Cervical vertebral maturation method and mandibular growth peak: a longitudinal study of diagnostic reliability

Giuseppe Perinetti[1], Jasmina Primozic[2], Bhavna Sharma[3], Iacopo Cioffi[3] and Luca Contardo[4]

[1]In Private Practice, Nocciano, Italy, [2]Department of Dental and Jaw Orthopaedics, Medical Faculty, University of Ljubljana, Slovenia, [3]Faculty of Dentistry, University of Toronto, Canada and [4]Department of Medical, Surgical and Health Sciences, School of Dentistry, University of Trieste, Italy

Correspondence to: Giuseppe Perinetti, Via San Lorenzo 69/1, 65010 Nocciano (PE), Italy. E-mail: G.Perinetti@yahoo.com

Summary

Background/objectives: The capability of the cervical vertebral maturation (CVM) method in the identification of the mandibular growth peak on an individual basis remains undetermined. The diagnostic reliability of the six-stage CVM method in the identification of the mandibular growth peak was thus investigated.

Subjects/Methods: From the files of the Oregon and Burlington Growth Studies (data obtained between early 1950s and middle 1970s), 50 subjects (26 females, 24 males) with at least seven annual lateral cephalograms taken from 9 to 16 years were identified. Cervical vertebral maturation was assessed according to the CVM code staging system, and mandibular growth was defined as annual increments in Co–Gn distance. A diagnostic reliability analysis was carried out to establish the capability of the circumpubertal CVM stages 2, 3, and 4 in the identification of the imminent mandibular growth peak.

Results: Variable durations of each of the CVM stages 2, 3, and 4 were seen. The overall diagnostic accuracy values for the CVM stages 2, 3, and 4 were 0.70, 0.76, and 0.77, respectively. These low values appeared to be due to false positive cases.

Limitations: Secular trends in conjunction with the use of a discrete staging system. In most of the Burlington Growth Study sample, the lateral head film at age 15 was missing.

Conclusions/Implications: None of the CVM stages 2, 3, and 4 reached a satisfactorily diagnostic reliability in the identification of imminent mandibular growth peak.

Introduction

Nowadays, the search for a reliable indicator of the skeletal maturity is still of crucial importance to enhance the efficiency of functional treatment, especially when dealing with skeletal Class II malocclusions (1). Among the indicators of skeletal maturity, the well-known cervical vertebral maturation (CVM) method (in different variants) has been proposed (2–4). This method was originally reported over four decades ago (5) and has become popular both in clinical practice and research. However, only very few investigations (6–12) correlated the CVM method with mandibular growth according to a longitudinal study design. Some of these studies (6–8) reported the mandibular growth peak to occur during CVM stages 3 and 4, while others (3, 9–12) failed to detect a significant correlation between the maturation of the cervical vertebrae and mandibular growth peak. However, as correlation between parameters does not necessarily imply diagnostic accuracy, the reliability of the CVM method in the identification of the mandibular growth peak on an individual basis remains undetermined (1). To date only one longitudinal study (13) has reported on the diagnostic reliability, i.e. sensitivity or specificity, of the CVM stages 3 and 4 in the identification of mandibular

growth peak. However, that investigation was limited as the results were derived from a data set of the University of Michigan Growth Study from which the CVM method itself was derived (7).

By using files from the Oregon and Burlington Growth Studies, the present investigation aimed at elucidating the diagnostic reliability of the CVM method in the identification of the mandibular growth peak. To achieve this goal, the individual CVM stages and mandibular growth were recorded longitudinally in a group of subjects. Subsequently, a comprehensive diagnostic reliability analysis of the circumpubertal CVM stages 2, 3, and 4 in the identification of the mandibular growth peak has been carried out as previously reported (13–15).

Subjects and methods

Study design

Subjects were selected from the records of the Oregon and Burlington Growth Studies (data obtained between early 1950s and middle 1970s), and extracted from the American Association of Orthodontists Foundation (AAOF) Craniofacial Growth Legacy Collection (www.aaoflegacycollection.org). Subjects were selected for inclusion if they had a series of consecutive annual lateral cephalograms from the age of 9 to 16 years and if cervical vertebrae 2, 3, and 4 were visible in all films. Exceptions were made for 23 out of 27 Burlington Growth Study cases missing the recording at 15 years (see below). An attempt was made for the inclusion of subjects with normal growth having an overall ANB angle between 0 degree and 6 degrees and total facial divergence (as SN–CoGn angle) between 25 degrees and 42 degrees. Exclusion criteria were: 1. incomplete records; 2. radiographs of poor diagnostic quality; 3. subjects with visible signs of craniofacial syndromes; and 4. evident orthodontic treatment, even minimal including the use of a space maintainer.

From the original samples available on the AAOF Craniofacial Legacy Collection, 50 cases (26 females, 24 males) were included. A total of 23 and 27 cases were derived from the Oregon and Burlington Growth Study collections, respectively. Part of the data on the assessment of maturation of the cervical vertebrae (from the Oregon Growth Study) and mandibular growth peak (from the Burlington Growth Study) are subsets of larger samples reported in two previous investigations (14, 15). With little approximation, magnification factors of 8 and 10 per cent were adopted for the Oregon and Burlington Growth Study samples, respectively.

Assessment of the CVM stages

In the present study, the six-stage CVM method as described by Baccetti *et al*. (2) was assessed according to the objective CVM code staging system (15). Briefly, CVM staging was carried out following a digitization regimen and analysis (16) with a customized cephalometric software (Viewbox, version 3.0, dHAL Software, Kifissia, Greece). For those cases where data was not retrieved from the previous investigation (15), cephalograms were traced by an operator (BS) and checked for accuracy by a second investigator (GP) with details reported elsewhere (15).

The retrieved data was used to the calculate presence/absence of concavity and shape of the vertebral body through a custom-made subroutine. Finally, individual CVM codes were retrieved according to the concavities of the cervical vertebrae 2–4 and shapes of the cervical vertebrae 3 and 4. In detail, lower borders of the cervical vertebrae 2–4 were reported as flat (F) or concave (C), while the shapes of the cervical vertebrae 3 and 4 were reported as trapezoidal (T), rectangular horizontal (H), squared (S), or rectangular vertical (V). Each case was thus defined by a five-letter code, as for instance, CCF-HT indicates concave cervical vertebrae 2 and 3, flat cervical vertebra 4, rectangular horizontal cervical vertebra 3, and trapezoidal cervical vertebra 4. The retrieved CVM codes, either falling within the reported norms (regular cases) or not (exception cases) were also converted into the CVM stages as previously reported (15) and summarized in Supplementary Material.

Assessment of the mandibular growth peak

The total mandibular length was defined as the distance between Condylion (Co) and Gnathion (Gn). For those cases where data was not retrieved from the previous investigation (14), cephalograms were traced by an operator (JP) and checked for accuracy by a second investigator (LC). A customized cephalometric software (Viewbox) was used.

Annual increments in Co–Gn distance were calculated for each subject from the 9–10 years interval to the last 15–16 years interval. Since annual intervals were not always equal to 12 months, annualized increments were derived. Finally, the annual age interval with the greatest increment in Co–Gn distance of the whole series, i.e. mandibular growth peak, was identified and subsequently used for diagnostic reliability analysis (13). Moreover, for the 23 Burlington Growth Study cases missing the 15 years, annualized increments in Co–Gn distance were derived from the bi-annual 14–16 years of age interval. According to this procedure, 3 cases out of 27 showed a mandibular growth peak during the bi-annual 14–16 years of age interval. Therefore, little approximation was followed by using the CVM stage at 14 years for diagnostic reliability analysis.

Method error and data analysis

Method error for the Co–Gn distance was assessed by the method of moments variance estimator on 30 pairs of recordings (15 per collection) randomly selected, and expressed as mean (95% confidence interval [CI]). The repeatability in the assignment of the CVM stage was performed on a further 30 pairs of cases (15 per collection) randomly selected by using the percentage of disagreement, and by both unweighted and linear weighted kappa coefficients presented as mean (95% CI) (17).

Diagnostic reliability analysis

For the pooled sample of the Oregon and Burlington Growth Studies, average growth curves were plotted for annualized increments of total mandibular length (as Co–Gn distance) for females and males. When constructing the curves, individual mandibular growth peaks were aligned (registered) according to the year of the peak, and each prior and succeeding intervals up to ±3 years were also reported.

Diagnostic reliability assessment (for the pooled sample) was carried out for each annual age interval and included sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy (18). Moreover, for each diagnostic parameter, the overall weighted mean was also calculated considering the paired nature of the data. Each diagnostic parameter has been presented as mean and 95 per cent CI. This analysis evaluated the capability of the circumpubertal CVM stages 2, 3, and 4 in the identification of the greatest individual increment of Co–Gn distance in the following year according to previously reported procedures (13, 14, 19). Moreover, diagnostic reliability assessment was also carried out for each of the Oregon and Burlington Growth Studies (Supplementary Data).

The Statistical Package for Social Science (SPSS software 13.0, SPSS Inc., Chicago, Illinois, USA), Comprehensive Meta-Analysis

software 2.0 (Biostat Inc., Englewood, New Jersy, USA), MedCalc software 12.3.3.0 (MedCalc Software, Mariakerke, Belgium), and the interactive Stats Calculator (http://ktclearinghouse.ca/cebm/practise/ca/calculators/statscalc) were used to perform the analyses.

Results

The method error for the Co–Gn distance was 0.8 mm (0.7–1.1). The overall percentage of disagreement for the CVM stages was 13.3 per cent (4 cases out of 30, with 1 stage of disagreement). The unweighted kappa coefficient was 0.84 (0.69–0.98) and the weighted kappa coefficient was 0.92 (0.85–0.99).

The full list of the individual CVM stages from 9 to 15 years (for the Oregon Growth Study sample) and from 9 to 14/15 years of age (for the Burlington Growth Study sample), and corresponding following annual increments in Co–Gn distance is summarized in Table 1. For the Burlington Growth Study sample, a total of 4 cases out of 27 also had recordings at 15 years of age.

With four exceptions (Oregon 250–1, 250–2 and 251, Burlington 366), all the cases had a CVM stage 1 at the beginning of the monitoring. Males showed generally later maturation as compared to that of females. Not all the cases showed the CVM stages 2, 3 and 4. In detail, CVM stage 2 was seen in 34 cases; CVM stage 3 was seen in 33 cases (including Oregon 089-1 and Burlington 544, 742 where it was recorded at 16 years of age); CVM stage 4 was seen in 37 cases (including Burlington 198 where it was recorded at 16 years of age). Whenever detected, the duration of each CVM stage 2, 3, and 4 was variable lasting from 1 to 3 years.

The values of greatest annualized increments in Co–Gn distance ranged from 2.9 mm (Burlington 198, 11–12 years of age) to 8.1 mm (Oregon 153, 13–14 years of age). Most of these mandibular growth peaks were seen at the 11–12 years (n = 10), 12–13 years (n = 8), and 13–14 years (n = 15) age intervals. Males showed generally mandibular growth peaks at later age intervals as compared to females, even though noteworthy variability has been seen over the different annual age intervals and among subjects. In only three Burlington subjects (183-2, 289, 366), mandibular growth peak occurred in the biannual 14–16 years of age.

The annualized increments of mandibular growth, in register according to the peaks, is shown in Figure 1. The greatest annualized mandibular growth values (as mean ± standard deviation) were 4.3 ± 0.2 mm and 5.4 ± 1.6 mm in females and males, respectively. All the other mean values were generally about 2 mm, ranging from 1.2 to 2.6 mm.

Results on the diagnostic reliability assessment of the CVM stages 2, 3, and 4 in the identification of the mandibular growth peak, according to each annual age interval for the whole sample, are summarized in Table 2. Generally, the CVM stages 2, 3, and 4 showed similar behavior with a slightly better diagnostic reliability seen for stages 3 and 4 as compared to stage 2. By excluding the 9–10 years and 10–11 years intervals (with very few mandibular growth peaks), the overall sensitivity ranged from 0.16 (CVM stages 2 and 3) to 0.25 (CVM stage 4), while the overall specificity ranged from 0.80 (CVM stage 2) to 0.87 (CVM stages 3 and 4). Overall PPVs ranged from 0.17 (CVM stage 2) to 0.29 (CVM stage 4), while overall NPVs ranged from 0.83 (CVM stage 2) to 0.87 (CVM stage 4). Finally, the overall diagnostic accuracy for the CVM stages 2, 3, and 4 was 0.70, 0.76, and 0.77, respectively.

The raw Co–Gn distances and related CVM codes for each case and chronologic age are summarized in Supplementary Table 1. For the whole sample, the frequencies of the different CVM codes for regular cases and exception cases are summarized in Supplementary Table 2. The detailed diagnostic parameters for the CVM stages 2, 3, and 4 for each of the Oregon and Burlington Growth Study samples are summarized in Supplementary Tables 3 and 4, respectively.

Discussion

The present longitudinal study assessed the diagnostic reliability of the CVM stages 2, 3, and 4 in the identification of the mandibular growth peak. An objective and transparent CVM code staging system has been applied revealing the existence of exception cases and noteworthy variability in the duration of the different stages. Finally, the CVM the method failed to show a satisfactorily diagnostic reliability.

Since only 3 out of 50 subjects, had a mandibular growth peak in the bi-annual 14–16 years of age interval, the approximations used in the present study were very little with no relevant effect on the results. In the whole sample, the mandibular growth peak occurred generally between 11 and 13 years in females, and between 13 and 14 years in males. These ages are greater as compared to those reported for the Bolton Brush Study sample (11) and similar to those reported from other investigations that recorded standing height instead of mandibular growth peak (20, 21). The annual rate of the total mandibular growth seen herein (Figure 1) is similar to previous evidence reported for Class I untreated subjects (22) and it could be easily identified as shown in Figure 1. Nevertheless, in several cases a very early or late mandibular growth peak was seen (Table 1), making the use of a determined age range unreliable. When analyzing old samples, such in this case, secular trends should be considered as a possible source of variability (23).

Regarding the CVM method, the present study used a recent objective code staging system (15) to analyze the longitudinal changes in a sample of growing subjects. The existence of exception cases (Supplementary Table 2) and their implications have been discussed elsewhere (15), while variable durations of the different CVM stages needs attention as it affects diagnostic reliability, and hence the clinical applicability of the method. The mean duration of each CVM stages from 2 to 5 has originally been reported to last 1 year (2, 7). However, a noteworthy inter-subject variability in terms of duration of the different CVM stages was seen for both the Oregon (15) and Burlington Growth Study samples (Table 1). In the whole sample, and excluding the extreme CVM stages 1 and 6, only 14 cases (28.0%) showed the full range of stages (11 cases could not be assessed due to limitations imposed by the age range). More interestingly, only 3 cases out of 50 (Oregon 76 and 105-1, Burlington 321) showed regular 1-year durations of CVM stages from 2 to 5 (Table 1). In several cases, a discontinuity between consecutive stages was seen with jumping of one or two stages (Table 1). This evidence may be due to fine morphological transitions, according to which a given stage may be apparently missing, while the preceding or the subsequent may last for 2 years. This aspect is an intrinsic limitation of any discrete maturity index, including the hand-and-wrist (24) and third finger middle phalanx (14) maturation methods, and it becomes of particular importance when a single film is available to assess skeletal maturity.

Results obtained from each of the Oregon and Burlington Growth Studies were generally similar, with the exception for the former sample in which most of the mandibular growth peaks occurred after CVM stages 3–5, while for the latter sample peaks occurred mostly after CVM stages 2–4. Although this apparent difference needs to be further investigated in other populations, the whole sample was analyzed since conclusions in terms of diagnostic

Table 1. The individual cervical vertebral maturation stages and corresponding following annualized increments in Co–Gn (mm) according to each annual age interval for the Oregon and Burlington Growth Study cases.

	9–10 years		10–11 years		11–12 years		12–13 years		13–14 years		14–15 years		15–16 years	
ID, Sex	CS	Co–Gn	CS	Co–Gn	CS	Co–Gn	CS	Co–Gn	CS	Co–Gn	CS	Co–Gn	CS	Co–Gn[a]
Oregon Growth Study														
76, F	1	2.6	1	1.9	2	2.2	3	2.4	4	1.9	5	**3.0**	5	1.0
083-1, F	1	1.9	1	1.9	4	**4.8**	6	0.6	6	0.9	6	0.9	NA	NA
100–1, F	1	3.0	3	3.0	4	2.5	4	**3.3**	4	1.9	5	2.1	5	0.2
100–2, F	1	2.2	2	3.1	3	2.5	4	1.9	4	**4.8**	5	1.5	5	2.2
132, F	1	1.7	2	2.0	3	2.7	4	2.4	4	**4.2**	5	0.7	5	2.5
248, F	1	1.9	2	1.0	2	2.3	2	2.5	4	**3.8**	5	1.1	6	1.8
250–1, F	3	2.5	3	2.7	4	3.2	5	**5.9**	6	1.0	6	–0.3	6	0.4
250–2, F	3	1.9	3	1.9	4	**5.6**	5	2.4	5	1.7	6	1.0	6	1.9
251, F	2	1.9	2	1.6	2	2.0	4	2.6	4	**5.5**	4	0.8	4	1.5
251-1, F	1	1.5	2	1.9	4	1.7	4	**4.8**	5	2.1	5	2.5	6	1.5
275, F	1	2.1	1	2.1	4	2.0	5	**3.9**	6	1.7	6	0.4	NA	NA
089-1, M	1	2.1	1	1.6	1	1.7	1	2.0	2	2.3	2	2.5	3	**5.3**
089-2, M	1	1.7	1	2.2	1	2.3	2	2.3	2	3.5	3	**4.7**	4	2.9
105-1, M	1	1.6	1	2.3	2	1.9	3	3.2	4	**4.9**	5	3.1	5	0.9
105-2, M	1	1.9	2	2.5	3	2.2	4	3.1	4	3.6	5	1.2	5	**4.0**
121–3, M	1	2.5	1	2.8	1	3.3	3	3.1	3	3.9	4	**5.5**	5	1.3
144, M	1	2.7	2	3.0	2	1.3	3	3.2	4	3.1	4	**5.6**	5	2.3
153, M	1	2.0	2	2.1	3	1.8	3	1.8	4	**8.1**	5	2.4	6	0.1
179, M	1	2.4	2	3.3	3	2.5	3	4.1	4	**5.9**	5	1.9	5	5.3
183-1, M	1	1.5	1	2.1	1	2.1	1	2.4	3	2.9	4	**4.3**	5	1.4
240, M	1	2.3	1	1.7	1	2.0	3	2.0	3	**7.2**	5	2.7	5	2.1
295, M	1	1.9	1	2.2	2	2.0	2	2.0	2	2.8	4	**7.9**	5	1.8
317-2, M	1	2.3	1	2.2	1	2.1	1	2.6	3	**7.6**	4	2.4	4	3.3
Burlington Growth Study														
153, F	1	**5.5**	1	1.7	2	2.9	2	1.3	5	1.3	5	NA	NA	1.6
163, F	1	1.0	2	1.5	2	1.5	2	**4.5**	3	1.3	4	NA	NA	1.4
188, F	1	3.3	1	**3.7**	4	2.2	4	0.1	4	2.2	4	NA	NA	1.8
198, F	1	2.2	1	2.3	1	**2.9**	1	1.6	1	0.8	3	NA	NA	1.2
208, F	1	1.2	1	**3.5**	2	3.2	5	2.3	5	3.0	5	NA	NA	2.0
316, F	1	4.3	2	1.8	3	**5.2**	4	0.2	4	0.2	5	NA	NA	1.2
321, F	1	2.2	2	**3.5**	3	2.1	4	1.4	5	0.7	6	NA	NA	2.6
391, F	1	2.9	2	1.9	3	**3.7**	4	1.6	4	2.2	4	NA	NA	0.8
487, F	1	2.2	2	0.9	3	**5.5**	3	3.3	4	1.6	5	NA	NA	2.4
595, F	1	**4.5**	3	1.6	3	1.4	4	0.5	4	2.9	5	NA	NA	4.6
602, F	1	3.2	1	2.8	1	**3.8**	3	1.6	6	0.5	6	NA	NA	3.2
619, F	1	3.1	1	3.0	1	**4.5**	3	4.1	4	1.4	5	NA	NA	0.8
631, F	2	2.6	2	1.7	3	2.7	3	**4.1**	4	0.8	4	NA	NA	3.4
855, F	1	1.7	1	2.3	1	**3.5**	2	0.2	3	0.6	4	NA	NA	2.8
1391, F	1	–0.1	1	0.8	1	1.5	2	0.7	2	**4.6**	4	NA	NA	3.8
135, M	1	1.7	1	2.7	1	3.0	2	0.3	3	**6.1**	3	NA	NA	1.4
166, M	1	1.9	2	2.5	3	3.1	4	1.7	4	**7.5**	4	NA	NA	4.8
183-2, M	1	1.0	1	3.0	1	1.7	2	0.8	2	2.4	2	NA	NA	**7.2**
231, M	1	2.0	1	2.5	1	0.5	1	**4.4**	2	2.9	5	NA	NA	4.0
289, M	1	1.9	1	1.1	1	2.9	1	1.3	2	4.2	2	NA	NA	**8.6**
366, M	2	2.0	2	1.5	3	2.1	3	1.2	4	2.3	4	NA	NA	**6.0**
392, M	1	3.9	1	1.5	1	1.5	2	0.7	2	**4.5**	2	NA	NA	4.2
544, M	1	1.2	2	2.8	2	0.3	3	**5.9**	4	3.2	4	1.3	5	0.7
742, M	1	2.0	1	0.9	1	1.4	2	2.2	2	**2.5**	2	–0.1	2	0.7
763, M	1	0.4	1	1.8	1	1.5	1	3.1	2	**6.0**	4	NA	NA	2.8
863, M	1	2.5	1	1.9	1	1.4	1	3.4	1	3.3	3	**7.5***	5	1.3
871, M	1	4.0	2	0.4	2	**4.5**	2	1.5	4	3.3	5	3.9	5	1.9

CS, cervical vertebral maturation stage; F, female; M, male; NA, not available. In bold, maximum individual annual increments in Co–Gn.

[a]For the Burlington Growth Study, bi-annual increments in Co-Gn (14–16 years interval) is reported instead of the annual increments (15–16 years interval) except for those cases where the 15 years recording was available. Part of this data was derived from previous investigations (13, 14).

reliability would be similar irrespective of the Growth Study under investigation (see Supplementary Tables 3 and 4).

In the whole sample, mandibular growth peak occurred after CVM stages 2, 3, and 4 in 9 (18%), 12 (24%), and 16 (32%) cases, respectively. Moreover, nine mandibular growth peaks occurred after CVM stage 1 and other 4 after the stage 5 (Table 1). Therefore, CVM stages 3 and 4 taken together, were able to identify no more than 56 per cent of the mandibular growth peaks. This evidence is

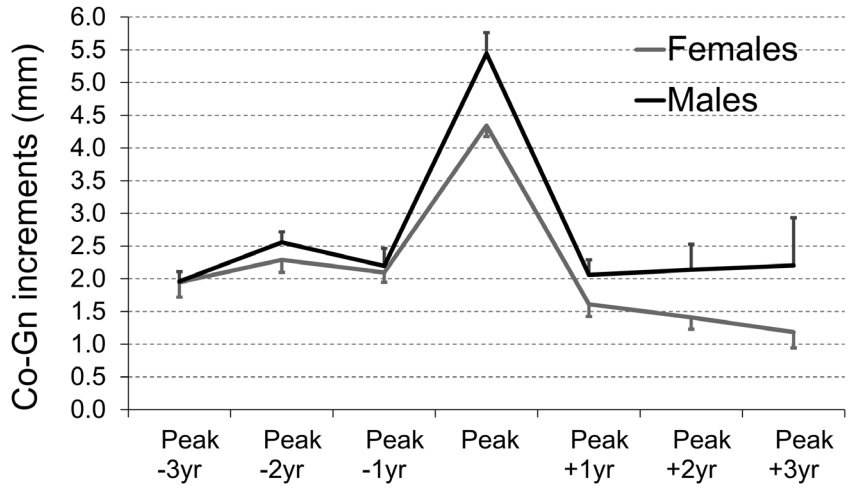

Figure 1. The mandibular growth peak defined as the greatest annualized increment in Co–Gn distance, by pooling the Oregon and Burlington Growth Study cases and according to the sexes. Data are presented as mean ± standard error.

Table 2. Diagnostic reliability of the cervical vertebral maturation stages 2–4 in the identification of the mandibular growth peak, as annualized increment in Co–Gn distance according to each annual age interval from 9 to 16 years by pooling the Oregon and Burlington Growth Study cases.

	Age intervals						
CVM stage/diagnostic parameter	9–10 years	10–11 years	11–12 years	12–13 years	13–14 years	14–15 years/ 14–16 years, a	Overall
CS2							
Sensitivity	0	0.25 (0–0.67)	0.10 (0–0.29)	0.13 (0–0.35)	0.27 (0.04–0.49)	0.20 (0–0.45)	0.16 (0–0.37)
Specificity	**0.94** (0.87–1.01)	**0.61** (0.47–0.75)	**0.75** (0.62–0.88)	**0.74** (0.61–0.87)	**0.83** (0.70–0.95)	**0.93** (0.84–1)	**0.80** (0.68–0.92)
PPV	**0**	**0.05** (0–0.15)	**0.09** (0–0.26)	**0.08** (0–0.24)	**0.40** (0.10–0.70)	**0.40** (0–0.83)	**0.17** (0–0.39)
NPV	**0.96** (0.90–1)	**0.90** (0.80–1)	**0.77** (0.64–0.90)	**0.82** (0.69–0.94)	**0.73** (0.59–0.86)	**0.82** (0.71–0.93)	**0.83** (0.72–0.95)
Accuracy	**0.90** (0.82–0.98)	**0.58** (0.44–0.72)	**0.62** (0.49–0.75)	**0.64** (0.51–0.77)	**0.66** (0.53–0.79)	**0.78** (0.67–0.89)	**0.70** (0.58–0.82)
CS3							
Sensitivity	**0**	**0**	**0.30** (0.02–0.58)	**0.25** (0–0.55)	**0.20** (0–0.40)	**0.20** (0–0.45)	**0.16** (0–0.36)
Specificity	**0.96** (0.9–1)	**0.91** (0.84–0.99)	**0.75** (0.62–0.88)	**0.74** (0.61–0.87)	**0.89** (0.78–0.99)	**0.95** (0.88–1.02)	**0.87** (0.77–0.96)
PPV	**0**	**0**	**0.23** (0–0.46)	**0.15** (0–0.35)	**0.43** (0.06–0.80)	**0.50** (0.01–0.99)	**0.22** (0–0.5)
NPV	**0.98** (0.94–1)	**0.93** (0.86–1)	**0.81** (0.68–0.94)	**0.84** (0.72–0.96)	**0.72** (0.59–0.85)	**0.83** (0.72–0.94)	**0.85** (0.75–0.96)
Accuracy	**0.94** (0.87–1)	**0.86** (0.76–0.96)	**0.66** (0.53–0.79)	**0.66** (0.53–0.79)	**0.68** (0.55–0.81)	**0.80** (0.69–0.91)	**0.76** (0.65–0.88)
CS4							
Sensitivity	**0**	**0**	**0.22** (0–0.49)	**0.25** (0–0.55)	**0.53** (0.28–0.79)	**0.50** (0.19–0.81)	**0.25** (0.03–0.47)
Specificity	**1**	**1**	**0.88** (0.78–0.98)	**0.76** (0.63–0.89)	**0.63** (0.47–0.79)	**0.73** (0.59–0.86)	**0.83** (0.72–0.94)
PPV	—	—	**0.29** (0–0.62)	**0.17** (0–0.38)	**0.38** (0.17–0.59)	**0.31** (0.09–0.54)	**0.29** (0.05–0.52)
NPV	**0.98** (0.94–1)	**0.94** (0.87–1)	**0.84** (0.73–0.95)	**0.84** (0.73–0.96)	**0.76** (0.60–0.91)	**0.85** (0.73–0.97)	**0.87** (0.76–0.98)
Accuracy	**0.98** (0.94–1)	**0.94** (0.87–1)	**0.76** (0.64–0.88)	**0.68** (0.55–0.81)	**0.60** (0.46–0.74)	**0.68** (0.55–0.81)	**0.77** (0.66–0.88)
n. peak	2	3	10	8	15	12	50

Data are presented as mean (95% confidence interval) with $n = 50$ in each age interval (23 and 27 cases form the Oregon and Burlington Growth Study, respectively). Bold values indicates the greatest of the series. NPV, negative predictive value. —, not derivable; PPV, positive predictive value.

[a]14–15 years interval and 14–16 years interval in the Oregon and Burlington Growth Study, respectively.

in contrast with original reports on the CVM method (2, 7) and in accordance with a previous investigation (10) on the Burlington Growth Sample. Other longitudinal investigations on the Bolton-Brush growth Study sample (11) and on a European sample (25) reported a poor correlation of the CVM method with the mandibular growth peak. However, given significant differences in study designs, cephalometric recordings and data analysis, comparisons of the conclusions of the above studies should be undertaken with caution (1). For instance, skeletal Class II females, with an expected minimal mandibular growth peak, were analyzed (22), or maturation of the cervical vertebrae has been recorded with specific methods (3) noteworthy different than those being reported here. More importantly, and irrespective of the version of the CVM method, none of the previous longitudinal studies reported data on the diagnostic reliability.

Overall sensitivity for each of CVM stages 2, 3, and 4 was very low being no greater than 0.25 (CVM stage 4). Even considering the full 95 per cent CI, this parameter remains low for each of the investigated stages. On the contrary, the specificity ranged from 0.80 to 0.87 for the CVM stages 2 and 3, respectively. However, when dealing with several possible clustering, an important diagnostic parameter is the PPV that gives an indication of the capability

of a given CVM stage in the identification of the mandibular growth peak (irrespective of the number of true negative cases belonging to the other stages). Herein, the overall PPVs were unsatisfactory with values not greater than 0.29 (CVM stage 4); on the contrary, the overall NPVs were high with values of at least 0.83 (CVM stage 4). This finding was due to the relatively large number of true negative cases in each annual age interval cluster. Therefore, irrespective of the CVM stage, misdiagnosis would mainly due to false positives consequent to the prolonged duration of the CVM stages 3 and 4 (Table 1). According to this evidence, the lack of a satisfactorily diagnostic reliability would be due to false positives. Although mandibular growth peak would not occur exclusively between the CVM stages 3 and 4 (2, 7), the CVM stages 2, 3, and 4 may still be defined as circumpubertal (Table 1).

The overall diagnostic accuracy of the circumpubertal CVM stages 2, 3, and 4 was limited being no greater than 0.77 (CVM stage 4). Therefore, precise identification of the imminent mandibular growth peak cannot rely on the CVM method alone. In the previous diagnostic study (13), the pooled CVM stages 3 and 4 in the identification of imminent mandibular growth peak, yielded a diagnostic accuracy ranging from 0.67 to 0.91 according to the different ages. Although a full comparison of these results is not feasible, the present evidence on diagnostic reliability is slightly worse as compared to that previously reported for the CVM method using the University of Michigan Growth Study sample (13). This may be explained by the concept that the sample used in the previous investigation was the same from which the CVM method was derived.

Recent meta-analyses (26, 27) on the supplementary mandibular elongation due to functional treatment in skeletal Class II patients, treated during the circumpubertal CVM stages 3 and 4, uncovered a clinically relevant effect as compared to matched untreated controls. However, a large variability in the skeletal effects has also been reported (26, 27). Considering that herein more than half of the cases showed a mandibular growth peak during or after the appearance of the CVM stages 3 and 4 (Table 1), the present results may explain, at least in part, the reported large variability. Accordingly, only a part of the patients included in those trials might have been treated during the occurrence of the mandibular growth peak.

Strength and limitations

In the present study, a full transparent analysis has been carried out with the reporting of the detailed case-by-case outcomes. The method used herein to assess the CVM code and corresponding CVM stage (15), although not identical, is very close to what has been previously reported (2).

A limitation might reside in the secular trend for both the onset of the pubertal growth peak (23) and maturation of the cervical vertebrae (28), especially considering that the Oregon and Burlington Growth Studies obtained data about half a century ago. The use of annual recordings has an inherent error when a discrete staging system is used. For the Burlington Growth Study sample, little approximation was necessary because of the lack of the 15 years recording.

Clinical implications

Diagnostic reliability of the circumpubertal CVM stages in the identification of imminent mandibular growth peak is not satisfactory. Unpredictable duration of the stages would further limit the clinical applicability of the method. This is especially the case when considering that lateral head films are usually available as a pre-treatment record and optimal treatment timing has to be delayed for an undetermined period after diagnosis. From a clinical standpoint, 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.

Conclusions

None of the CVM stages 2, 3, and 4 alone reached a satisfactory diagnostic reliability in the identification of imminent mandibular growth peak.

Acknowledgments

Part of this study was made possible by use of material from the Burlington Growth Centre, Faculty of Dentistry, University of Toronto, which was supported by funds provided by Grant 1. (No. 605-7-299) National Health Grant (Canada), (data collection); 2. Province of Ontario Grant PR 33 (duplicating), and 3. the Varsity Fund (for housing and collection). The Authors wish to express gratitude Dr. Sean Curry (University of the Pacific, CA) for providing images from the Oregon and Burlington Growth Studies, and to Dr. Alberto Bianchet and Dr. Valentina Sbardella (University of Trieste) for valuable assistance.

Conflict of interest

The authors declare no conflict of interest.

References

1. Perinetti, G. and Contardo, L. (2017) Reliability of growth indicators and efficiency of functional treatment for skeletal Class II malocclusion: current evidence and controversies. *BioMed Research International*, 2017, 1367691.
2. Baccetti, T., Franchi, L. and McNamara J.A., Jr. (2005) The cervical vertebral maturation (CVM) method for the assessment of optimal treatment timing in dentofacial orthopedics. *Seminars in Orthodontics*, 11, 119–129.
3. Beit, P., Peltomäki, T., Schätzle, M., Signorelli, L. and Patcas, R. (2013) Evaluating the agreement of skeletal age assessment based on hand-wrist and cervical vertebrae radiography. *American Journal of Orthodontics and Dentofacial Orthopedics*, 144, 838–847.
4. Hassel, B. and Farman, A.G. (1995) Skeletal maturation evaluation using cervical vertebrae. *American Journal of Orthodontics and Dentofacial Orthopedics*, 107, 58–66.
5. Lamparski, D.G. (1972) *Skeletal Age Assessment Utilizing Cervical Vertebrae*. Dissertation. The University of Pittsburgh.
6. O'Reilly, M.T. and Yanniello, G.J. (1988) Mandibular growth changes and maturation of cervical vertebrae–a longitudinal cephalometric study. *The Angle Orthodontist*, 58, 179–184.
7. Franchi, L., Baccetti, T. and McNamara, J.A., Jr. (2000) Mandibular growth as related to cervical vertebral maturation and body height. *American Journal of Orthodontics and Dentofacial Orthopedics*, 118, 335–340.
8. Chen, L.L., Xu, T.M., Jiang, J.H., Zhang, X.Z. and Lin, J.X. (2008) Quantitative cervical vertebral maturation assessment in adolescents with normal occlusion: a mixed longitudinal study. *American Journal of Orthodontics and Dentofacial Orthopedics* 134, 720 e721-e727.
9. Ball, G., Woodside, D., Tompson, B., Hunter, W.S. and Posluns, J. (2011) Relationship between cervical vertebral maturation and mandibular growth. *American Journal of Orthodontics and Dentofacial Orthopedics*, 139, e455–e461.

10. Gray, S., Bennani, H., Kieser, J.A. and Farella, M. (2016) Morphometric analysis of cervical vertebrae in relation to mandibular growth. *American Journal of Orthodontics and Dentofacial Orthopedics*, 149, 92–98.
11. Mellion, Z.J., Behrents, R.G. and Johnston, L.E. Jr. (2013) The pattern of facial skeletal growth and its relationship to various common indexes of maturation. *American Journal of Orthodontics and Dentofacial Orthopedics*, 143, 845–854.
12. Masoud, M.I., Marghalani, H.Y., Bamashmous, M., Alamoudi, N.M., El Derwi, D., Masoud, I.M., Allareddy, V. and Gowharji, N.F. (2015) Predicting changes in mandibular length and total anterior facial height using IGF-1, cervical stage, skeletal classification, and gender. *Progress in Orthodontics*, 16, 7.
13. Perinetti, G., Contardo, L., Castaldo, A., McNamara, J.A. Jr and Franchi, L. (2016) Diagnostic reliability of the cervical vertebral maturation method and standing height in the identification of the mandibular growth spurt. *The Angle Orthodontist*, 86, 599–609.
14. Perinetti, G., Sbardella, V. and Contardo, L. (2017) Diagnostic reliability of the third finger middle phalanx maturation (MPM) method in the identification of the mandibular growth peak. *European Journal of Orthodontics*, 39, 194–201.
15. Perinetti, G., Bianchet, A., Franchi, L. and Contardo, L. (2017) Cervical vertebral maturation: an objective and transparent code staging system applied to a 6-year longitudinal investigation. *American Journal of Orthodontics and Dentofacial Orthopedics*, 151, 898–906.
16. Perinetti, G., Caprioglio, A. and Contardo, L. (2014) Visual assessment of the cervical vertebral maturation stages: a study of diagnostic accuracy and repeatability. *The Angle Orthodontist*, 84, 951–956.
17. Perinetti, G. (2017) StaTips part III: assessment of the repeatability and rater agreement for nominal and ordinal data. *South European Journal of Orthodontics and Dentofacial Research*, 4, 3–4.
18. Greenhalgh, T. (1997) How to read a paper. Papers that report diagnostic or screening tests. *BMJ (Clinical research ed.)*, 315, 540–543.
19. Perinetti, G., Sossi, R., Primozic, J., Ierardo, G. and Contardo, L. (2017) Diagnostic reliability of mandibular second molar maturation in the identification of the mandibular growth peak: a longitudinal study. *The Angle Orthodontist*, 87, 665–671.
20. Hägg, U. and Taranger, J. (1982) Maturation indicators and the pubertal growth spurt. *American Journal of Orthodontics*, 82, 299–309.
21. Fishman, L.S. (1982) Radiographic evaluation of skeletal maturation. A clinically oriented method based on hand-wrist films. *The Angle Orthodontist*, 52, 88–112.
22. Stahl, F., Baccetti, T., Franchi, L. and McNamara, J.A. Jr. (2008) Longitudinal growth changes in untreated subjects with Class II Division 1 malocclusion. *American Journal of Orthodontics and Dentofacial Orthopedics*, 134, 125–137.
23. Parent, A.S., Teilmann, G., Juul, A., Skakkebaek, N.E., Toppari, J. and Bourguignon, J.P. (2003) The timing of normal puberty and the age limits of sexual precocity: variations around the world, secular trends, and changes after migration. *Endocrine Reviews*, 24, 668–693.
24. Houston, W.J. (1980) Relationships between skeletal maturity estimated from hand-wrist radiographs and the timing of the adolescent growth spurt. *European Journal of Orthodontics*, 2, 81–93.
25. Engel, T.P., Renkema, A.M., Katsaros, C., Pazera, P., Pandis, N. and Fudalej, P.S. (2016) The cervical vertebrae maturation (CVM) method cannot predict craniofacial growth in girls with Class II malocclusion. *European Journal of Orthodontics*, 38, 1–7.
26. Perinetti, G., Primožič, J., Furlani, G., Franchi, L. and Contardo, L. (2015) Treatment effects of fixed functional appliances alone or in combination with multibracket appliances: a systematic review and meta-analysis. *The Angle Orthodontist*, 85, 480–492.
27. Perinetti, G., Primozic, J., Franchi, L. and Contardo, L. (2015) Treatment effects of removable functional appliances in pre-pubertal and pubertal Class II patients: a systematic review and meta-analysis of controlled studies. *PLoS One*, 10, 1–35.
28. Antoun, J.S., Cameron, C., Sew Hoy, W., Herbison, P. and Farella, M. (2015) Evidence of secular trends in a collection of historical craniofacial growth studies. *European Journal of Orthodontics*, 37, 60–66.