Original Article

Computational Formulation of Orthodontic Tooth-Extraction Decisions

Part I: To Extract or Not To Extract

Kenji Takada[a]; Masakazu Yagi[b]; Eriko Horiguchi[c]

ABSTRACT

Objective: To develop a mathematical model that simulates whether or not to extract teeth in optimizing orthodontic treatment outcome and to formulate the morphologic traits sensitive to optimizing the tooth-extraction/nonextraction decisions.

Materials and Methods: A total of 188 conventional orthodontic records of patients with good treatment outcomes were collected, and dentofacial morphologic traits, along with their degrees of influence in the optimized model, were determined.

Results: The rate of coincidence between the recommendations given by the optimized model and the actual treatments performed was found to be 90.4%. The major morphologic traits and their corresponding influences in improving the simulation accuracy of the model were the incisor overjet (3.0) and the size of the basal arch relative to the sum of the mesiodistal crown diameters of the upper dentition (2.4) and the lower dentition (2.0). The remaining 22 morphologic-trait variables were also found to be indispensable in achieving robust simulation readings.

Conclusion: A mathematical model that simulates whether or not to extract teeth in optimizing orthodontic treatment outcomes with a success rate of 90.4% at its prediction performance was developed. This model has 25 morphologic traits with four major categories (sagittal dentoskeletal and soft tissue relationship, vertical dentoskeletal relationship, transverse dental relationship, and intra-arch conditions) that affected the accuracy in determining optimal tooth extractions/nonextractions. (*Angle Orthod.* 2009;79:885–891.)

KEY WORDS: Tooth; Extraction; Simulation; Modeling; Orthodontics

INTRODUCTION

Despite the popularity of extracting teeth in orthodontic practice, there are no objective standards to be used by orthodontists to decide whether to extract or not to extract teeth. During the past century, clinical experiences that were thought to be useful in explaining criteria for orthodontic tooth extraction have been stocked, and nowadays they are used in orthodontic education and practice as the general knowledge. This knowledge, however, has mostly comprised descriptions based on an individual practitioner's often-fragmented experiences, and thus is unlikely to systematically provide a rationale basis in choosing either extraction or nonextraction of teeth. In fact, this is one of the reasons why there is a considerable degree of discordance in opinions regarding the judgments on tooth extractions delivered by orthodontists.[1,2]

However, in the past two decades, the development of orthodontic-treatment planning systems has been pioneered.[3,4] Diagnosis and treatment planning is a computational procedure that compresses and models knowledge, including scientific findings and personal clinical experiences held by medical experts, into a finite number of discrete feature elements, and finally transforms them into mathematical descriptions. Pattern-matching techniques have been confirmed to be quite effective in clinical use[5,6] by visualizing and modeling the many knowledge elements (experiential knowledge and empirically learned rules) previously usd by orthodontic experts as aids in their judgments. To date, however, there is no developed mathematical model that automatically simulates the orthodontic

[a] Professor and Chair, Department of Orthodontics and Dentofacial Orthopedics, Graduate School of Dentistry, Osaka University, Suita, Japan.

[b] Associate Professor, The Center for Advanced Medical Engineering and Informatics, Osaka University, Suita, Japan.

[c] Private practice, Minoh, Japan.

Corresponding author: Dr Kenji Takada, Osaka University, 1-8 Yamadaoka, Suita, Japan 565-0871
(e-mail: opam@dent.osaka-u.ac.jp)

Accepted: November 2008. Submitted: August 2008.

tooth-extraction/nonextraction decisions that would logically lead to a guaranteed optimum treatment outcome. If the degrees of influence of the knowledge elements usd by orthodontists in deciding extraction/nonextraction of teeth were quantified, it would help practitioners and their patients understand that the rationale used is derived from a logical sequence for achieving optimum treatment outcome, whether it means extracting teeth or not.

The purposes of Part I of the present study are (1) to develop a mathematical model that simulates experts' decisions of whether or not to extract a tooth/teeth based on pretreatment conditions of orthodontic patients and (2) to examine what kinds of knowledge elements are sensitive in formulating orthodontic tooth-extraction/nonextraction decisions.

MATERIALS AND METHODS

One hundred and eighty-eight women (mean age, 17 years 5 months; age range, 12 years 1 month to 36 years 0 months) who had achieved good orthodontic treatment outcomes at the university's dental hospital were selected according to order of registration. Medical charts, dental casts, and lateral and posteroanterior head films of each subject were used. All subjects had full permanent dentitions except the third molar teeth, had no abnormalities of the craniofacial forms or skeletal deformities, or had no history of surgical orthodontic treatment. Good treatment outcome was defined as a posttreatment condition with a greater than 70% decrease in the peer assessment rating (PAR) index score[7] (compared with the pretreatment state). The PAR index scores were measured for the pretreatment records of the subjects in the extraction group (mean, 31.2; range, 11–60) and the nonextraction group (mean, 26.0; range, 8–55).

Twenty-seven feature variables, that is, elements that characterize orthodontic problems that were assumed to be important in deciding whether or not to extract teeth were selected subjectively on the basis of knowledge and clinical experiences held by the authors and measured on the pretreatment records. Definitions of the variables are provided in Table 1.[8–12] Transformation of the functions into nonlinear forms made with reference to the normative values[13] and the authors' expertise knowledge were then applied to the variables. To generate feature vectors, 1000 combinations of feature variables were heuristically.[6] A feature vector is represented by a set of multiple feature variables as mentioned previously. For each case, a feature vector and the actual treatment recorded in the medical charts were paired to generate a knowledge data set.

Architecture of the model is illustrated in Figure 1. The knowledge data sets were stored as templates in the modeling system. Feature vectors were generated from the pretreatment records of input cases, and the top N_m templates nearest to the input feature vectors ($N_m = 1, 3, 5, 7, 9, 11, 13$) were searched mathematically in the system. The optimum decisions of whether a tooth/teeth should be extracted were predicted by applying majority voting to the selected templates. Weighting coefficients (*W*s) having values ranging from 0.1 to 3.5 with a resolution of 0.1 were given to each feature variable. This resulted in 8000 ways with varying amount of *W*s. The range and resolution of *W*s were determined in a preliminary experiment.

Each of the 188 knowledge data sets was used as an input to the model, the remaining knowledge data sets were used as templates, and the performance accuracy of the model was evaluated.[14] If the model's recommendation and the actual treatment coincided, the case was labeled "coincided." The ratio of the number of coincided cases with respect to the entire number of cases was computed as the rate of coincidence (ROC). The model with the highest ROC was selected as the optimized model, and the knowledge elements that influenced tooth extraction/nonextraction decisions were thus determined.

For the cases that were not deemed coincided, three orthodontists (two men and one woman) who had clinical experiences longer than 8 years separately examined the records, except the medical charts, and judged whether they could be susceptible to tooth extraction. The cases in which the three judges shared agreements were assigned to the "typical" set of cases. When the three judges did not agree, the cases were assigned to the "borderline" set of cases. The software program (MATLAB, MathWorks, Natick, MA) performed the model simulations.

RESULTS

The ROC of the optimized model was 90.4% in case of $N_m = 7$. The feature vector elements adopted in the optimized model and their corresponding weighting coefficients are shown in Table 2. Twenty-five elements were used in the optimized model and classified into the four major categories, that is, the sagittal dentoskeletal and soft tissue relationship, the vertical dentoskeletal relationship, the transverse dental relationship, and the intra-arch conditions. The 18 cases that showed disagreement between the system's recommendations and the treatments actually done are shown in Table 3. The numbers of the borderline cases, the typical extraction cases, and the typical nonextraction cases were 11, 7, and 0, respectively.

Table 1. Definitions of Feature Variables Used in the Present Study

Variable	Definition
Sagittal relationship	
Sk2 and Sk3	Membership grades that designate the severity of skeletal Class II and Class III jaw relationships, respectively, as defined by the (TSS) Takada-Sorihashi-Stephens analysis.[8]
FMIA (°)	Lower central incisor angulation to the mandibular plane.
$\underline{1}$ to NA and $\overline{1}$ to NB[9] (mm)	Distances from the upper and lower incisor tips to the N-A line and the N-B line, respectively; protrusiveness of the upper central incisors.
L1a-Cli and L1a-Cla[10] (mm)	Distances from lower incisor apex (L1a) to Cli and Cla, which are the points on the trajectory of the hypothetical tipping movement of the mandibular central incisor root around the center of rotation and on the inner contour of the posterior and anterior cortical plate, respectively.
EL-ls and EL-li[11] (mm)	Upper and lower lip protrusions; distances from upper lip to E-line and from lower lip to E-line, respectively.
OJ (mm)	Distance from the upper central incisor tip to a plane tangential to the lower incisor labial surface and parallel to the occlusal plane.
Molar-R and Molar-L (mm)	Distance from the upper to the lower first molar teeth in the maximal intercuspal position on the right and left sides. When the lower first molar tooth was located mesially, a negative value was assigned.
Vertical relationship	
FMA (°)	Mandibular plane angle.
ALFH (%)	Lower anterior face height; The ratio of the Me to the palatal plane distance with the N to the palatal plane distance.
OB (mm)	Distance from the upper central incisor tip to the lower central incisor tip and perpendicular to the occlusal plane.
Transverse relationship	
MDL_U and MDL_L (mm)	Upper and lower dental midline deviations from the facial midline. When the variable took a negative value, the dental midline was assumed to shift left from the facial midline.
Intra-arch conditions	
II_U and II_L (mm)	Modified irregularity indices[12] for the upper and lower dentitions, respectively. Sum of the linear distances from an anatomic contact point to its adjacent anatomic contact point between the first molar tooth on one side and in each dental arch on the opposite side.
SCD_U and SCD_L (mm)	Tooth materials, that is, the sum of the mesiodistal crown diameters of the 12 teeth (SCD)—the incisors, cuspids, bicuspids, and first molars in the upper and lower dentitions, respectively.
CAL_U/SCD_U and CAL_L/SCD_L	The ratio of the first bicuspid coronal arch length (CAL) to SCD in the upper and lower dentitions, respectively.
CAW_U/SCD_U and CAW_L/SCD_L	The ratio of upper first bicuspid coronal arch width (CAW) to SCD in the upper and lower dentitions, respectively.
BAL_U*BAW_U/SCD_U and BAL_L*BAW_L/SCD_L	The ratio of the product of first bicuspid basal arch length (BAL) and first bicuspid basal arch width (BAW) to SCD in the upper and lower dentitions, respectively.

DISCUSSION

In a previous study,[7] malocclusions were judged as ''having been improved'' if there was a posttreatment reduction of more than 30% in the PAR index values compared with the pretreatment state. In the present study, we used a more stringent criterion in order to develop a highly accurate simulation performance model and, thus, to guarantee good treatment outcome, a posttreatment reduction of more than 70% was required.

When orthodontists judge whether or not to perform tooth extraction, the intraexaminer reproducibility and the interexaminer agreement are known to range between 80% and 98%, and 50% to 90%, respectively,[1] showing a considerable diversity. This prompted us to formulate the elements that affect orthodontic experts' decision-making for achieving optimum treatment outcome.

The rate of coincidence between the current optimized model's recommendations and the actual treatment was 90.4%. The judgment criteria our model uses reflect the nature of the samples whose treatments have been managed by 77 orthodontists from diagnosis to completion and whose treatment out-

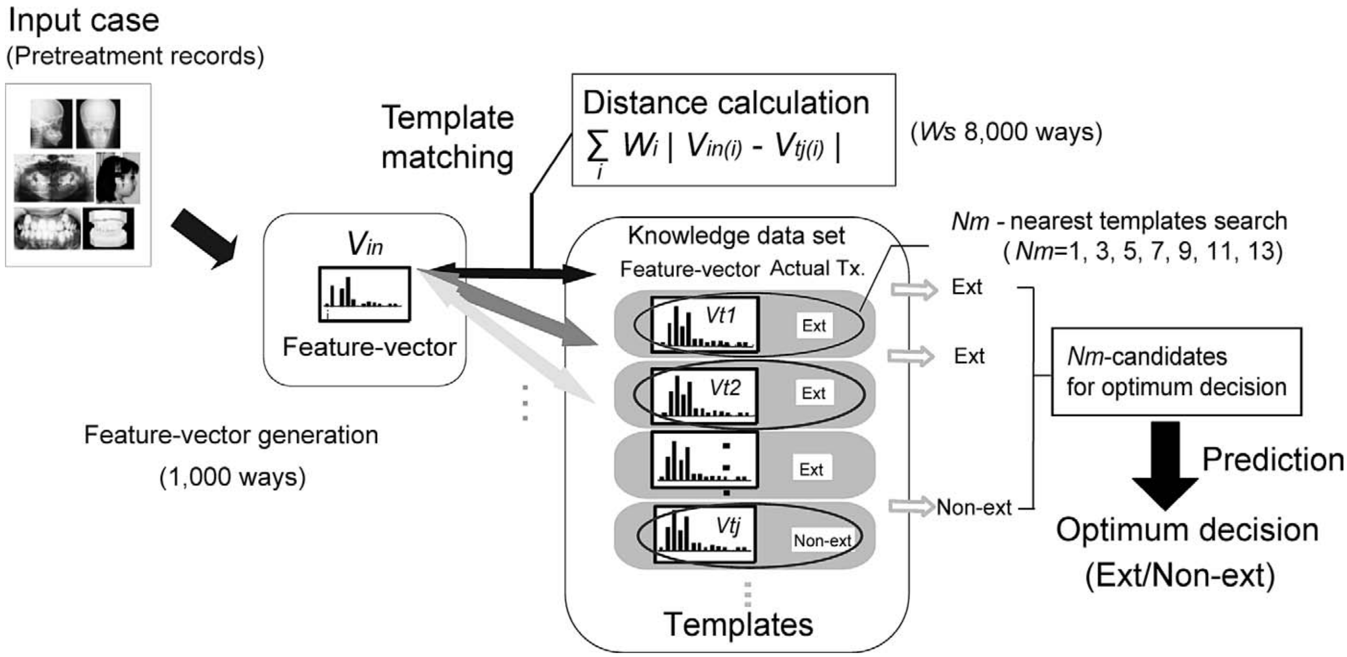

Figure 1. Architecture of the current model. A feature vector was generated from the pretreatment records of an input case, and the optimum decision of whether or not to extract teeth was predicted by means of a template-matching technique with nearest neighbor search (*Nm*).

Table 2. Feature Variables and the Corresponding Weighting Coefficients Adopted in the Optimized Model

Variable	Weighting Coefficient
Sagittal relationship	
OJ	3.0
Molar-R	1.6
Molar-L	1.6
EL-ls	1.3
$\overline{1}$ to NB	1.3
$\underline{1}$ to NA	1.1
FMIA	1.0
Sk3	0.8
Sk2	0.5
EL-li	0.4
Vertical relationship	
ALFH	1.0
OB	1.0
FMA	0.6
Transverse relationship	
MDL_U	1.1
MDL_L	1.1
Intra-arch conditions	
BAL_U*BAW_U/SCD_U	2.4
BAL_L*BAW_L/SCD_L	2.0
II_U	1.5
CAW_U/SCD_U	1.2
CAW_L/SCD_L	1.2
CAL_U/SCD_U	1.1
CAL_L/SCD_L	1.1
II_L	1.1
SCD_U	1.0
SCD_L	1.0

comes had been confirmed objectively and quantitatively as highly successful according to the PAR index scores.

The feature variables adopted in the optimized model were classified into four major subcategories: the sagittal dentoskeletal and soft tissue relationship, the vertical dentoskeletal relationship, the transverse dental relationship, and the intra-arch conditions.

With the sagittal relationship, the overjet was the highest weighting coefficient at 3.0. In other words, the evaluation of the sagittal position of the upper incisors and the possibility of its improvement, that is, the esthetic factor of the face and its solution, was found to be a most important key element in properly simulating whether to or not to extract teeth in the model. The subsequent vector elements were Molar-R and Molar-L, both representing the molar relationships in the sagittal direction. Also, it should be noted that achieving Class I molar relationship at the final stage of the active orthodontic treatment can lead to establishing tight tooth intercuspation of the full dentition and functional occlusal stability. The results of this study suggest that the sagittal molar relationships, because of their influence in achieving improvements of occlusal function and facial proportion, must be taken into account in the orthodontists' decision of tooth-extraction/nonextraction of teeth.

The current model also adopted the overbite and the midline deviation elements that represent the structural traits in the vertical and transverse relationships. Research shows that tooth extraction likely increases

Table 3. Distribution of the Cases that were not Coincided According to Each Case Category

Factual Treatment	Recommendation by the Present Model	Case Category Defined by Experts' Judgment	Number of Cases
Extraction	Nonextraction	Borderline case	7
		Typical extraction case	7
		Typical nonextraction case	0
Nonextraction	Extraction	Borderline case	4
		Typical extraction case	0
		Typical nonextraction case	0

incisor overbite,[15] and this is why tooth extraction is not preferred in patients with a deep overbite; it is recommended in patients exhibiting a decreased incisor overbite because it will increase, or at least not decrease, the pretreatment overbite. Our results also revealed that the vertical incisor relationship contributes to orthodontists' decision making when considering the extraction matter. In regard to the upper and lower dental midline deviations, they are not merely esthetic problems, but they also disturb the establishment of the tight intercuspation of teeth bilaterally. If there is a midline deviation, extraction of a tooth on the nondeviated side in the affected dental arch, or of teeth on the deviated side in both dental arches, is recommended.[16] In agreement with this, our results revealed that those feature elements representing the midline discrepancy were also sensitive in an accurate modeling of tooth-extraction decisions.

Concerning the intra-arch conditions, size of the basal arch relative to the sum of the mesiodistal crown diameters was found to be effective in increasing the ROC, next to the incisor overjet. It should be noted that weights for these feature variables were greater than those of the irregularity indices that designate the severity of tooth crowding. The latter gives a linear summation of the deviations between neighboring teeth, and the former (basal arch variables) serves as an important clue to practitioners in predicting if the apical base, that is, the skeletal frame, can provide enough space for the teeth to be orthodontically moved in; expansion of dental arches as a way to correct tooth crowding may not necessarily guarantee a stable posttreatment occlusion because the direction and distance of tooth movement largely depend on the size of the apical base.[17–19] Hence, the results of the present study indicate that, at a sight where tooth extraction/nonextraction is a possibility, orthodontists are more likely to question whether stable occlusion can be established after teeth realignment without extracting any teeth. The severity of tooth crowding is known to be relevant to the size of the apical base relative to the sum of the mesiodistal crown diameters.[20] The results of the present study reveal that recording the severity of crowding in the upper dentition increases the model simulation accuracy subsequent to the sagittal molar relationships. On the contrary, the severity of crowding in the lower dentition was not as influential as in the upper dentition. These results mean that in deciding whether or not to extract teeth, orthodontists give more concern to the severity of tooth crowding of the upper dental arch rather than its opponent. It would be logical to assume that the first reason orthodontists focus on the crowding condition in the maxillary dentition is because of the greater visibility of the upper incisors compared with the lower antagonists. One of the practical measures used to gauge the severity of the tooth crowding condition is arch length discrepancy.[21,22] The measurement reproducibility of the arch length discrepancy, however, is not robust.[22] The prediction accuracy of the model depends on the reproducibility in measuring the feature vector elements; accordingly, we used the modified irregularity indices, which show less measurement errors[12] as feature elements representing tooth-crowding conditions. In the present study, 7 of the 11 borderline cases had received extraction of teeth but were simulated as being nonextraction cases; for the remaining four borderline cases, actual treatments were completed without extracting teeth, but the model recommended those patients as being in need of extraction. Theoretically, if the treatment should proceed without extracting teeth in patients who were diagnosed as typically in need of tooth extraction, it is anticipated that a good alignment of teeth, stable occlusal state, and beautiful facial proportion would likely be difficult to achieve. In contrast, if orthodontic treatment is done with the extraction of teeth in patients who were judged as being typical nonextraction cases, treatment period may likely be prolonged because of the time required for tooth movement and closure of the extracted space. Previous studies,[23,24] however, did not find any significant difference between the treatment outcomes that were achieved with and without the extraction of teeth, because when dealing with borderline cases, orthodontists hesitate to judge whether to extract or not to extract. Hence, we consider the model's recommendations given for the borderline cases appropriate. Regarding the seven typical extraction cases, however, the treatments were actually done with the extraction of teeth, but the model recommended nonextraction.

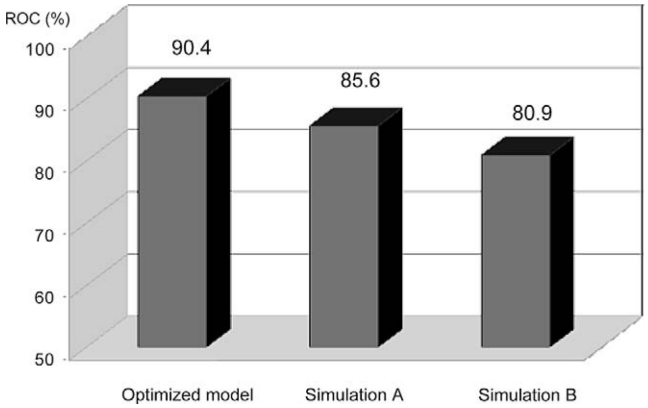

Figure 2. Simulation results for the model with conditions of weights on the feature vector elements different from those adopted in the optimized model. Simulation results for the model with conditions of weights on the feature vector elements different from those adopted in the optimized model. Simulation A: all the elements in the optimized model had equal weight of 1.0. Simulation B: the elements having the weights of greater than 1.5 in the optimized model were employed in the simulation model.

We consider this to be entirely the system's misjudgments. In orthodontics, it is sometimes desirable to start the treatment without extracting teeth tentatively in order to reduce the possible risk of an irreversible treatment, that is, tooth extraction. Based on this, the fact that there was no instance in the current typical nonextraction cases simulated by our system as being extraction cases, indicates the eminence of the present model in avoiding risks.

Using the clinical facts to solve the inverse problem developed the current optimized model. Hence, we applied the model to solve direct problems, that is, simulations of experts' thoughts based on possible choices of feature elements from those adopted in the optimized model and of different conditions of weighting coefficients. As shown in Figure 2, the ROC in the simulation with the model having the same feature vector elements adopted in the optimized model, but with equal weights, was decreased to 85.6%. The simulation model that had only those elements adopted in the optimized model with weights greater than 1.5 showed the even more degraded ROC of 80.9%. These results thus indicate that though there are major feature elements that contribute to experts' judgments, optimum decisions that guarantee the excellent orthodontic treatment outcomes can only be made through experts' elaborate thinking about the multiple elements used in the present study holistically and simultaneously.

Finally, though the current model may not yet suffice to achieve complete agreement with human judgments, it should be noted that it has an advantage in that the system can independently improve its prediction accuracy by adding new patient records as templates just as orthodontists might increase their clinical knowledge and experiences. This means the model will become robust clinically for making decisions for individual patient treatment. The model would be useful to test in different ethnic populations, in populations treated by different orthodontists, or in decisions made in different time periods to study how decision making varies.

CONCLUSIONS

- A mathematical model was developed that can simulate experts' decisions regarding whether or not to extract teeth based on pretreatment conditions of orthodontic patients with a success rate of 90.4% at its prediction performance.
- Twenty-five elements with four major subcategories—sagittal dentoskeletal and soft tissue relationship, vertical dentoskeletal relationship, transverse dental relationship, and intra-arch conditions, were found to be susceptible in optimizing the model.

ACKNOWLEDGMENT

Grant partly supported by the Global COE Program, ''a center of excellence for an *in silico* medicine–oriented worldwide open platform,'' by the Government of Japan.

REFERENCES

1. Ribarevski R. Consistency of orthodontic extraction decisions. *Eur J Orthod.* 1996;18:77–80.
2. Burrow SJ. To extract or not to extract: a diagnostic decision, not a marketing decision. *Am J Orthod Dentofacial Orthop.* 2008;133:341–342.
3. Sims-Williams IH, Brown ID, Matthewaman A. A computer-controlled expert system for orthodontic advice. *Br Dent J.* 1987;163:161–166.
4. Stephens C, Mackin N. The validation of an orthodontic expert system rule-base for fixed appliance treatment planning. *Eur J Orthod.* 1998;20:569–578.
5. Fujita K, Takada K, QianRong G, Shibata T. Patterning of human dental arch wire blanks using a vector quantization algorithm. *Angle Orthod.* 2002;72:285–294.
6. Tanikawa C, Kakiuchi Y, Yagi M, Miyata K, Takada K. Knowledge-dependent pattern classification of human nasal profiles. *Angle Orthod.* 2007;77:821–830.
7. Richmond S, Shaw WC, Roberts CT, Andrews M. The PAR Index (Peer Assessment Rating): methods to determine outcome of orthodontic treatment in terms of improvement and standards. *Eur J Orthod.* 1992;14:180–187.
8. Sorihashi Y, Stephens CD, Takada K. An inference modeling of human visual judgement of sagittal jaw-base relationships based on cephalometry: Part II. *Am J Orthod Dentofacial Orthop.* 2000;117:303–311.
9. Steiner CC, Bilis B. Cephalometrics for you and me. *Am J Orthod.* 1953;39:729–755.
10. Yamada C, Kitai N, Kakimoto N, Murakami S, Furukawa S, Takada K. Spatial relationships between the mandibular central incisor and associated alveolar bone in adults with mandibular prognathism. *Angle Orthod.* 2007;77:766–772.
11. Ricketts RM. Esthetics, environment, and the law of lip relation. *Am J Orthod.* 1968;54:272–289.

12. Little MR. The Irregularity Index: a quantitative score of mandibular anterior alignment. *Am J Orthod.* 1975;68:554–563.
13. Wada K. A study on the individual growth of maxillofacial skeleton by means of lateral cephalometric roentgenograms. *J Osaka Univ Dent Soc.* 1977;22:239–269.
14. Goutte C. Note on free lunches and cross-validation. *Nural Computation.* 1997;9:1245–1249.
15. Nance HN. The removal of second premolars in orthodontic treatment. *Am J Orthod.* 1949;35:685–695.
16. Nanda R, Margolis MI. Treatment strategies for midline discrepancies. *Semin Orthod.* 1996;2:84–89.
17. Case CS. The question of extraction in orthodontia. National Dental Association. Cleveland, Ohio, July 15; reprinted in *Am J Orthod.* 1964;50:656–911.
18. Lundstrom AF. Malocclusion of the teeth regarded as a problem in connection with the apical base. *J Oralsur Rad.* 1925;7:591–602.
19. Zentner A, Peylo S, Brothag D. Predictive value of morphologic parameters for successful correction of Class II Division 2 malocclusion. *Eur J Orthod.* 2003;123:279–285.
20. Norderval K, Wisth PJ, Boe OE. Mandibular anterior crowding in relation to tooth size and craniofacial morphology. *Scand J Dent Res.* 1975;83:267–273.
21. Bishara SE, Jakobsen JR, Treder JE, Stasi MJ. Changes in the maxillary and mandibular tooth size arch length relationship from early adolescence to early adulthood. *Am J Orthod Dentofacial Orthop.* 1989;95:46–59.
22. Rudge SJ, Jones PT, Hepenstal S, Bowden DEJ. The reliability of study model measurement in the evaluation of crowding. *Eur J Orthod.* 1983;5:225–231.
23. Zierhut CE, Joondeph RD, Artun J, Little MR. Long-term profile changes associated with successfully treated extraction and nonextraction class II division 1 malocclusions. *Angle Orthod.* 2000;70:208–219.
24. Xu TM, Liu Y, Yang MZ, Huang W. Comparison of extraction versus nonextraction orthodontic treatment outcomes for borderline Chinese patients. *Am J Orthod Dentofacial Orthop.* 2006;129:672–677.