Preliminary validation of serotransferrin and vitamin D binding protein in the gingival crevicular fluid as candidate biomarkers for pubertal growth peak in subjects with Class I and Class II malocclusion

Xi Wen and Yan Gu
Beijing, China

Introduction: Identification of pubertal growth peak is of great importance for the orthopedic treatment of Class II malocclusion. Our previous work demonstrated that vitamin D binding protein (DBP) and serotransferrin (TF) in gingival crevicular fluid (GCF) could be candidate biomarkers of pubertal growth peak. This research aimed to preliminarily validate TF and DBP in subjects with Class I and Class II malocclusion, to compare their diagnostic accuracy, and to construct a statistic model to help the diagnosis of skeletal pubertal peak. **Methods:** Sixty-six circumpubertal subjects were recruited, including 32 subjects with Class I malocclusion and 34 subjects with Class II malocclusion. All subjects were divided into prepubertal, pubertal, and postpubertal groups according to their cervical vertebral maturation stages. GCF samples were collected, and the concentration of DBP and TF were detected by enzyme-linked immunosorbent assay. **Results:** Percentage of TF in GCF was significantly higher in pubertal than in prepubertal and postpubertal groups, in subjects with Class I and Class II malocclusion, whereas the difference observed in DBP was less significant. The diagnostic accuracy of TF was better than DBP and chronological age. The most optimal thresholds of maxillary and mandibular TF in distinguishing pubertal from nonpubertal subjects were 4.20% and 4.09%, respectively. The combination of TF and age exhibited the best diagnostic accuracy. **Conclusions:** TF in GCF could be considered as a potential biomarker of pubertal peak and can assist the diagnosis of skeletal pubertal peak. (Am J Orthod Dentofacial Orthop 2021;159:415-25)

It is well known that pubertal growth peak plays a vital role in decision making of orthodontic treatment timing, especially for subjects with Class II malocclusion with retrusive mandible.[1-5] Traditional x-ray methods were usually applied to identify skeletal maturation stage and to determine the most appropriate treatment timing, such as hand wrist method[6] and cervical vertebral maturation method (CVM).[7] However, the hand wrist method has its limitation as patients have to suffer from extra radiation exposure. The CVM stage can be evaluated on a lateral cephalogram, which is a part of routine pretreatment examination. In addition, CVM was proved to be a reliable method to identify the pubertal peak.[8] Reproducibility of CVM assessment was also satisfying in the circumstance that the observer had accepted a professional training course.[9]

However, because of the qualitative nature of the CVM method, researchers have been trying to explore new quantitative indicators as alternatives to predict pubertal growth peak, among which biomarkers in body fluids such as serum[10,11] and saliva[12] turn out to be the most attractive. Gingival crevicular fluid (GCF) was considered as an ideal source of biomarkers for its simple

From the Department of Orthodontics, National Engineering Laboratory for Digital and Material Technology of Stomatology, Beijing Key Laboratory of Digital Stomatology, Peking University School and Hospital of Stomatology, Beijing, China.
All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest, and none were reported.
This work was supported by the National Natural Science Foundation of China (grant nos. NSFC81771027, 81970979) and the Ministry of Science and Technology of the People's Republic of China (grant no. 2018FY101005).
Address correspondence to: Yan Gu, Department of Orthodontics, Peking University School and Hospital of Stomatology, #22 Zhongguancun South Avenue, Haidian District, Beijing 100081, China; e-mail, guyan96@126.com.
Submitted, October 2019; revised, December 2019; accepted, January 2020.
0889-5406/$36.00

https://doi.org/10.1016/j.ajodo.2020.01.025

and noninvasive collection procedure.[13,14] In our previous work,[15] we collected GCF samples from 20 pubertal and 20 postpubertal subjects and applied liquid chromatography-tandem mass spectrometry combined with tandem mass tags labeling to compare the whole GCF proteome between pubertal and postpubertal subjects. Through gene ontology analysis, we found that vitamin D binding protein (DBP) was the most significantly enriched protein (P <0.0001). Meanwhile, serotransferrin (TF) in terms of "iron ion homeostasis" exhibited a similar function and was reported to be enriched in the serum of pubertal subjects.[15] Therefore, TF and DBP were chosen as candidate GCF biomarkers indicating pubertal peak.[15]

TF was synthesized in the liver and can bind to and transport iron ions in circulation between sites of absorption, storage, and use.[16] Anttila et al[17] reported that serum TF was significantly higher in pubertal boys than in nonpubertal boys. Misaki et al[18] also found that in adolescent boys, the rate of growth in height was significantly correlated with serum TF levels. Malecki et al[19] created an animal model of hypotransferrinemia and found that lack of TF would compromise the mechanical properties of the femur and that TF was important for normal bone mineralization.[19] However, whether TF level in GCF was associated with puberty remains unclear. DBP transports 80%-90% of serum 25-hydroxy vitamin D and 1,25-dihydroxy vitamin D and can promote intestinal calcium absorption and regulate maturation and mineralization of bone.[20,21] A previous study showed that 1,25-dihydroxy vitamin D levels peak during puberty to meet the high demand for calcium in this critical phase of bone development.[22] More evidence is still needed to reveal the relationship between DBP level and pubertal growth peak.

Our study aimed to primarily validate TF and DBP as biomarkers for pubertal growth peak in subjects with Class I and Class II malocclusion, to compare their diagnostic accuracy, and to explore their most optimal threshold in distinguishing pubertal subjects from nonpubertal subjects. Ultimately, we combined chronological age with GCF biomarkers to construct a statistic model to help in the diagnosis of pubertal peak.

MATERIAL AND METHODS

Subjects were collected from patients who sought treatment at the Department of Orthodontics, Peking University School and Hospital of Stomatology from March 2017 to August 2019. This research was approved by the Biomedical Ethics Committee of Peking University Health Science Center (approval no. PKUSSIRB-201735069) and was performed in accordance with the ethical standard in the 1964 Declaration of Helsinki. Informed consents were received from patients and their parents.

We recruited subjects according to the inclusion criteria as follows: (1) aged between 6 and 18 years; (2) healthy periodontal status; (3) skeletal Class I ($0 < ANB < 5°$) or Class II ($ANB > 5°$) malocclusion; (4) no history of orthodontic treatment; (5) no systematic disease; (6) not taken any medication in the past 3 months; and (7) Mongolian.

All subjects took lateral cephalograms as part of routine pretreatment examination to determine the sagittal skeletal relationship and CVM stages. Lateral cephalograms were analyzed using Huazheng System (version 1.2; Peking University School of Stomatology, Beijing, China).

ANB angle (°), overjet (mm), and molar relationship were used to distinguish subjects with Class I and Class II malocclusion. Subjects with $ANB > 5°$, overjet > 3 mm, and distal molar relationship were included in the Class II malocclusion group, whereas subjects with $0 < ANB < 5°$, 0 < overjet < 3 mm, and neutral molar relationship were included in the Class I group.

CVM stage was determined according to the method proposed by Baccetti et al.[23] Subjects at CVM stages 1 and 2 were included in the prepubertal group. In contrast, subjects at CVM stages 3 and 4 were included in the pubertal group, and subjects at CVM stages 5 and 6 were included in the postpubertal group.

The interobserver agreement of the CVM stage was tested by 2 different orthodontists (Y.G. and X.W.). Intraobserver agreement of CVM stage was determined by the same orthodontist (X.W.) initially and 1 week later. Two orthodontists were both blinded to the basic information of studied subjects.

The minimum sample size was calculated on the basis of the formula applying to a 1-way ANOVA 2-tailed test proposed by Chow et al.[24] The type I error probability was set as 0.05, and type II error probability was 0.2. According to the result of our preliminary study, for DBP, the value of μA-μB and σ were 0.11 and 0.17, respectively, and hence the minimal sample size should be 50. For TF, the value of μA-μB and σ were 3.0 and 5.2, respectively, and hence, the minimal sample size should be 63.

Periodontal examination was performed on each subject before GCF collection. Only those with probing depth ≤ 3 mm, bleeding index ≤ 3, and no attachment loss were recruited. All subjects accepted ultrasonic supragingival scaling and oral hygiene instruction 1 week before GCF collection.

GCF sample was collected between 8 AM and 10 AM. Before collection, the tooth surface was carefully dried,

and cotton rolls were used to avoid contamination of saliva. GCF was collected in mesiolabial and distolabial sites of both maxillary and mandibular central incisors. Paper points (no. 30; Tianjin Dayading Medical Treatment Appliance Company, Tianjin, China) were gently inserted into the gingival sulcus until minimal resistance was felt and then were held in situ for 60 seconds. Paper points contaminated by blood would be discarded. After 2 minutes, the collection procedure was repeated using new paper points to ensure adequate volume of GCF samples were collected.

GCF samples from maxillary and mandibular incisors were separated into different 1.5 mL centrifuge tubes (Axygen, Corning Company, Union City, Calif). Each tube contained 8 paper points. The tips of paper points were cut from where they were visibly wetted by GCF and were incubated in 50 μL phosphate buffered solution. After vortexed for 1 minute, GCF samples were centrifuged at 5000 rpm for 10 minutes at 4°C. The supernatant was collected. The same procedure was repeated 3 times, and the elution solution was pooled and stored at −80°C for further analysis.

A bicinchoninic acid kit was used to determine the total protein concentration of GCF samples (Beyotime Biotechnology, Beijing, China). Standard samples were prepared according to the instruction of the manufacturer. An EI×808 spectrophotometer (BioTek, Colchester, Vt) was used to detect the optical densities and calculate the total protein concentration.

Commercially available enzyme-linked immunosorbent assay kits of human DBP and TF (Qisong Company, Beijing, China) were applied to determine the concentration of DBP and TF in each GCF sample according to the instruction of the manufacturer. Each sample had 3 replicates, and the mean value was calculated for further statistical analysis.

To further evaluate the diagnostic accuracy of TF and DBP in distinguishing pubertal from nonpubertal subjects, all samples were regrouped. Subjects at CVM stages 1-2 or 5-6 were included in the nonpubertal group, whereas the rest were included in the pubertal group. Thus, the dependent variable was transformed into a binary variable *pubertal peak*, for which at peak or not were defined as *1* and *0*, respectively. Independent variables, including GCF biomarkers and chronological age, were continuous variables. Hence, logistic regression (LR) model analysis was performed to quantificationally describe the relationship between pubertal peak and its influencing factors (version 3.6.1; R Foundation for Statistical Computing, Vienna, Austria). Through the LR model, the predictive probability of subjects being at pubertal peak could be calculated. Backward stepwise method was used to screen for significant independent variables. The basic form of the LR model was as follows:

$$\text{Predictive probability} = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots}}{1+e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots}}$$

In our study, there were 5 potential influencing factors that could have an impact on *pubertal peak*, including TF, DBP, chronological age, sex, and skeletal type. Specifically, chronological age exhibited parabolic distribution relative to *pubertal peak* and entered the model in the form of ($\beta_0 + \beta_1 x^2 + \beta_2 x$), whereas the sex and skeletal type entered as interactive variables with chronological age. To separately evaluate and compare their diagnostic accuracy, LR models were constructed using different combinations of independent variables. They were listed as follows:

(1) Model 1: Maxillary TF (Max-TF).
(2) Model 2: Mandibular TF (Md-TF).
(3) Model 3: Maxillary DBP (Max-DBP).
(4) Model 4: Mandibular DBP (Md-DBP).
(5) Model 5: Chronological age.
(6) Model 6: Comprehensive model including Max-TF, Md-TF, Max-DBP, Md-DBP, age, sex, and skeletal type.

Receiver operator characteristic (ROC) curve were used to evaluate and compare the diagnostic accuracy of 6 models mentioned above. For each model, the true positive rate (TPR) and false positive rate (FPR) of each candidate threshold were calculated. ROC curves were first calculated using back substitution method, which was also applied in some of the previous studies.[25,26] However, this method used the same samples for training models and verification, which might exaggerate the final result. Therefore, we further developed a 10-fold cross-validation algorithm using R software to test the diagnostic accuracy of GCF biomarkers. All samples were randomly divided into the training group and verification group, and the regression model learned from the training group was used to predict the verification group. This procedure were repeated 10 times until all samples had been predicted. ROC curves by 10-fold cross-validation method were drawn by R software as well.

Through LR models mentioned earlier, the predictive probability of a subject being at pubertal peak could be calculated, and an optimal threshold was needed to classify this subject into the pubertal or nonpubertal group. Youden index was defined as *Sensitivity + Specificity-1* and was calculated on the basis of the ROC curve to determine the most optimal threshold. When the Youden

index reached the maximum, the corresponding predictive probability could be used to reversely calculate the value of GCF biomarkers through the formula of LR models, and this value of GCF biomarkers was considered as the most optimal threshold in distinguishing pubertal from nonpubertal subjects.

Statistical analysis

To reduce the error caused by different GCF volume in each sampling procedure, we normalized the concentration of DBP and TF by total protein concentration, which was calculated by dividing DBP or TF concentration by GCF total protein concentration, and the results were presented as percentage of DBP or TF in GCF total protein.

Kappa values were used to evaluate the interobserver and intraobserver agreement of CVM stages. The normality of all variables was tested by the Shapiro-Wilk test. If the variables were normally distributed, independent sample *t* tests (for 2 independent samples) and 1-way ANOVA (for 3 independent samples) were used to analyze the difference between groups. If variables were not normally distributed, then they were analyzed using the Kruskal-Wallis H test (for 2 independent samples) and Mann-Whitney U test (for 3 independent samples).

Differences with $P < 0.05$ were considered statistically significant. All statistical analysis was accomplished using SPSS (version 19.0; IBM, Armonk, NY) and R software. All figures were created using GraphPad Prism 5 (version 5.01; GraphPad Software, San Diego, Calif) and R software.

RESULTS

A total of 66 patients were recruited, including 32 subjects with Class I malocclusion (mean age, 11.8 ± 2.4 years) and 34 subjects with Class II malocclusion (mean age, 12.2 ± 2.4 years). Detailed information was shown in Table I and Supplementary Figure. Kappa values of CVM stages for interobserver and intraobserver reliability are 0.84 and 0.96, respectively.

The percentage of TF in GCF was significantly higher in pubertal subjects than in prepubertal and postpubertal subjects, both in maxilla and mandible (Fig 1, *A*). A similar trend was also observed in the DBP change pattern (Fig 1, *B*). However, the difference in DBP between pubertal subjects and nonpubertal subjects was not statistically significant.

At the postpubertal stage, mandibular TF was significantly higher than maxillary TF (Fig 1, *A*). At the prepubertal and pubertal stage, mandibular DBP was significantly higher than maxillary DBP (Fig 1, *B*).

Table I. Detailed information of recruited subjects with Class I and Class II malocclusion

	Subjects		*Sex*		
Groups	*Class I*	*Class II*	*Male*	*Female*	*Chronological age*
Prepubertal	12	11	14	9	10.48 ± 2.19
Pubertal	10	13	14	9	11.78 ± 1.04
Postpubertal	10	10	7	13	14.05 ± 2.35

When comparing the percentage of TF (Fig 2, *A* and *C*) and DBP (Fig 2, *B* and *D*) between subjects with Class I and Class II malocclusion, no significant difference was observed at the pubertal stage, no matter in maxilla or mandible.

For maxilla, the percentage of DBP in GCF was higher in subjects with Class II malocclusion than in subjects with Class I malocclusion at the pubertal stage (Fig 2, *B*). For mandible, the percentage of DBP in GCF was lower in subjects with Class II malocclusion than in subjects with Class I malocclusion at the pubertal stage (Fig 2, *D*).

The change pattern of TF (Fig 3, *A* and *C*) and DBP (Fig 3, *B* and *D*) between males and females were also compared. However, no statistically significant difference was observed, which indicated that the level of DBP and TF in GCF were similar between males and females during the prepubertal, pubertal, and postpubertal stages.

Six LR models using different combinations of independent variables were shown in Table II. For Model 6 that included all possible variables, after a backward stepwise screening, only Max-TF and age stayed in the model, which indicated that among all studied factors, Max-TF and age were the most significant influencing factors in the diagnosis of pubertal peak, at least from a statistical perspective.

ROC curves of 6 different models drawn by back substitution method (Fig 4, *A-D*, *I*, and *J*) and 10-fold cross-validation method (Fig 4, *E-H*, *K*, and *L*) were presented separately, and the area under the curve (AUC) was calculated to quantitatively evaluate the diagnostic accuracy (Fig 4, *M* and *N*). First, ROC curves of 2 candidate GCF biomarkers (Model 1-4) were analyzed and compared (Fig 4, *A-H*). Results showed that the AUC of TF was greater than DBP in both maxilla or mandible, among which the Max-TF exhibited the greatest accuracy in distinguishing pubertal from nonpubertal subjects, and this conclusion was verified by both statistical methods. AUC of Max-TF was more than 0.8 and could be considered to have satisfactory diagnostic accuracy.

From Figure 4, *I* and *K*, it was noticed that the diagnostic accuracy of chronological age (Model 5) in

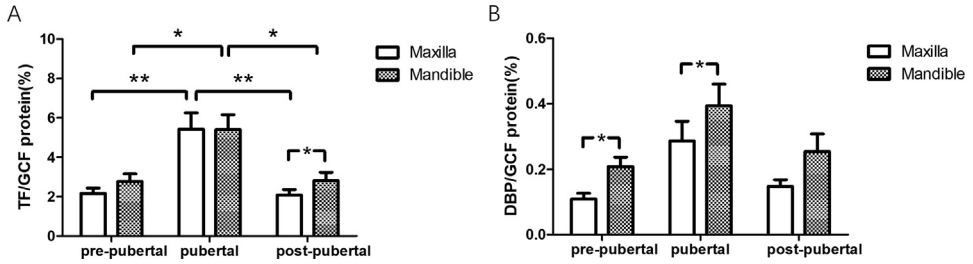


Fig 1. Percentage of TF **(A)** and DBP **(B)** in GCF of subjects at prepubertal, pubertal, and postpubertal stage (*P <0.05; **P <0.01).

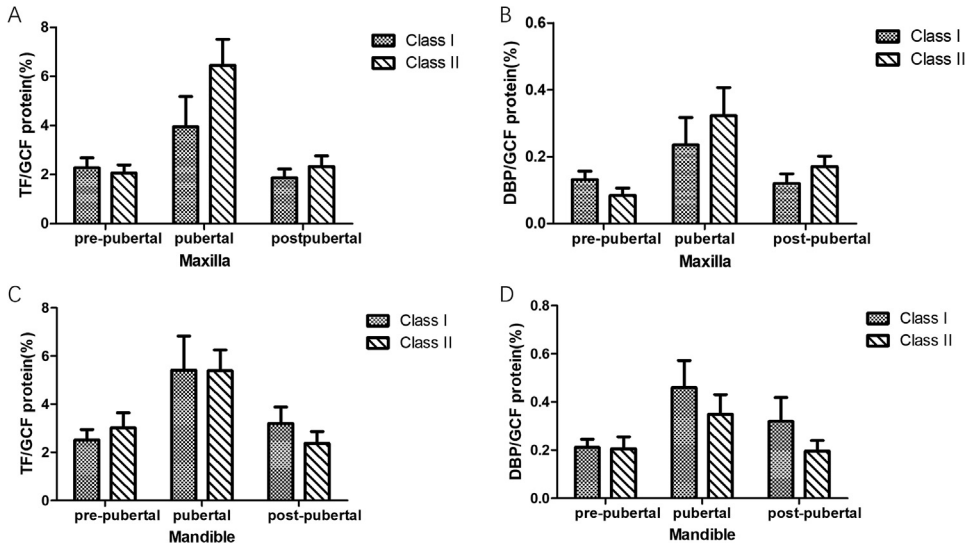


Fig 2. Comparison of TF **(A, C)** and DBP **(B, D)** between subjects with Class I and Class II malocclusion.

predicting pubertal peak was not satisfactory. The AUC of age was also lower than TF. However, after we combined chronological age with GCF biomarkers together as a comprehensive prediction model (Model 6; Fig 4, *J* and *L*), the diagnostic accuracy was greatly improved and ranked the first among all prediction models with AUC up to 0.9.

From this chart, we can also find that the ROC curve, drawn using back substitution method, generally exhibited better diagnostic accuracy results than cross-validation method (Fig 4, *N*).

To determine the most optimal threshold of GCF biomarkers in distinguishing pubertal from nonpubertal subjects, we calculated the Youden index, and the results were showed in Figure 5. Coordinates of the maximal Youden index were labeled, and the x value referred to the most optimal threshold of predictive probability of each LR model. According to the LR model and the coordinates labeled in Figure 5, the most optimal threshold of biomarkers could be reversely calculated. Results showed that for Max-TF and Md-TF, the most optimal threshold distinguishing pubertal from nonpubertal subjects turned out to be 4.20% and 4.09%, respectively (Table III).

The sensitivity and specificity of the most optimal threshold of each model were shown in Table III. It was observed that for Max-TF (Model 1), the specificity was ideal (97.7%), whereas the sensitivity was less satisfying (54.5%). On the contrary, the chronological age (Model 5) presented extraordinary sensitivity (86.3%) but poor specificity (54.6%). Consequently, Model 6 combining Max-TF and age exhibited specificity of 100% and sensitivity of 68.2%.

Special attention should be paid when elucidating the result of Model 6 (Table III). Considering that Model 6 was affected by 2 different variables (Max-TF and age) at the same time, it was more meaningful to calculate the most optimal threshold of the *predictive probability*

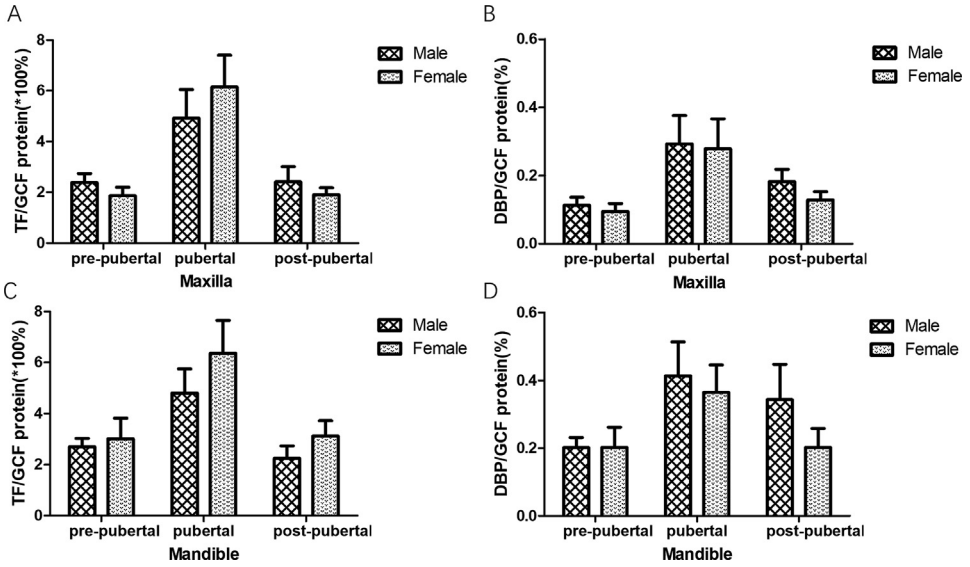


Fig 3. Comparison of TF **(A, C)** and DBP **(B, D)** between males and females.

Table II. LR models using different combinations of independent variables

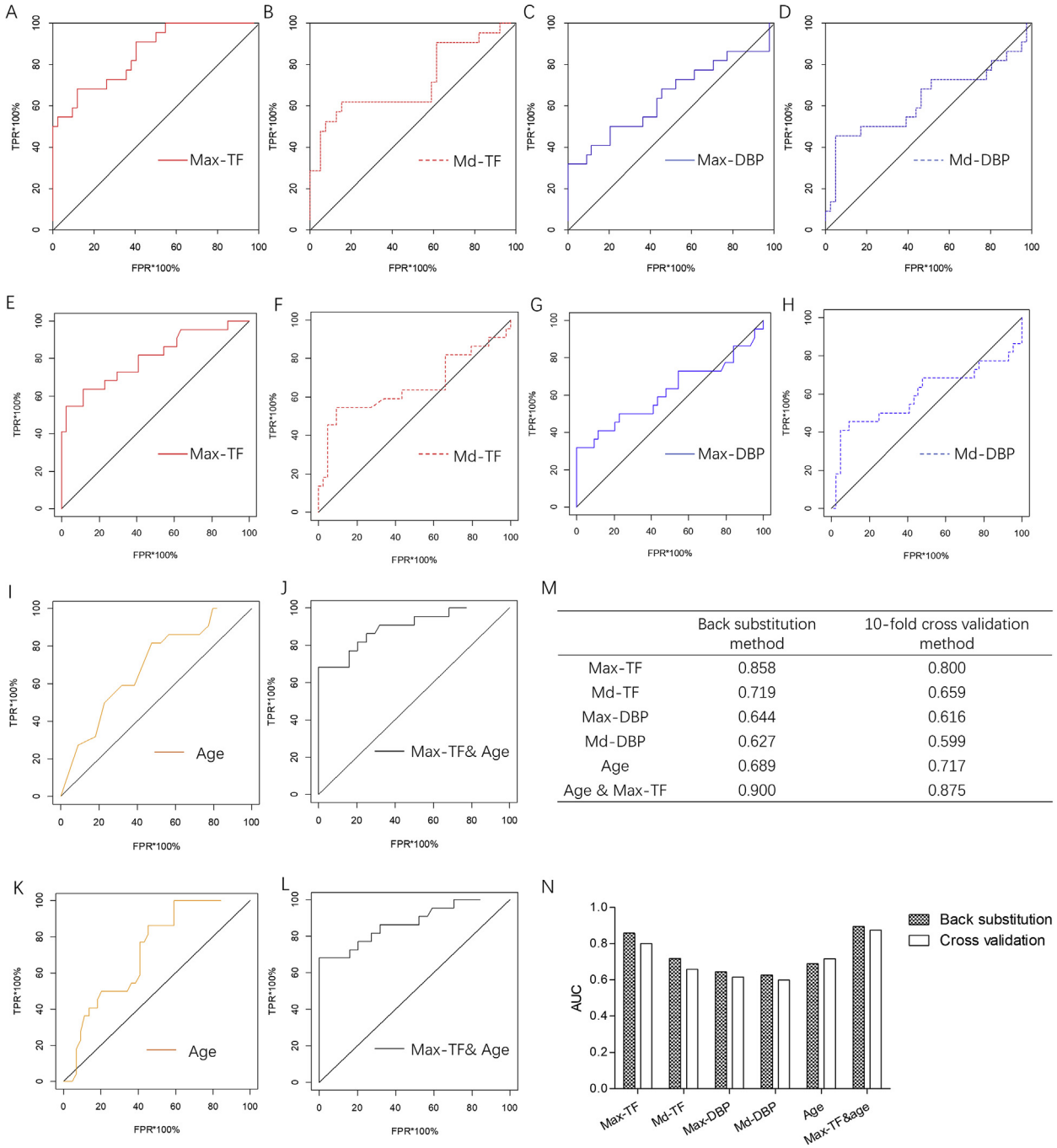
Models	*Variables included*	*Backward stepwise screening*	*Coefficient*	*Standard error*	P *value*
1	Max-TF	Constant	−2.4493	0.5517	0.00000902**
		Max-TF	0.6494	0.1817	0.000351**
2	Md-TF	Constant	−2.2494	0.5829	0.00011**
		Md-TF	0.3974	0.1291	0.00208**
3	Max-DBP	Constant	−1.6244	0.4267	0.00014**
		Max-DBP	5.0421	1.8286	0.00582**
4	Md-DBP	Constant	−1.5497	0.4522	0.00061**
		Md-DBP	2.8417	1.1650	0.01472*
5	Age	Constant	−57.7874	23.4117	0.012*
		Age^2	−0.4059	0.1640	0.013*
		Age	9.7096	3.9271	0.013*
6	Age, sex, skeletal type, Max-TF, Md-TF, Max-DBP, Md-DBP	Constant	−57.2438	26.5475	0.031*
		Age^2	−0.4275	0.3369	0.003**
		Age	9.7342	0.1954	0.028*
		Max-TF	0.9859	4.5526	0.032*

*P <0.05; **P <0.01.

rather than that of either independent variable. Because the predictive probability corresponding to the maximal Youden index was 0.659, the most optimal threshold of this model should be 0.659, with corresponding prediction specificity of 100% and sensitivity of 68.2%.

DISCUSSION

Considering the great impact that pubertal peak detection has on the functional treatment effect in subjects with Class II malocclusion, we paid special attention to the change pattern of TF and DBP in subjects with Class I and Class II malocclusion. Results showed that the percentage of TF in GCF of pubertal subjects was significantly higher than prepubertal and postpubertal subjects, both in maxilla and mandible (Fig 1), which further validate our previous hypothesis that TF can be considered as a candidate biomarker of pubertal growth peak. However, no significant difference was observed in DBP, indicating that the correlation between DBP and pubertal peak might be weaker than we had expected, and a larger sample size may be needed for further validation.



	Back substitution method	10-fold cross validation method
Max-TF	0.858	0.800
Md-TF	0.719	0.659
Max-DBP	0.644	0.616
Md-DBP	0.627	0.599
Age	0.689	0.717
Age & Max-TF	0.900	0.875

Fig 4. ROC curves of Model 1-6 drawn by back substitution method and 10-fold cross-validation method and their corresponding AUC. **A-D, I, J,** ROC curve of Model 1-6 drawn by back substitution method. **E-H, K, L,** ROC curve of Model 1-6 drawn by 10-fold cross-validation method. **M, N,** Descriptive data of the AUC.

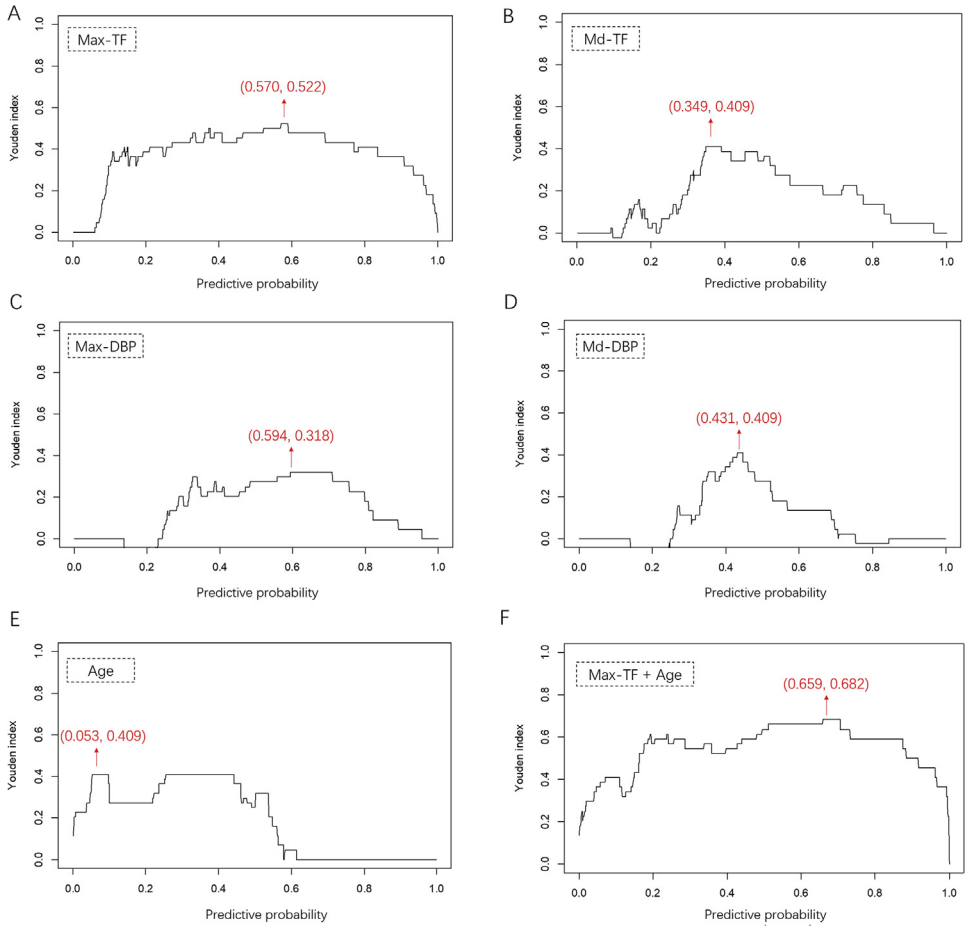


Fig 5. Youden index as a function of the predictive probability threshold and the corresponding coordinates of the maximal Youden index of Model 1-6. **(A)** Youden index of Max-TF; **(B)** Youden index of Md-TF; **(C)** Youden index of Max-DBP; **(D)** Youden index of Md-DBP; **(E)** Youden index of age; **(F)** Youden index of Max-TF combined with age.

Table III. The most optimal threshold and corresponding sensitivity and specificity of 6 LR models

Models	*Significant independent variables*	*Most optimal threshold of GCF biomarkers (%)*	*Most optimal threshold of the predictive probability*	*Sensitivity (%)*	*Specificity (%)*	*Maximal Youden index*
1	Max-TF	4.20	0.570	54.5	97.7	0.522
2	Md-TF	4.09	0.349	54.5	86.4	0.409
3	Max-DBP	0.40	0.594	31.8	100	0.318
4	Md-DBP	0.45	0.431	45.4	95.5	0.409
5	Age	/	0.053	86.3	54.6	0.409
6	Max-TF, Age	/	0.659	68.2	100	0.682

In addition, we found that mandibular GCF biomarkers were higher than maxillary biomarkers at the pubertal stage, especially for DBP (Fig 1, *B*). This phenomenon could be explained by the theory of craniofacial growth that the growth peak of the maxilla was earlier than the mandible. Chen et al[27,28] observed longitudinal lateral cephalograms of 87 subjects with normal occlusion and found that the greatest relative growth rate of maxillary length and height was in quantitative CVM stage I, whereas the greatest relative growth

rate of mandibular length and height was in quantitative CVM stage II. Nielsen et al[29] examined head films taken semiannually from 10 male monkeys and reported that the growth peak of the maxilla was earlier than the mandible, and the growth velocity of the mandible was greater than the maxilla. The CVM method was proposed mainly on the basis of the growth velocity of the mandible rather than growth velocity of the maxilla.[7] Therefore, at the *pubertal stage* defined by the CVM method, mandibular biomarkers were higher than maxillary biomarkers, possibly because of the difference in growth velocity between maxilla and mandible.

Franchi et al[8] and Baccetti et al[30] showed that the greatest increment of mandibular length occurred during the interval from cervical stage 3 (CS3) to CS4. Therefore, we can infer that when the increment of mandibular length comes up to the greatest value, the percentage of GCF biomarkers also reached the peak. Hence, we further hypothesized that GCF biomarkers might be sensitive indicators of growth increment of the jaw and may reflect the subtle difference in mandibular growth between subjects with Class I and Class II malocclusion. Previous studies showed that differences did exist in the growth velocity of the mandible between subjects with Class I and Class II malocclusion during the pubertal stage. Stahl et al[31] analyzed longitudinal lateral cephalograms of 17 subjects with Class I and Class II malocclusion from CS1 to CS6 stage and found that at the pubertal stage (CS3-CS4), the increment of mandibular length in subjects with Class II malocclusion was significantly lower than subjects with Class I relationship. Ngan et al[32] and Baccetti et al[33] also had similar conclusions. Unfortunately, according to our results, neither TF nor DBP showed a significant difference between the Class I and Class II groups. However, the tendency of DBP seemed to approximate our hypothesis. At the pubertal stage, the percentage of Max-DBP was higher in subjects with Class II malocclusion, whereas Md-DBP was lower in subjects with Class II malocclusion than in subjects with Class I malocclusion (Fig 2, *B* and *D*). One of the possible explanations was that individual varieties of GCF biomarkers were greater, and a larger sample size is needed to further validate this assumption.

The difference in growth increment of the jaw between males and females has been commonly recognized. Jacob et al[34] analyzed the longitudinal lateral cephalograms of 130 untreated adolescences aged from 10 to 15 years and found that the increment of mandibular length and ramus height of males were significantly higher than females. Similarly, Ochoa et al[35] collected longitudinal data on 28 subjects aged from 6 to 20 years and concluded that the duration of mandibular growth of males was significantly longer than females, and the increment of mandibular length was significantly greater than females. However, our result did not show any significant difference of TF or DBP in GCF between males and females, probably because of the limited sample size and the individual variety of GCF biomarkers. A previous study has reported the sex difference of total iron pool and serum transferrin-bound iron,[16] but the sex difference of TF in GCF remains unclear. Interestingly, DBP in males tended to be higher than females during the pubertal stage (Fig 3, *B* and *D*), whereas the TF showed the opposite tendency (Fig 3, *A* and *C*). This finding might suggest that different GCF biomarkers could exhibit different patterns of change. However, a larger sample size will be required to test this speculation.

To further evaluate and compare the diagnostic accuracy of TF and DBP, LR analysis was performed. Six models were created using different combinations of independent variables. Models 1-4 included only 1 of candidate GCF biomarkers, namely Max-TF, Md-TF, Max-DBP, and Md-DBP, respectively, to assess their diagnostic accuracy separately. Apart from the CVM method, chronological age was also one of the indexes applied to evaluate the growth stage. Therefore, chronological age was used to construct Model 5, to evaluate its accuracy, and to compare it with GCF biomarkers. Model 6 included all related variables into the model to see if better accuracy could be achieved. After backward stepwise screening for significant variables, only Max-TF and age stayed in the model, which suggested that GCF biomarker Max-TF and age dominated in the diagnosis of pubertal peak.

The ROC curve is used most commonly in medicine as a means of evaluating diagnostic tests.[36] In our study, the ROC curve was used to compare the diagnostic accuracy of each model. AUC was calculated for quantitative comparison. Results showed that the diagnostic accuracy of GCF biomarkers followed the following order: Max-TF > Md-TF > Max-DBP > Md-DBP. TF showed better accuracy than DBP, especially in the maxilla, of which the AUC exceeded 0.8. In contrast, the AUC of age was only 0.689, which confirmed that chronological age was not a reliable indicator of skeletal growth peak.[37,38] In Model 6, combining Max-TF with age, the diagnostic accuracy was improved, and the AUC exceeded 0.9. This finding suggested that GCF biomarkers TF could achieve the best accuracy in combination with age. A similar result was reported by a recent study by Alhazmi et al[12] that the combination of the salivary biomarker alkaline phosphatase activity and chronological age may provide the best prediction of the CVM stage.

Calculation of the Youden index helped to figure out the most optimal thresholds of diagnostic models and their corresponding sensitivity and specificity. Results showed that the most optimal threshold of Max-TF and Md-TF in distinguishing pubertal from nonpubertal subjects was 4.20% and 4.09%, respectively, in which best sensitivity and specificity can be achieved. For Model 6, when the threshold of the predictive probability was set at 0.659, the specificity would reach up to 100%, and the sensitivity would be 68.2%. Taking both sensitivity and specificity into consideration, Model 6 exhibited the most satisfactory diagnostic accuracy. The treatment timing plays a vital role in the orthopedic treatment effect, and either misdiagnose or missing diagnosis would compromise the final result. From this perspective, the model, including both Max-TF and age, was considered to be a relatively satisfactory diagnostic model in distinguishing pubertal from nonpubertal subjects.

GCF serves as a medium transmitting the underlying message of bone development in the oral environment and should be deeply explored for its promising diagnostic value. Indeed, GCF volume is limited compared with saliva, and the experimental protocol in the laboratory is relatively time-consuming at the current stage. However, the collection procedure of GCF is safe and noninvasive and can be repeated several times until enough volume is collected. In the future, we will work on developing a more convenient, rapid, and sensitive approach to replace the paper points method and complex experimental procedure for a better adaptation of clinical usage.

CONCLUSIONS

1. The percentage of TF in GCF was significantly higher in pubertal than in prepubertal and postpubertal subjects, whereas DBP showed a relatively weaker correlation with pubertal peak.
2. Maxillary TF exhibited the best diagnostic accuracy among GCF biomarkers with AUC exceeding 0.8.
3. The most optimal thresholds of maxillary and mandibular TF in distinguishing pubertal from nonpubertal subjects were 4.20% and 4.09%, respectively.
4. Chronological age was not a reliable indicator of pubertal growth peak. However, after combined with maxillary TF, the diagnostic accuracy was greatly improved, and the AUC could reach up to 0.9.
5. For Model 6, combining maxillary TF with age, the most optimal threshold of the predictive probability was 0.659, and the specificity could reach up to 100%.
6. TF in GCF could be considered as a potential biomarker of pubertal peak and could achieve acceptable accuracy as an assistant of the chronological age in the diagnosis of pubertal peak.

SUPPLEMENTARY DATA

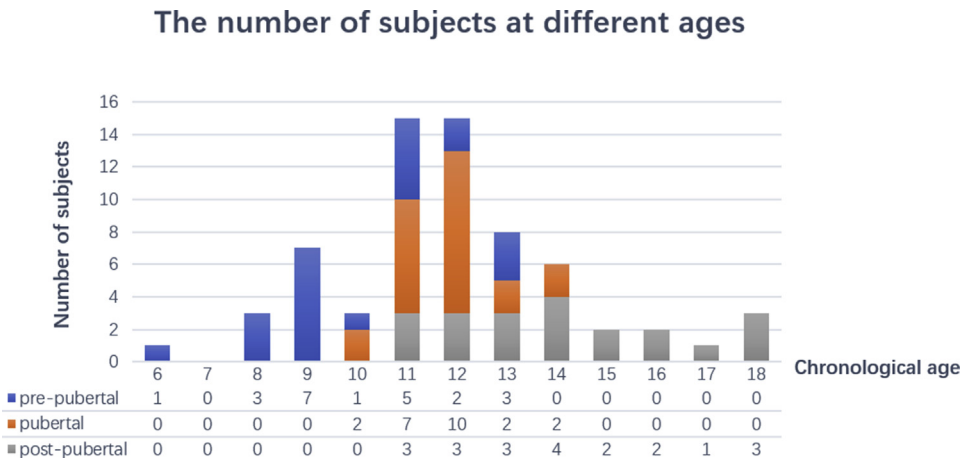
Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ajodo.2020.01.025.

REFERENCES

1. Baccetti T, Franchi L, Toth LR, McNamara JA Jr. Treatment timing for Twin-block therapy. Am J Orthod Dentofacial Orthop 2000;118:159-70.
2. Faltin KJ, Faltin RM, Baccetti T, Franchi L, Ghiozzi B, McNamara JA Jr. Long-term effectiveness and treatment timing for bionator therapy. Angle Orthod 2003;73:221-30.
3. Martina R, Cioffi I, Galeotti A, Tagliaferri R, Cimino R, Michelotti A, et al. Efficacy of the Sander bite-jumping appliance in growing patients with mandibular retrusion: a randomized controlled trial. Orthod Craniofac Res 2013;16:116-26.
4. Uzuner DF, Darendeliler N, Yucel E. Combined fixed-functional treatment of skeletal class II malocclusions with the EVAA appliance: a preliminary study. J Orofac Orthop 2014;75:275-86.
5. Pavoni C, Lombardo EC, Lione R, Faltin K Jr, McNamara JA Jr, Cozza P, et al. Treatment timing for functional jaw orthopaedics followed by fixed appliances: a controlled long-term study. Eur J Orthod 2018;40:430-6.
6. Fishman LS. Radiographic evaluation of skeletal maturation. A clinically oriented method based on hand-wrist films. Angle Orthod 1982;52:88-112.
7. O'Reilly MT, Yanniello GJ. Mandibular growth changes and maturation of cervical vertebrae–a longitudinal cephalometric study. Angle Orthod 1988;58:179-84.
8. Franchi L, Baccetti T, McNamara JA Jr. Mandibular growth as related to cervical vertebral maturation and body height. Am J Orthod Dentofacial Orthop 2000;118:335-40.
9. 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.
10. Ishaq RA, Soliman SA, Foda MY, Fayed MM. Insulin-like growth factor I: a biologic maturation indicator. Am J Orthod Dentofacial Orthop 2012;142:654-61.
11. Szulc P, Seeman E, Delmas PD. Biochemical measurements of bone turnover in children and adolescents. Osteoporos Int 2000;11:281-94.
12. Alhazmi N, Trotman CA, Finkelman M, Hawley D, Zoukhri D, Papathanasiou E. Salivary alkaline phosphatase activity and chronological age as indicators for skeletal maturity. Angle Orthod 2019;89:637-42.
13. Uitto VJ. Gingival crevice fluid–an introduction. Periodontol 2000 2003;31:9-11.
14. Griffiths GS. Formation, collection and significance of gingival crevice fluid. Periodontol 2000 2003;31:32-42.
15. Wen X, Franchi L, Chen F, Gu Y. Proteomic analysis of gingival crevicular fluid for novel biomarkers of pubertal growth peak. Eur J Orthod 2018;40:414-22.
16. Elsayed ME, Sharif MU, Stack AG. Transferrin saturation: a body iron biomarker. Adv Clin Chem 2016;75:71-97.

17. Anttila R, Siimes MA. Serum transferrin and ferritin in pubertal boys: relations to body growth, pubertal stage, erythropoiesis, and iron deficiency. Am J Clin Nutr 1996;63:179-83.
18. Misaki M, Shima T, Yano Y, Okazaki T. Serum transferrin as a marker of bone growth in boys: correlation with serum alkaline phosphatase activity, plasma insulin-like growth factor 1 and rate of growth in height. Horm Metab Res 1991;23:230-2.
19. Malecki EA, Buhl KM, Beard JL, Jacobs CR, Connor JR, Donahue HJ. Bone structural and mechanical properties are affected by hypotransferrinemia but not by iron deficiency in mice. J Bone Miner Res 2000;15:271-7.
20. Anderson PH, Lam NN, Turner AG, Davey RA, Kogawa M, Atkins GJ, et al. The pleiotropic effects of vitamin D in bone. J Steroid Biochem Mol Biol 2013;136:190-4.
21. Nimitphong H, Sritara C, Chailurkit LO, Chanprasertyothin S, Ratanachaiwong W, Sritara P, et al. Relationship of vitamin D status and bone mass according to vitamin D-binding protein genotypes. Nutr J 2015;14:29.
22. Ilich JZ, Badenhop NE, Jelic T, Clairmont AC, Nagode LA, Matkovic V. Calcitriol and bone mass accumulation in females during puberty. Calcif Tissue Int 1997;61:104-9.
23. 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.
24. Chow S, Shao J, Wang H. Sample Size Calculations in Clinical Research. 2nd ed. Boca Raton: Chapman & Hall/CRC Press; 2008.
25. Kong C, Li SM, Yang H, Xiao WD, Cen YY, Wu Y, et al. Screening and combining serum biomarkers to improve their diagnostic performance in the detection of intestinal barrier dysfunction in patients after major abdominal surgery. Ann Transl Med 2019;7:388.
26. Yi J, Ren L, Wu J, Li W, Zheng X, Du G, et al. Apolipoprotein C1 (APOC1) as a novel diagnostic and prognostic biomarker for gastric cancer. Ann Transl Med 2019;7:380.
27. Chen L, Liu J, Xu T, Lin J. Longitudinal study of relative growth rates of the maxilla and the mandible according to quantitative cervical vertebral maturation. Am J Orthod Dentofacial Orthop 2010;137:736.e1-8: discussion 736-7.
28. Chen LL, Xu TM, Jiang JH, Zhang XZ, Lin JX. Quantitative cervical vertebral maturation assessment in adolescents with normal occlusion: a mixed longitudinal study. Am J Orthod Dentofacial Orthop 2008;134:720.e1-7: discussion 720-1.
29. Nielsen IL, Bravo LA, Miller AJ. Normal maxillary and mandibular growth and dentoalveolar development in Macaca mulatta. A longitudinal cephalometric study from 2 to 5 years of age. Am J Orthod Dentofacial Orthop 1989;96:405-15.
30. Baccetti T, Franchi L, McNamara JA Jr. An improved version of the cervical vertebral maturation (CVM) method for the assessment of mandibular growth. Angle Orthod 2002;72:316-23.
31. Stahl F, Baccetti T, Franchi L, McNamara JA Jr. Longitudinal growth changes in untreated subjects with Class II Division 1 malocclusion. Am J Orthod Dentofacial Orthop 2008;134:125-37.
32. Ngan PW, Byczek E, Scheick J. Longitudinal evaluation of growth changes in Class II division 1 subjects. Semin Orthod 1997;3:222-31.
33. Baccetti T, Franchi L, McNamara JA Jr, Tollaro I. Early dentofacial features of Class II malocclusion: a longitudinal study from the deciduous through the mixed dentition. Am J Orthod Dentofacial Orthop 1997;111:502-9.
34. Jacob HB, Buschang PH. Mandibular growth comparisons of Class I and Class II division 1 skeletofacial patterns. Angle Orthod 2014;84:755-61.
35. Ochoa BK, Nanda RS. Comparison of maxillary and mandibular growth. Am J Orthod Dentofacial Orthop 2004;125:148-59.
36. Carter JV, Pan J, Rai SN, Galandiuk S. ROC-ing along: evaluation and interpretation of receiver operating characteristic curves. Surgery 2016;159:1638-45.
37. Alkhal HA, Wong RW, Rabie AB. Correlation between chronological age, cervical vertebral maturation and Fishman's skeletal maturity indicators in southern Chinese. Angle Orthod 2008;78:591-6.
38. Baccetti T, Franchi L, De Toffol L, Ghiozzi B, Cozza P. The diagnostic performance of chronologic age in the assessment of skeletal maturity. Prog Orthod 2006;7:176-88.

SUPPLEMENTARY DATA



Supplementary Fig. The number of subjects at different ages. The subjects recruited in this study were divided according to their chronological ages. The bars with different colors referred to different cervical stages (blue: pre-pubertal stage; yellow: pubertal stage; grey: post-pubertal stage).