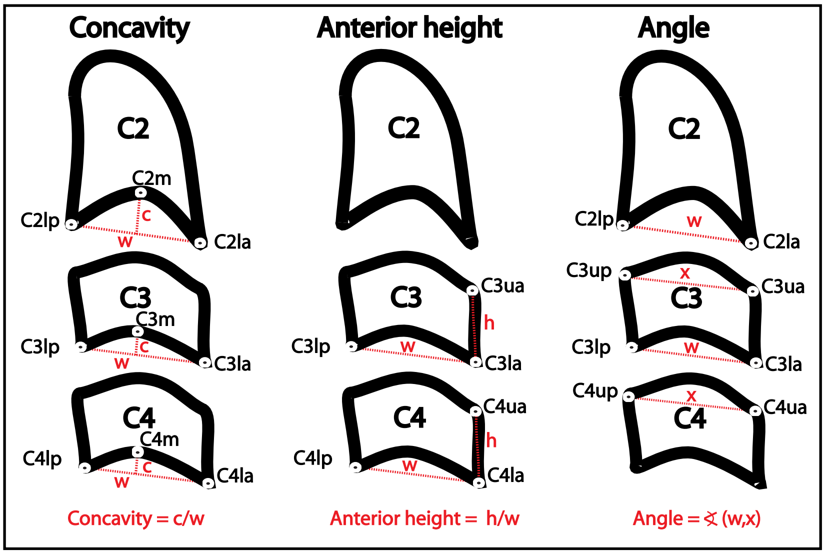


**Fig 1.** Tracings of cervical vertebrae C2, C3, and C4. The corner points are landmarked as well as the most cranial point of the vertebrae's caudal border (*C2m*, *C3m*, *C4m*). Abbreviations: *u*, upper; *l*, lower; *a*, anterior; *p*, posterior; *m*, most cranial point. To obtain the measurements used (concavity, anterior height, and angle), ratios and angles of the following linear measurements were taken: *w*, width, the distance between *lp* and *la*; *c*, concavity, perpendicular line to *w* through *m*; *h*, height, distance between *ua* and *la*; *x*, line between *up* and *ua*.

**Table II.** Ratios and angles used in the cephalometric analysis

| *Variable* | *Applied on* | *Description* |
|---|---|---|
| Concavity | C2, C3, C4 | The ratio between the lower border (C2/3/4lp – C2/3/4la) of the vertebra and its perpendicular distance through C2/3/4m |
| Anterior height | C3, C4 | The ratio between the anterior border (C3/4ua – C3/4la) and the lower border (C3/4lp – C3/4la) |
| Angle | C3, C4 | The angle between the upper (C3/4up - C3/4ua) and lower (C3/4lp – C3/4lp) borders of the vertebrae |

See Figure 1 for abbreviations.

horizontal plane was set parallel to the floor, and the teeth were in centric occlusion. The radiographs were taken with a focus-to-coronal plane distance of 200 cm and an enlargement of 7.5%. Hand-wrist radiographs were taken in a standardized fashion with a focus-to-film distance of 215 cm and a 30° angulation of the thumb to allow depiction of the sesamoid bone.

The 730 lateral cephalograms were traced and landmarked by 3 dedicated investigators who shared the workload equally (these landmarks are defined in Table II and shown in Fig 1). The digitizing was performed with a tablet digitizer, AccuGrid (Numonics, Landsdale, Pa), with a resolution of 1 mili-inch. Custom-made software was used for the calculation of the cephalometric values.

Ratios of the linear measurements or angles were used to account for radiographic enlargement and interindividual size differences. All analyzed variables are shown in Figure 1 and correspond to the morphologic changes described in the literature.[13] The analysis included the following parameters: concavity of C2, C3, and C4; anterior height of C3 and C4; and the angle between the upper and lower borders of C3 and C4.

To appraise the reliability of the cervical spine evaluation, it was examined against the skeletal age assigned by the method of Greulich and Pyle,[8] the gold standard chosen for this study.

One of the above-mentioned investigators who traced the cephalograms repeated the tracings of 38 cephalograms, more than 6 months later; 19 of this investigator's tracings and an additional 19 tracings of another investigator were repeated to determine the intraobserver and interobserver reproducibilities. Since only 1 investigator (P.B.) assessed all hand-wrist radiographs, the reliability tests were confined to intraobserver repeatability. Thirty hand-wrist radiographs were analyzed a second time, 12 months later, by the same investigator, to determine intraobserver reliability.

**Table III.** Pearson correlation of morphometric measurements of the cervical vertebrae (*C2*, *C3*, and *C4*) and skeletal age: correlation coefficients and *P* values for each sex

| | *Concavity* | | | *Anterior height* | | *Angle* | |
|---|---|---|---|---|---|---|---|
| | *C2* | *C3* | *C4* | *C3* | *C4* | *C3* | *C4* |
| Boys | | | | | | | |
| Correlation coefficient | 0.649 | 0.676 | 0.675 | 0.762 | 0.706 | 0.235 | 0.087 |
| *P* value | 0.000* | 0.000* | 0.000* | 0.000* | 0.000* | 0.000* | 0.102 |
| Girls | | | | | | | |
| Correlation coefficient | 0.587 | 0.690 | 0.717 | 0.746 | 0.750 | 0.343 | 0.365 |
| *P* value | 0.000* | 0.000* | 0.000* | 0.000* | 0.000* | 0.000* | 0.102 |

*Significance level at $P < 0.01$.

## STATISTICAL ANALYSIS

A standard statistical software package (PASW Statistics version 18; IBM, Armonk, New York, NY) was used for the statistical analysis. The intraclass correlation coefficient (ICC) for absolute agreement based on a 1-way random-effects analysis of variance (ANOVA) was calculated to determine intraobserver and interobserver reliability for every assessment method separately. Descriptive statistics for the measurements were computed. The Pearson correlation coefficient was performed to evaluate correlations between the morphologic measurements and the skeletal age by hand-wrist assessment. Correlating variables were included in a multiple linear regression to generate a calculated skeletal age. To establish the agreement between the Greulich and Pyle[8] evaluation and the calculated skeletal age, Bland-Altman plots[27] were performed, and limits of agreement were identified. Cross-tables (before and after peak height velocity) with sensitivity, specificity, positive and negative predictive values, and likelihood ratios were computed. Similarly, the agreement between the Greulich and Pyle evaluation and each subject's chronologic age was estimated for comparison. *P* values smaller than 0.05 were considered statistically significant.

## RESULTS

The ICC values showed excellent reproducibility for all cephalometric measurements, with 0.948 (95% confidence interval [CI], 0.931-0.966) for intraobserver repeatability and 0.933 (95% CI, 0.915-0.952) for interobserver repeatability. The ICC of 0.992 (95% CI, 0.985-0.996) for intraobserver repeatability of hand-wrist radiographs assessment was similarly high.

Possible correlations between morphologic changes of the vertebrae and skeletal age, as determined by the method of Greulich and Pyle,[8] are shown in Table III. Excellent correlations were found for concavity of C2, C3, and C4 as well as for anterior height of C3 and C4. In contrast, although statistically highly significant, angle C3 had only a low correlation coefficient ($r = 0.235$), and angle C4 did not correlate at all. Variables without significant correlations and with a correlation coefficient lower than 0.5 were excluded from the multiple regression equation. As a result, angles C3 and C4 were not included in the regression analysis.

A multiple regression equation as a function of the variables of concavity C2, C3, and C4 as well as anterior height C3 and C4 was calculated to estimate the skeletal age, as shown in Table IV. Based on multiple linear regression analyses, Equations 1 and 2 were established for the calculated skeletal age (CSA).

The multiple linear regressions are based on highly significant relationships, as shown in Table IV. Moreover, the adjusted coefficients of determination (boys, adjusted $R^2 = 0.693$; girls, adjusted $R^2 = 0.671$) demonstrate that the regression lines fit the data fairly well and indicate that future outcomes are likely to be predictable by the models.

$$\text{CSA boys} = 4.559 + 9.897 \times \text{concavity C2} + 6.866 \times \text{concavity C3} + 10.066 \times \text{concavity C4} + 6.193 \times \text{anterior height C3} + 2.844 \times \text{anterior height C4} \quad \text{(Equation 1)}$$

$$\text{CSA girls} = 5.242 + 5.758 \times \text{concavity C2} + 6.629 \times \text{concavity C3} + 9.176 \times \text{concavity C4} + 2.953 \times \text{anterior height C3} + 4.306 \times \text{anterior height C4} \quad \text{(Equation 2)}$$

**Table IV.** Multiple linear regression analysis of the correlating measurements

| | Independent variable | Concavity | | | Anterior height | |
|---|---|---|---|---|---|---|
| | | *C2* | *C3* | *C4* | *C3* | *C4* |
| Boys | | | | | | |
| Regression coefficient | 4.559 | 9.897 | 6.866 | 10.066 | 6.193 | 2.844 |
| SE | 0.508 | 2.148 | 2.410 | 2.417 | 1.022 | 1.164 |
| *t* value | 8.977 | 4.608 | 2.849 | 4.164 | 6.062 | 2.444 |
| Significance | 0.000† | 0.000† | 0.005† | 0.000† | 0.000† | 0.015* |
| Girls | | | | | | |
| Regression coefficient | 5.242 | 5.758 | 6.629 | 9.176 | 2.953 | 4.306 |
| SE | 0.481 | 1.714 | 2.102 | 2.071 | 0.971 | 1.067 |
| *t* value | 10.891 | 3.358 | 3.154 | 4.430 | 3.040 | 4.034 |
| Significance | 0.000† | 0.001† | 0.002† | 0.000† | 0.003† | 0.001† |

*Significance level at $P<0.05$; †significance level at $P<0.01$.

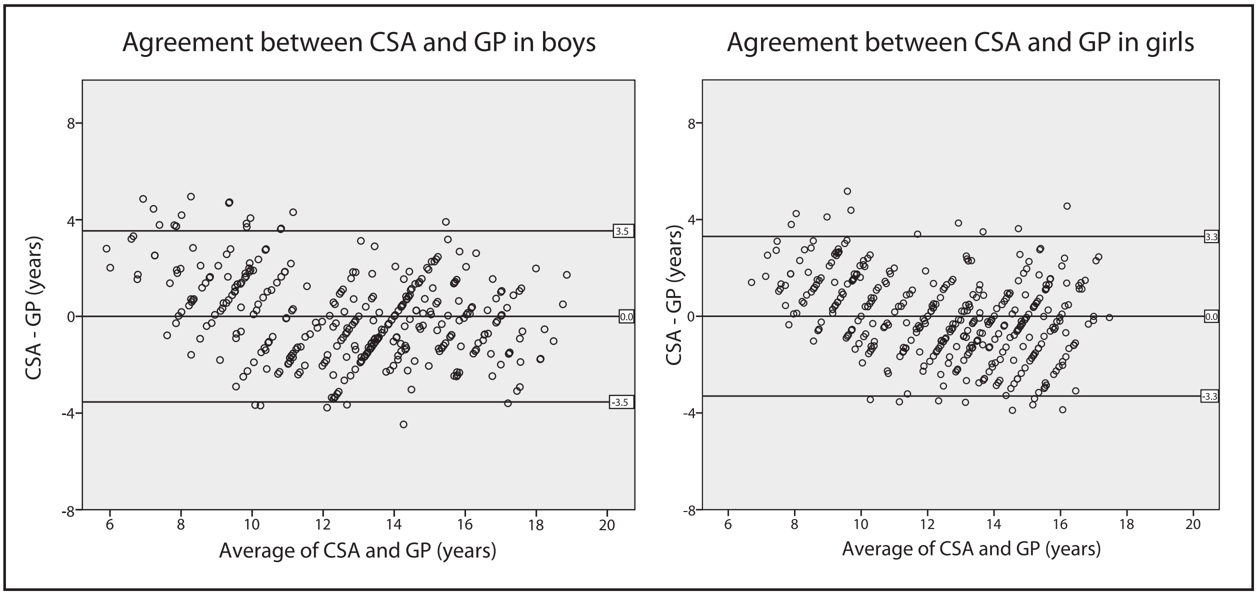


**Fig 2.** Bland-Altman plots to evaluate the agreement between the calculated skeletal age (*CSA*), based on the regression analysis of the cervical vertebrae, and the age evaluated with the method of Greulich and Pyle[8] (*GP*) for boys (*left*) and girls (*right*): *middle line*, mean; *outer lines*, limits of agreement.

To appraise the reliability, calculated skeletal age was yet again examined against the gold standard. The agreement between the calculated skeletal age and the skeletal age based on the Greulich and Pyle[8] assessment was investigated with Bland-Altman plots[27] as shown in Figure 2. The mean of the differences was 0 years for both boys and girls, and the limits of agreement (1.96 × the standard deviation) were ±3.5 years for boys and ±3.3 years for girls. To put these values into context, additional Bland-Altman plots were performed to evaluate the agreement between chronologic age and the skeletal age obtained with the Greulich and Pyle assessment, as shown in Figure 3. The limits of agreement were considerably smaller (boys, +2.1 to −1.7 years; girls, +2.2 to −1.2 years) than in Figure 2. A discrete bias could be discerned in the difference of means, accounting for 0.2 years in boys and 0.5 years in girls.

Cross-tables were computed to compare the agreement between calculated skeletal age, skeletal age based on the method of Greulich and Pyle,[8] and chronologic age. All variables were transformed into binary outcomes before and after peak height velocity, which was defined according to the literature at skeletal ages of 14 years in boys and 12 years in girls,[19,28] and the skeletal age according to Greulich and Pyle was used as a

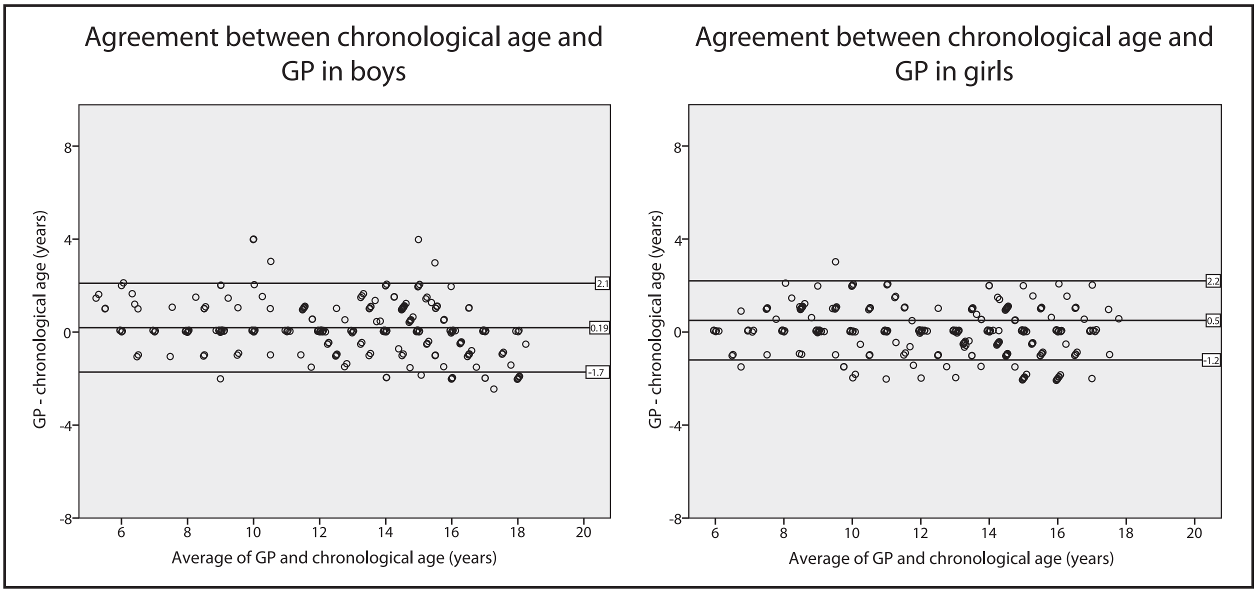


**Fig 3.** Bland-Altman plots to evaluate the agreement between chronologic age and the age evaluated with the method of Greulich and Pyle[8] (*GP*) for boys (*left*) and girls (*right*): *middle line*, mean; *outer lines*, limits of agreement. The unusual distribution is because all records were taken close to each subject's chronologic birthday.

**Table V.** Cross-table for binary splitting of the cases (before vs after peak height velocity), with agreement of calculated skeletal age (*CSA*) of the Greulich and Pyle[8] hand-wrist assessment

| *Skeletal age according to Greulich and Pyle (y)* | *CSA (y)* <14 | *CSA (y)* >14 | *Total* |
|---|---|---|---|
| Boys | | | |
| <14 | 178 | 14 | 192 |
| >14 | 53 | 107 | 160 |
| Total | 231 | 121 | 352 |
| Girls | | | |
| <12 | 140 | 23 | 163 |
| >12 | 26 | 189 | 215 |
| Total | 166 | 212 | 378 |

reference. The agreement between calculated skeletal age and skeletal age as defined by Greulich and Pyle is given in Table V. As a comparison, the agreement between chronologic age and skeletal age as defined by Greulich and Pyle is given in Table VI.

To evaluate the cross-tables statistically, the chronologic age and the calculated skeletal age were considered tests, and "peak height velocity not occurred" was defined as a positive finding. Reliance on the calculated skeletal age led in boys to 15.0% (n = 53) false-positive and 3.9% (n = 14) false-negative results. Conversely, reliance on chronologic age led to only 1.7% (n = 6) false-positive and 5.4% (n = 19) false-negative results.

**Table VI.** Cross-table for binary splitting of the cases (before vs after peak height velocity), with agreement of chronologic age with the Greulich and Pyle[8] hand-wrist assessment

| *Skeletal age according to Greulich and Pyle (y)* | *Chronologic age (y)* <14 | *Chronologic age (y)* >14 | *Total* |
|---|---|---|---|
| Boys | | | |
| <14 | 173 | 19 | 192 |
| >14 | 6 | 154 | 160 |
| Total | 179 | 173 | 352 |
| Girls | | | |
| <12 | 138 | 25 | 163 |
| >12 | 3 | 212 | 215 |
| Total | 141 | 237 | 378 |

The findings were similar for girls. Calculated skeletal age yielded 6.9% (n = 26) false-positive and 6% (n = 23) false-negative results; for chronologic age, there were 0.8% (n = 3) false-positive and 6.6% (n = 25) false-negative results.

Moreover, sensitivity, specificity, predictive values, and likelihood ratios were computed and are shown in Table VII. High positive predictive values to identify "peak height velocity not occurred" in accordance with the method of Greulich and Pyle[8] were confirmed for chronologic age (>96%), but not for calculated skeletal age (boys, 77.0%; girls, 84.3%). Negative predictive

**Table VII.** Sensitivity, specificity, positive (*PPV*) and negative (*NPV*) predictive values, and likelihood ratios (*LR+* and *LR−*) for Tables V and VI

| | *Sensitivity (%)* | *Specificity (%)* | *PPV (%)* | *NPV (%)* | *LR+* | *LR−* |
|---|---|---|---|---|---|---|
| Calculated skeletal age | | | | | | |
| Boys | 92.71 | 66.88 | 77.06 | 88.43 | 2.80 | 0.11 |
| Girls | 85.89 | 87.91 | 84.33 | 89.15 | 7.10 | 0.16 |
| Chronologic age | | | | | | |
| Boys | 90.10 | 96.25 | 96.65 | 89.02 | 24.03 | 0.10 |
| Girls | 84.66 | 98.60 | 97.87 | 89.45 | 60.67 | 0.16 |

Calculated skeletal age and chronologic age were termed as tests, and "peak height velocity not occurred" was defined as a positive test outcome.

values to identify "peak height velocity occurred" in accordance with the method of Greulich and Pyle were about 89% for both chronologic age and calculated skeletal age. Similarly, the likelihood ratio positive differed substantially between both methods (>24.0 vs <7.1), but not the likelihood ratio negative.

## DISCUSSION

Many radiologic methods to assess skeletal age have been described and evaluated. Of all methods, hand-wrist radiography is considered to be the gold standard for orthopedic purposes.[17,18] Nevertheless, the assessment of skeletal age based on vertebral morphology as depicted on a lateral cephalogram is also established in the clinical setting. Many authors and even some orthodontic societies,[29,30] as well as the European Union guidelines in dental radiology,[31] have recommended refraining from hand-wrist radiography because of the additional radiation exposure.

However, more recently, skeletal age assessment based on the cervical spine has been critically reviewed, and its standing in both reduction of radiation exposure[20] and clinical reliability[21-25] has been seriously questioned.

In this study, we aimed to enhance the reliability of cervical spine assessment by circumventing the problems associated with staging. Good to excellent correlations between calculated skeletal age and skeletal age assessed with the method of Greulich and Pyle[8] were found in girls and boys alike, and the adjusted coefficients of determination suggested that the regression models fit the data reasonably well.

Yet, the agreement was weak. The limits of agreement were significantly wider when skeletal age, as assessed by Greulich and Pyle,[8] was compared with calculated skeletal age than when it was compared with chronologic age. Calculated skeletal age performed equally poorly when agreement was evaluated for binary splitting (before vs after peak height velocity). Based on the assumption that the Greulich and Pyle assessment renders the truth, the calculated skeletal age generated more false allocations of subjects than would have been the case had chronologic age been used.

The results of this investigation clearly imply that excellent correlations and solid regression models might still be insufficient to consolidate a robust agreement. The reason is simple: morphologic changes in the cervical spine are age-dependent, and it is not surprising that these changes correlate with everything else that alters with age.

Therefore, the test of significance will prove that the 2 methods are related, but it would be remarkable if the 2 methods designed to measure age-dependent changes were not related.[27] Regrettably, most previous studies evaluating the agreement of the cervical vertebral assessment with any other skeletal age evaluation restricted their statistical approach to correlations and regressions. Our study raises concern about the proper interpretation of these previous studies because it demonstrates that a test of significance can be misleading and does not reflect the agreement.

The correct approach to identify whether the 2 methods of measurement are interchangeable is to evaluate by how much one method is likely to differ from the other.[27,32] We evaluated the agreement in a twofold manner: the data were assessed as continuous values with the Bland-Altman method[27] and were later converted into categorical data to obtain predictive values. Both analyses proved that the agreement between calculated skeletal age and skeletal age by the method of Greulich and Pyle[8] was lower than the agreement between chronologic age and the method of Greulich and Pyle.[8]

Our findings are supported by a recent study. Using geometric morphometric evaluation to determine the relationship of spine anatomy to skeletal maturation, Chatzigianni and Halazonetis[33] noted that the shape of the vertebrae alone could not predict skeletal maturation. Although their approach differed by using the common cervical vertebral method (and not calculated skeletal age), their results showed that the assessment of the cervical spine is not more informative than simple chronologic age.

The disagreement between the Greulich and Pyle[8] appraisal and the chronologic age was not negligible. This incongruity could either be based on the polymorphism in ossification sequencing of the hand-wrist bones[34] or be the result of the study design. Yet, similar to our study, Suri et al[35] evaluated the agreement between chronologic age and skeletal age assessment based on the method of Greulich and Pyle with Bland-Altman plots.[27] Their results concur with the established limits of agreement in our study.

Based on the above, it seems that chronologic age might be more indicative than skeletal age assessment of the cervical spine. This requires an explanation.

Expanding on this evidence, the following rationale might be given. The measurements used correlated highly significantly to skeletal age and made the use of staging obsolete. Contrary to the previous belief that the paucity in accuracy is due to staging, our findings show that age evaluation based on cervical spine morphology is unreliable, even without staging. The most probable explanation is that although vertebral morphology changes during growth, there is an inherent problem because it simply does not contain enough information for accurate age estimation.

The approach of our study—to evaluate the morphologic changes as continuous data without staging—has recently been attempted in a few other studies.[36-39] However, these studies suffer from certain inadequacies. First, the samples were too small, or various samples were mixed.[37,38] Second, it is methodologically questionable to use absolute measurements instead of ratios or angles.[39] Third, all used unsuitable statistical evaluations such as correlation analyses.[36-39]

In orthopedic planning, knowledge of whether peak height velocity has occurred is often more important than an exact skeletal age. Our cross-tables confirm the weakness of the cervical vertebral assessment. In general, calculated skeletal age was less reliable than chronologic age. Of special clinical relevance, however, is the amount of false-negative results, where the peak height velocity is still anticipated although it has already been passed. Using chronologic age results in 0.8% to 1.7% false-negative predictions, whereas applying calculated skeletal age results in 6.9% to 15.0% false-negative predictions. To elucidate this point, we termed the different age assessments as tests to identify whether the subject was still before peak height velocity. As a highly specific assessment method (>96%), chronologic age will rarely miss peak height velocity. Use of chronologic age yielded a higher precision rate (ie, higher positive predictive values) to identify subjects before peak height velocity than did the use of calculated skeletal age, and the likelihood ratios for chronologic age, being more distant from the value of 1 than calculated skeletal age, similarly indicate more conclusive results.

This finding is highly relevant for clinicians because it implies that just knowing the chronologic age will suffice to rarely miss the peak height velocity. Nonetheless, the evaluation of skeletal age in conjunction with the timing of the growth spurt is still important in avoiding premature treatment.

A rationale should be given as to why peak height velocity was defined as occurring at skeletal ages of 14 years in boys and 12 years in girls. There is a strong relationship between skeletal maturity and somatic growth, allowing skeletal age to be used to assess development progress.[40] Several studies demonstrate that peak height velocity is linked to skeletal age.[19,28,41] A recent study showed that the skeletal ages of 14 years in boys and 12 years in girls represent peak height velocity best in the population evaluated.[19] However, peak height velocity should not be determined based on chronologic age because it is influenced by many factors such as secular trend, genetic factors (mediated by environmental factors), and weight (nutritional status).[42,43]

For this study, the values obtained with the Greulich and Pyle[8] assessment were used as the ground truth. Various hand-wrist assessments exist, and there are contradictory reports as to which method is superior.[44-47] Although many studies weigh intraobserver and interobserver reproducibilities of different appraisals against each other, no study has been performed to assess the reliability of any method to assign the correct endocrine maturational stage. The decision to use the Greulich and Pyle assessment for this study was based on its ascertained high reproducibility, sensitivity, and accuracy in assessing peak height velocity because of the appearance of the sesamoid bone as a clearly identifiable landmark.[45-47]

## CONCLUSIONS

1. Calculated skeletal age based on cervical vertebral morphology does not contain enough information for accurate age estimation (inherent problem).
2. Correlation analyses and regression models are insufficient to prove that cervical vertebral morphology assessment can replace other established evaluations.
3. The cervical spine evaluation offers no advantage over chronologic age in either assessing skeletal age or predicting the pubertal growth spurt.
4. Based on the examined population, peak height velocity will be missed in only a few patients if chronologic age is used.

## ACKNOWLEDGMENTS

We thank Nikolaos Pandis for his guidance in statistical analysis, Urs Eiholzer for sharing his expertise in pediatric endocrinology, and Goldie Maxwell for her help in editing the manuscript.

## REFERENCES


1. Baccetti T, Franchi L, Kim LH. Effect of timing on the outcomes of 1-phase nonextraction therapy of Class II malocclusion. Am J Orthod Dentofacial Orthop 2009;136:501-9.
2. Fishman LS. Chronological versus skeletal age, an evaluation of craniofacial growth. Angle Orthod 1979;49:181-9.
3. Hägg U, Taranger J. Maturation indicators and the pubertal growth spurt. Am J Orthod 1982;82:299-309.
4. Pirinen S. Endocrine regulation of craniofacial growth. Acta Odontol Scand 1995;53:179-85.
5. Masoud M, Masoud I, Kent RL Jr, Gowharji N, Cohen LE. Assessing skeletal maturity by using blood spot insulin-like growth factor I (IGF-I) testing. Am J Orthod Dentofacial Orthop 2008;134:209-16.
6. Franchi L, Baccetti T, McNamara JA. Mandibular growth as related to cervical vertebral maturation and body height. Am J Orthod Dentofacial Orthop 2000;118:335-40.
7. Hägg U, Taranger J. Menarche and voice change as indicators of the pubertal growth spurt. Acta Odontol Scand 1980;38:179-86.
8. Greulich WW, Pyle SI. Radiographic atlas of skeletal development of the hand and wrist. 2nd ed. Redwood City, Calif: Stanford University Press; 1959.
9. Fishman LS. Radiographic evaluation of skeletal maturation. Angle Orthod 1982;52:88-112.
10. Flores-Mir C, Nebbe B, Major PW. Use of skeletal maturation based on hand-wrist radiographic analysis as a predictor of facial growth: a systematic review. Angle Orthod 2004;74:118-24.
11. Tanner JM, Whitehouse RH, Cameron N, Marshall WA, Healy MJR, Goldstein H. Assessment of skeletal maturity and prediction of adult height (TW2 method). 2nd ed. London, United Kingdom: Academic Press; 1975.
12. Hassel B, Farman AG. Skeletal maturation evaluation using cervical vertebrae. Am J Orthod Dentofacial Orthop 1995;107:58-66.
13. Baccetti T, Franchi L, McNamara J. An improved version of the cervical vertebral maturation (CVM) method for the assessment of mandibular growth. Angle Orthod 2002;72:316-23.
14. Roman PS, Palma JC, Oteo MD, Nevado E. Skeletal maturation determined by cervical vertebrae development. Eur J Orthod 2002;24:303-11.
15. Baccetti T, Franchi L, McNamara JA Jr. The cervical vertebral maturation (CVM) method for the assessment of optimal treatment timing in dentofacial orthopedics. Semin Orthod 2005;11:119-29.
16. Lamparski D. Skeletal age assessment utilizing cervical vertebrae. Am J Orthod 1975;67:458-9.
17. Hägg U, Taranger J. Skeletal stages of the hand and wrist as indicators of the pubertal growth spurt. Acta Odontol Scand 1980;38:187-200.
18. Gilli G. The assessment of skeletal maturation. Horm Res 1996;45(Suppl 2):49-52.
19. Molinari L, Gasser T, Largo R. A comparison of skeletal maturity and growth. Ann Hum Biol 2013;40:333-40.
20. Patcas R, Signorelli L, Peltomaki T, Schatzle M. Is the use of the cervical vertebrae maturation method justified to determine skeletal age? A comparison of radiation dose of two strategies for skeletal age estimation. Eur J Orthod 2013;35:604-9.
21. Santiago RC, de Miranda Costa LF, Vitral RW, Fraga MR, Bolognese AM, Maia LC. Cervical vertebral maturation as a biologic indicator of skeletal maturity. Angle Orthod 2012;82:1123-31.
22. Nestman TS, Marshall SD, Qian F, Holton N, Franciscus RG, Southard TE. Cervical vertebrae maturation method morphologic criteria: poor reproducibility. Am J Orthod Dentofacial Orthop 2011;140:182-8.
23. Gabriel DB, Southard KA, Qian F, Marshall SD, Franciscus RG, Southard TE. Cervical vertebrae maturation method: poor reproducibility. Am J Orthod Dentofacial Orthop 2009;136:478.e1-7: discussion, 478-80.
24. Zhao XG, Lin J, Jiang JH, Wang Q, Ng SH. Validity and reliability of a method for assessment of cervical vertebral maturation. Angle Orthod 2012;82:229-34.
25. Fudalej P, Bollen AM. Effectiveness of the cervical vertebral maturation method to predict postpeak circumpubertal growth of craniofacial structures. Am J Orthod Dentofacial Orthop 2010;137:59-65.
26. Federal Authorities of the Swiss Confederation. Expertenkommission für das Berufsgeheimnis in der medizinischen Forschung. Sonderbewilligung zur Offenbarung des Berufsgeheimnisses zu Forschungszwecken im Bereich der Medizin und des Gesundheitswesens. Bundesblatt 2011;7640-2.
27. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-10.
28. Hewitt D, Acheson RM. Some aspects of skeletal development through adolescence. II. The inter-relationship between skeletal maturation and growth at puberty. Am J Phys Anthropol 1961;19:333-44.
29. Turpin DL. British Orthodontic Society revises guidelines for clinical radiography. Am J Orthod Dentofacial Orthop 2008;134:597-8.
30. Isaacson KG, Thom AR, Horner K, Whaites E. Orthodontic radiographs—guidelines for the use of radiographs in clinical orthodontics. 3rd ed. London: United Kingdom: British Orthodontic Society; 2008.
31. European Union. European guidelines on radiation protection in dental radiology. Luxembourg, Luxembourg: Office for Official Publications of the European Communities; 2004.
32. Ludbrook J. Statistics in biomedical laboratory and clinical science: applications, issues and pitfalls. Med Princ Pract 2008;17:1-13.
33. Chatzigianni A, Halazonetis DJ. Geometric morphometric evaluation of cervical vertebrae shape and its relationship to skeletal maturation. Am J Orthod Dentofacial Orthop 2009;136:481.e1-9:discussion, 481-3.
34. Garn SM, Rohmann CG, Blumenthal T. Ossification sequence polymorphism and sexual dimorphism in skeletal development. Am J Phys Anthropol 1966;24:101-15.
35. Suri S, Prasad C, Tompson B, Lou W. Longitudinal comparison of skeletal age determined by the Greulich and Pyle method and chronologic age in normally growing children, and clinical interpretations for orthodontics. Am J Orthod Dentofacial Orthop 2013;143:50-60.
36. Mito T, Sato K, Mitani H. Cervical vertebral bone age in girls. Am J Orthod Dentofacial Orthop 2002;122:380-5.
37. Chen L, Liu J, Xu T, Long X, Lin J. Quantitative skeletal evaluation based on cervical vertebral maturation: a longitudinal study of adolescents with normal occlusion. Int J Oral Maxillofac Surg 2010;39:653-9.
38. Caldas Mde P, Ambrosano GM, Haiter Neto F. Computer-assisted analysis of cervical vertebral bone age using cephalometric radiographs in Brazilian subjects. Braz Oral Res 2010;24:120-6.
39. Varshosaz M, Ehsani S, Nouri M, Tavakoli MA. Bone age estimation by cervical vertebral dimensions in lateral cephalometry. Prog Orthod 2012;13:126-31.
40. Nahhas RW, Sherwood RJ, Chumlea WC, Towne B, Duren DL. Predicting the timing of maturational spurts in skeletal age. Am J Phys Anthropol 2013;150:68-75.

41. Hewitt D, Acheson RM. Some aspects of skeletal development through adolescence. I. Variations in the rate and pattern of skeletal maturation at puberty. Am J Phys Anthropol 1961;19:321-31.
42. Karlberg J. Secular trends in pubertal development. Horm Res 2002;57(Suppl 2):19-30.
43. Delemarre-van de Waal HA. Secular trend of timing of puberty. Endocr Dev 2005;8:1-14.
44. Bull RK, Edwards PD, Kemp PM, Fry S, Hughes IA. Bone age assessment: a large scale comparison of the Greulich and Pyle, and Tanner and Whitehouse (TW2) methods. Arch Dis Child 1999;81:172-3.
45. Schmidt S, Nitz I, Ribbecke S, Schulz R, Pfeiffer H, Schmeling A. Skeletal age determination of the hand: a comparison of methods. Int J Legal Med 2013;127:691-8.
46. King DG, Steventon DM, O'Sullivan MP, Cook AM, Hornsby VP, Jefferson IG, et al. Reproducibility of bone ages when performed by radiology registrars: an audit of Tanner and Whitehouse II versus Greulich and Pyle methods. Br J Radiol 1994; 67:848-51.
47. Cole AJ, Webb L, Cole TJ. Bone age estimation: a comparison of methods. Br J Radiol 1988;61:683-6.