Original Article / Özgün Araştırma

The Effect of Clinical Experience on Cervical Vertebral Maturation Assessment: Is a Hand-wrist Radiograph Necessary?

Klinik Deneyimin Servikal Vertebral Maturasyonun Değerlendirilmesine Etkisi: El-bilek Radyografileri Gerekli Midir?

Serpil Çokakoğlu, Merve Çakır, Alp İlgenli

Pamukkale University Faculty of Dentistry, Department of Orthodontics, Denizli, Turkey

Keywords
Hand-wrist, orthodontic experience, vertebral maturation

Anahtar Kelimeler
El-bilek, ortodontik deneyim, vertebral maturasyon

Received/Geliş Tarihi : 14.01.2021
Accepted/Kabul Tarihi : 08.04.2021

doi:10.4274/meandros.galenos.2021.94103

Address for Correspondence/Yazışma Adresi:
Serpil Çokakoğlu Asst. Prof.,
Pamukkale University Faculty of Dentistry, Department of Orthodontics, Denizli, Turkey
Phone : +90 258 296 44 85
E-mail : serpilcokakoglu@gmail.com
ORCID ID: orcid.org/0000-0002-1240-6951

Abstract

Objective: To evaluate the role of hand-wrist radiographs (HWRs) in the assessment of growth and development based on the cervical vertebral maturation (CVM) method and to determine whether clinical experience levels would affect the decision-making process.

Materials and Methods: The lateral cephalometric radiographs and HWRs of 48 patients aged 11 to 16 years were assessed in this study. The clinicians first determined the growth stage by the CVM method based on lateral cephalometric images (T1). Then, the clinicians filled out the same questionnaire with HWRs on each question (T2). Subsequently, the clinicians were divided into groups according to their experience levels. Intraclass correlation coefficients were calculated to evaluate intraobserver agreements. The data were analyzed with chi-square or Fisher-Freeman-Halton Exact tests for intergroup comparisons.

Results: For all evaluators, the percentage of agreement changed minimally after T2. The most experienced clinicians demonstrated the most consistent decisions for overall CVM stages at both evaluations. Concerning each stage, the percentages of changed decisions showed no significant differences between the groups except for stages CS4 and CS5. There were significant differences between the rates of agreement for CS3 in each group from T1 to T2 ($p<0.05$).

Conclusion: The level of orthodontic experience affects the decision-making process. The presence of HWRs provided no advantage in the assessment of CVM.

Öz

Amaç: Çalışmanın amacı el-bilek radyografilerinin servikal vertebral maturasyon (SVM) yöntemine dayalı büyüme ve gelişimin değerlendirilmesindeki rolünü değerlendirmek ve klinik deneyim düzeyinin karar verme sürecine etkisi olup olmadığını belirlemektir.

Gereç ve Yöntemler: Çalışmamızda 11 ila 16 yaşları arasındaki 48 hastanın lateral sefalometrik ve el-bilek radyografileri değerlendirildi. Klinisyenler ilk olarak lateral sefalometrik görüntülere (T1) dayalı olarak SVM yöntemi ile büyüme evresini belirlediler. Daha sonra aynı anketi, lateral sefalometrik görüntülerin her birini el-bilek radyografileri ile değerlendirerek doldurdular (T2). Anketin sonunda klinisyenler deneyim düzeylerine göre gruplara ayrıldı. Gözlemci içi güvenilirliği değerlendirmek için sınıf içi korelasyon katsayıları hesaplandı. Verilerin gruplar arası karşılaştırmaları için ki-kare ve Fisher-Freeman-Halton testleri kullanıldı.

Bulgular: Tüm değerlendiriciler için uyum yüzdesi T2'den sonra minimum düzeyde değişti. En deneyimli klinisyenler, her iki değerlendirmede de toplamda SVM

aşamaları için en tutarlı kararları gösterdiler. CS4 ve CS5 haricindeki diğer aşamalar için değişen kararların yüzdeleri gruplar arasında önemli bir farklılık göstermedi. T1'den T2'ye her grupta CS3 için uyum oranları arasında önemli farklılıklar gözlendi ($p<0,05$).
Sonuç: Ortodontik deneyim düzeyinin karar verme süreci üzerinde etkisi vardır. El-bilek radyografilerin varlığı SVM'nin değerlendirilmesinde avantaj sağlamamıştır.

Introduction

In orthodontic practice, the determination and prediction of growth and development (G&D) is crucial for many treatment protocols. Various methods exist for assessing the maturation stage, such as chronological age, height and weight increases, sexual characteristics, dental development, and skeletal age (1). Clinicians generally prefer the hand-wrist radiograph (HWR) and cervical vertebral maturation (CVM) methods, which provide more accurate data for the assessment of skeletal maturation (2,3).

In a systematic review and meta-analysis, the CVM method has been demonstrated to be an alternative to HWR for predicting the pubertal growth acceleration period (4). The CVM method has also been stated as reliable and reproducible (5,6). In many studies, a significant correlation has been found between the maturations of cervical vertebrae and hand-wrist bones (1,4,7-9). On the other hand, many researchers have demonstrated that the CVM method is affected by serious methodological issues (8,10,11). Predko-Engel et al. (10) reported that the reliability of the CVM method is questionable. Additionally, Perinetti et al. (11) stated that the CVM stages did not exactly reflect mandibular growth. For this reason, the researchers recommended accounting for additional biological indicators of whether orthodontic treatment should be started during the maximum growth period (10-12). Considering this evidence, the determination of G&D by HWR can be accepted as the gold standard (9).

The quality of images, training, experience, and various assessments have been shown to be among the reasons for the poor reliability of the CVM method (13-15). Perinetti et al. (13) concluded that the visual assessment can be accurate and repeatable after training. Rongo et al. (14) reported that clinicians with lower experience levels showed the best results for the reproducibility of the CVM method. Khajah et al. (15) concluded that the highest agreement in CVM staging was obtained on 2D-digital lateral cephalograms with training.

However, no study to date has examined whether the determination of G&D with the CVM method would change when an HWR is included and whether the outcome is affected by the experience level. Therefore, the first aim of this study was to investigate whether HWR is necessary. The second aim was to examine whether assessments are affected by clinical experience. The null hypothesis tested was that there is no significant effect of either HWRs or clinical experience on the decision-making process for determining CVM stages.

Materials and Methods

This study was approved by the Ethics Committee of Pamukkale University Faculty of Medicim (decision number: 12, date: 02.07.2019). The study population was composed of clinicians registered to the Turkish Orthodontic Society. The power analysis (G* Power Ver.3.1.9.2, Kiel, Germany) showed that 34 individuals provided a power of 82% at an alpha error probability of 0.05 and a 0.435 effect size based on a previous study (14).

The subjects participated in this study via e-mailed questionnaires. This was a two-phase questionnaire study. For the questionnaires, the lateral cephalograms and HWRs of 48 patients (24 female and 24 male) were randomly selected according to the following criteria: chronological age (range: 11-16 years), no malformations in any vertebrae or hand-wrist bones, and high-quality radiographs. The informed consent was obtained from the patients for the use of their records. During the selection of cephalograms, those without clear and visible vertebrae were excluded from the study. All lateral cephalogram images were identified in terms of gender and cropped to include only the C2-C4 vertebrae to eliminate any additional information.

During the first evaluation, the participants were asked to determine the skeletal maturation with the CVM method on 50 lateral cephalometric images and with the addition of HWRs in the second evaluation. Before assessments, the participants were informed

about the CVM method introduced by Baccetti et al. (3). The second questionnaire link was deactivated after two months. The participants were divided into three groups according to their experience level following the completion of the questionnaires.

- Group 1: Clinicians with experience of two years or fewer.
- Group 2: Clinicians with experience of two to four years.
- Group 3: Clinicians with experience of five years or more.

An objective analysis of CVM stages was determined by an independent observer (B.K.A) according to the hard copy of the schematic representation of the CVM method and the two samples of each stage in Baccetti's original article (3).

Statistical Analysis

The data were statistically analyzed using SPSS version 23.0 (IBM Corp., Armonk, NY, USA). The Shapiro-Wilk test was used to test normality. Chi-square or Fisher-Freeman-Halton Exact tests were used for comparison between independent variables. The intraclass correlation coefficient (ICC) was used to assess intraobserver reliability. All tests were performed with a significance level of $p<0.05$.

Results

The first and second questionnaires were fully answered by 60 and 55 clinicians, respectively. A total of 34 clinicians completed both questionnaires. The demographic characteristics are shown in Table 1. The clinicians with experience of two years or fewer were significantly younger than the others.

There were 1632 assessments from all participants. The records of two patients were used to determine intraobserver reliability. ICC values ranged from 0.80 to 0.86, which demonstrated good intraobserver reliability (Table 2).

Irrespective of experience level, there were 741 (45%) agreements and 891 (55%) disagreements compared with the reference standard at T1 and 791 (48%) agreements and 841 (52%) disagreements at T2 (Figure 1). The percentage of CVM stage agreement increased from 42% to 45% in Group 1, from 44.5% to 48.5% in Group 2, and from 49.5% to 51.6% in Group 3 (Figure 2).

The distributions of the agreements regarding CVM stages for the different groups are shown in Table 3. There were no significant differences among the groups in terms of CS1, CS2, CS3, and CS6 agreement at T1 and T2. However, there were significant differences for CS4 among groups at T1 ($p=0.001$). Additionally, CS5 agreements showed significant differences among groups at T1 and T2 ($p<0.05$).

For CS4, group 2's percentage of agreement (36.4%) was significantly lower than those of group 1 (51.2%) and group 3 (59.8%) at T1 ($p=0.001$). For CS5, the group 1's percentage of agreement (27.3%) was significantly lower than that of group 3 (48.8%) at T1 ($p=0.019$). Group 1's percentage of agreement (26.0%) was also significantly lower than groups 2 (54.5%) and 3 (48.8%) at T2 ($p=0.001$). For total percentage of agreement, group 1 (42.0%) was significantly lower than group 3 (49.5%) at T1 ($p=0.04$).

Intragroup evaluation showed that there were significant differences for CS3 in each group and for CS4 and CS6 agreements in group 3 from T1 to T2 ($p<0.05$). For CS3, the percentage of agreements increased significantly, from 28.0% to 44.8% in group 1, 37.8% to 50.3% in group 2, and 31.4% to 46.7% in group 3, from T1 to T2. The agreements in group 3 for CS4 decreased significantly from T1 to T2 (59.8% to

Table 2. ICC values for intraobserver reliability

Groups	ICC	Range
Group 1	0.866	0.628-0.961
Group 2	0.813	0.481-0.945
Group 3	0.801	0.475-0.938
Total	0.858	0.750-0.924

ICC: Intraclass correlation coefficients

Table 1. The demographic characteristics

Groups	n	Female (%)	Male (%)	p*	Age (mean ± SD)	p**
Group 1	11	7 (63.6)	4 (36.4)	0.225	27.45±3.503[a]	0.001*
Group 2	11	3 (27.3)	8 (72.7)		33.36±4.843[b]	
Group 3	12	6 (50.0)	6 (50.0)		36.75±6.369[b]	

*Chi-square test, **One Way ANOVA. [a,b]Same letters showed no significant differences between groups

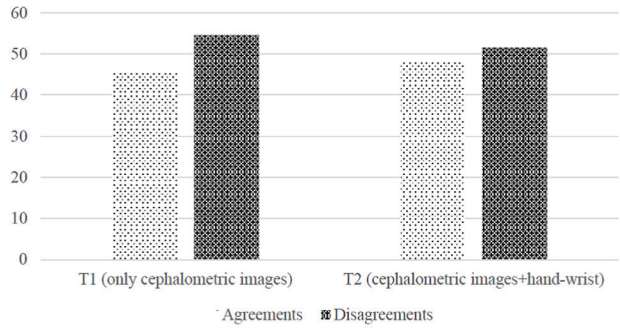

Figure 1. Frequency of observers' total agreement and disagreement at two time points

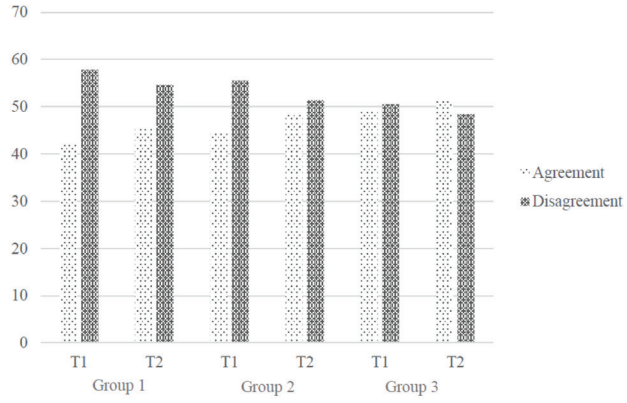

Figure 2. Frequency of groups' agreement and disagreement at two time points

40.2%) while increasing (44.4% to 80.6%) for CS6.

Discussion

Over the years, the CVM method has become the most widely used G&D assessment among clinicians, although HWR is accepted as the gold standard (9,10). This study evaluated the influence of the hand-wrist method on the decision-making process for determining CVM stages. In the first questionnaire (T1), clinicians evaluated the CVM stage without HWR. In the second evaluation (T2), the same lateral cephalometric images were presented, but with the addition of HWRs, for assessing the CVM stage.

The participants evaluated the skeletal maturation on cropped images that contained only the C2-C4 vertebrae in order to eliminate any additional information, such as dentition stage, that could cause method bias among observers. Moreover, the lateral cephalometric images were presented in a random order in the questionnaire to avoid bias.

Our findings showed that the reliability of observers' individual answers demonstrated good agreement. Irrespective of experience levels, about 45% agreement was found for overall stages in the absence of HWRs. However, the percentage of disagreement among clinicians was still almost 50% after the second evaluation. Because training has a major effect on clinicians' accuracy in determining the curvature and shape of vertebrae (15). Therefore, it was inevitable that we would obtain a lower overall level of agreement from the untrained participants, unlike previous studies (16,17).

The clinicians with experience of more than five years had the most consistent decisions (49.5%) at T1 for overall stages. During the first evaluation, they also presented all vertebral stages, except for CS3 and CS6, with the most consistent decisions. These stages were more correctly scored by clinicians with two to four years of experience. In addition, the agreement level for CS2 and CS5 at T1 increased with experience. However, the distributions of total agreements were changed minimally in the presence of HWRs.

When experience level was taken into account, the overall level of agreement among the least experienced clinicians was significantly lower than that of the most experienced clinicians at T1. A reasonable explanation was that experience level might have an impact on the decision process. This finding raised the question of whether different stages would yield different results. However, no pronounced agreement differences were found among the groups in assessing stages 1, 2, 3 and 6 at T1 or T2. Therefore, in both evaluations, clinical experience had no major influence on CVM determination for all stages, except for stages 4 and 5.

For the T1 evaluation, the clinicians with two to four years of experience showed significantly fewer agreements in terms of CS4 compared with the others. For CS5, the least experienced clinicians showed a significantly lower percentage of agreement than the most experienced clinicians. Nevertheless, the addition of HWR in T2 had no positive effect on the level of agreement. Unlike CS4, the clinicians with two to four years' experience showed an increased percentage of agreements, and no significant differences were determined between the moderately and most experienced clinicians' decisions. At this point of the study, it was difficult to explain the differences between CS4 and CS5 in

Table 3. Distribution and comparison of groups' agreement and disagrement according to different cervical vertebra stages

		Group 1		Group 2		Group 3		p (x^2)
		Agree (%)	Disagree (%)	Agree (%)	Disagree (%)	Agree (%)	Disagree (%)	
CS1	T1	69.7	30.3	66.7	33.3	77.8	22.2	0.569
	T2	57.6	42.4	45.5	54.5	69.4	30.6	0.131
	p	0.306		0.083		0.422		
CS2	T1	51.2	48.8	52.9	47.1	54.5	45.5	0.871
	T2	57.0	43.0	51.2	48.8	57.6	42.4	0.541
	p	0.367		0.797		0.620		
CS3	T1	28.0	72.0	37.8	62.2	31.4	68.6	0.199
	T2	44.8	55.2	50.3	49.7	46.7	53.3	0.631
	p	0.003*		0.032*		0.005*		
CS4	T1	51.2[a]	48.8	36.4[b]	63.6	59.8[a]	40.2	0.001*
	T2	40.5	59.5	38.0	62.0	40.2	59.8	0.912
	p	0.094		0.790		0.001*		
CS5	T1	27.3[a]	72.7	40.3[a,b]	59.7	48.8[b]	51.2	0.019*
	T2	26.0[a]	74.0	54.5[b]	45.5	48.8[b]	51.2	0.001*
	p	0.855		0.076		1.000		
CS6	T1	42.4	57.6	60.6	39.4	44.4	55.6	0.266
	T2	57.6	42.4	57.6	42.4	80.6	19.4	0.065
	p	0.218		0.802		0.002*		
Total	T1	42.0[a]	58.0	44.5[a,b]	55.5	49.5[b]	50.5	0.04*
	T2	45.5	54.5	48.5	51.5	51.6	48.4	0.128
	p	0.264		0.195		0.479		

*p<0.05. [a,b]The same letters show no significant differences between groups at T1 and T2

the moderately experienced group. Perinetti et al. (13) found that disagreement was seen mostly in stages 4 and 5 in terms of diagnostic accuracy and repeatability among raters with different experience levels. As demonstrated, the evaluations should be carefully made to avoid unreliable decision-making processes, and other developmental indicators should be evaluated when the visual assessment of the CVM staging is uncertain, especially for stages 4 and 5. In line with this finding, Perinetti et al. (11) concluded that the CVM method may be helpful only when a lateral head film is indicated for other reasons and in combination with other indicators to increase diagnostic reliability.

According to our results, the most disagreements were found among the least experienced clinicians for CS5. The clinicians with between two and four years of experience had the highest disagreement frequency for CS4 at both T1 and T2. On the other hand, the most experienced clinicians showed the highest disagreement rate for CS3 at T1. For these participants, determining CS3 was the most difficult of all the evaluations. In such cases, Cunha et al. (6) recommended adding the HWRS as a way to mitigate any doubt and to help confirm the treatment timing relative to the pubertal growth spurt. Supporting this view, significant improvements in the level of agreement were found in each group when assessing the CVM with HWRs in terms of CS3. Although the percentages of agreement increased in each group, the overall agreement was still about 50%. This meant that clinicians had difficulty agreeing with the determination of the maximum growth period even with the addition of the hand-wrist method. As stated by Gabriel et al. (16) the CVM method can be used only to support the observations of

clinicians in making clinical decisions. On the other hand, significant disagreements appeared in terms of CS4, which can be explained by the evaluation of maturation by the most experienced clinicians with the help of cephalometric film, and confusion may have arisen due to their rarely using hand-wrist films. The difficulty in classifying the vertebral bodies of C3 and C4 as trapezoidal, rectangular horizontal, square, or rectangular vertical may lead to its inadequacy as a strict clinical guideline for the timing of orthodontic treatment (18). In another study, Khanum et al. (17) stated that more careful assessment of CVM stages 5 and 6 should be done in order to avoid unreliable diagnoses. Although McNamara and Franchi (19) reported that the most difficult stage to determine was CS6, the most experienced clinicians showed significantly increased agreement when assessing the CVM in the presence of hand-wrist films during this study. This finding revealed that these participants correctly measured the length of the posterior and inferior borders of the cervical vertebrae when using visual assessment. Based on the findings of this study, the null hypothesis was partially rejected. Additionally, it is obvious that further studies with larger sample sizes are necessary to gain a better understanding of the effects of HWRs on clinicians' decision-making processes.

Conclusion

The HWRs had minimal effects on the determination of the CVM stage. The orthodontic experience of clinicians had an effect on the decision-making process.

Acknowledgement

We thank to Burak Kerem Apaydın, Assoc. Prof., for his support during the objective analysis of maturation stages.

Ethics

Ethics Committee Approval: This study was approved by the Ethics Committee of Pamukkale University Faculty of Medicine (decision number: 12, date: 02.07.2019).

Informed Consent: The informed consent was obtained from the patients for the use of their records.

Peer-review: Externally peerreviewed.

Authorship Contributions

Consept: S.Ç., M.Ç., A.İ., Design: S.Ç., M.Ç., A.İ., Data Collection or Processing: S.Ç., M.Ç., A.İ., Analysis or Interpretation: S.Ç., M.Ç., A.İ., Literature Search: S.Ç., M.Ç., A.İ., Writing: S.Ç., M.Ç., A.İ.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Hassel B, Farman AG. Skeletal maturation evaluation using cervical vertebrae. Am J Orthod Dentofacial Orthop 1995; 107: 58-66.
2. 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.
3. Baccetti T, Franchi L, McNamara JA Jr. The cervical vertebral maturation method for the assessment of optimal treatment timing in dentofacial orthopedics. Semin Orthod 2005; 11: 119-29.
4. Cericato GO, Bittencourt MAV, Paranhos LR. Validity of the assessment method of skeletal maturation by cervical vertebrae: a systematic review and meta-analysis. Dentomaxillofac Radiol 2015; 44: 20140270.
5. Rainey BJ, Burnside G, Harrison JE. Reliability of cervical vertebral maturation staging. Am J Orthod Dentofacial Orthop 2016; 150: 98-104.
6. Cunha AC, Cevidanes LH, Sant'Anna EF, Guedes FR, Luiz RR, McNamara JA, et al. Staging hand-wrist and cervical vertebrae images: a comparison of reproducibility. Dentomaxillofac Radiol 2018; 47: 20170301.
7. Uysal T, Ramoglu SI, Basciftci FA, Sari Z. Chronologic age and skeletal maturation of the cervical vertebrae and hand-wrist: is there a relationship? Am J Orthod Dentofacial Orthop 2006; 130: 622-8.
8. Santiago RC, de Miranda Costa LF, Vitral RWF, Fraga MR, Bolognese AM, Maia LC. Cervical vertebral maturation as a biologic indicator of skeletal maturity. Angle Orthod 2012; 82: 1123-31.
9. Szemraj A, Wojtaszek-Słomińska A, Racka-Pilszak B. Is the cervical vertebral maturation (CVM) method effective enough to replace the hand-wrist maturation (HWM) method in determining skeletal maturation? A systematic review. Eur J Radiol 2018; 102: 125-8.
10. Predko-Engel A, Kaminek M, Langova K, Kowalski P, Fudalej PS. Reliability of the cervical vertebrae maturation (CVM) method. Bratisl Lek Listy 2015; 116: 222-6.
11. Perinetti G, Primozic J, Sharma B, Cioffi I, Contardo L. Cervical vertebral maturation method and mandibular growth peak: a longitudinal study of diagnostic reliability. Eur J Orthod 2018; 40: 666-72.

12. 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.
13. Perinetti G, Caprioglio A, Contardo L. Visual assessment of the cervical vertebral maturation stages: A study of diagnostic accuracy and repeatability. Angle Orthod 2014; 84: 951-6.
14. Rongo R, Valleta R, Bucci R, Bonetti GA, Michelotti A, D'Antò V. Does clinical experience affect the reproducibility of cervical vertebrae maturation method? Angle Orthod 2015; 85: 841-7.
15. Khajah A, Tadinada A, Allareddy V, Kuo CL, Nanda R, Uribe F. Influence of type of radiograph and levels of experience and training on reproducibility of the cervical vertebral maturation method. Am J Orthod Dentofacial Orthop 2020; 157: 228-39.
16. 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: 478-80.
17. Khanum A, Kalia A, Nene S, Mirdeghan N, Joshi J, Mithani R. Visual assessment of cervical vertebral maturation stages: A study of diagnostic accuracy and repeatability. J Oral Dent Health Res 2019; 1: 1-5.
18. 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.
19. McNamara JA Jr, Franchi L. The cervical vertebral maturation method: A user's guide. Angle Orthod 2018; 88: 133-43.