Original Article

Assessment of malalignment factors related to Invisalign treatment time aided by automated imaging processes

Sanghee Lee[a]; Tai-Hsien Wu[b]; Toru Deguchi[c]; Ai Ni[d]; Wei-En Lu[e]; Sumeet Minhas[a]; Shaun Murphy[a]; Ching-Chang Ko[f]

ABSTRACT

Objectives: To identify predictors regarding the type and severity of malocclusion that affect total Invisalign treatment duration based on an intraoral digital scan.

Materials and Methods: The subjects of this retrospective clinical cohort were 116 patients treated with Invisalign. A deep learning method was used for automated tooth segmentation and landmark identification of the initial and final digital models. The changes in the six degrees of freedom (DOF), representing types of malalignment, were measured. Linear regression was performed to find the contributing factors associated with treatment time. In addition, the Peer Assessment Rating (PAR) score and a composite score combining 6 DOF were correlated separately to the treatment time.

Results: The number of trays differed between sexes ($P = .0015$). The absolute maximum torque was marginally associated with the total number of trays ($P = .0518$), while the rest of the orthodontic tooth movement showed no correlation. The composite score showed a higher correlation with the total number of trays ($P = .0045$) than did individual tooth movement. Pretreatment upper and lower anterior segment PAR scores were positively associated with the treatment time ($P < .001$).

Conclusions: There is not enough evidence to conclude that certain types of tooth movement affect the total aligner treatment time. A composite score seems to be a better predictor for total treatment time than do individual malalignment factors in aligner treatment. Upper and lower anterior malalignment factors have a significant effect on the treatment duration. (*Angle Orthod.* 2023;93:144–150.)

KEY WORDS: Invisalign; Aligner; Artificial intelligence; Deep learning

INTRODUCTION

Approximately 6 million Americans undergo orthodontic treatment every year. A growing proportion of this group comprises adult patients seeking esthetic treatment with clear plastic aligners. According to Shalish et al.,[1] patients treated with aligners claimed a higher health-related quality of life compared to those treated with labial, lingual orthodontic appliances in terms of patient adaptability.

Previous studies[2,3] reported that certain types of orthodontic tooth movements were difficult to accomplish with clear aligners, and the effectiveness of extrusion and rotation were inferior to those associated with other types of tooth movement. However, methodologies incorporated to calculate how much planned movement has been achieved are not as reliable and accurate as measurements based on a constructed three-dimensional coordinate system.[4–9]

With regard to efficiency, it is controversial to conclude that aligners can achieve their goals in a

[a] Resident, Division of Orthodontics, The Ohio State University College of Dentistry, Columbus, Ohio.

[b] Postdoctoral Scholar, Division of Orthodontics, The Ohio State University College of Dentistry, Columbus, Ohio.

[c] Professor and Graduate Program Director, Division of Orthodontics, The Ohio State University College of Dentistry, Columbus, Ohio.

[d] Professor, Division of Biostatistics, The Ohio State University College of Dentistry, Columbus, Ohio.

[e] PhD Candidate, Division of Biostatistics, The Ohio State University College of Dentistry, Columbus, Ohio.

[f] Professor and Chair, Division of Orthodontics, The Ohio State University College of Dentistry, Columbus, Ohio.

Corresponding author: Dr Ching-Chang Ko, Division of Orthodontics, College of Dentistry, 4088 Postle Hall, The Ohio State University, 305 W 12th Ave, Columbus, OH 43210-1267 (e-mail: ko.367@osu.edu)

Accepted: September 2022. Submitted: March 2022.
Published Online: November 3, 2022

 DOI: 10.2319/031622-225.1

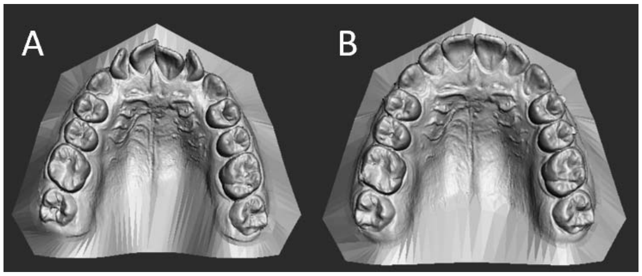

Figure 1. Pretreatment (A) and posttreatment (B) intraoral digital scan.

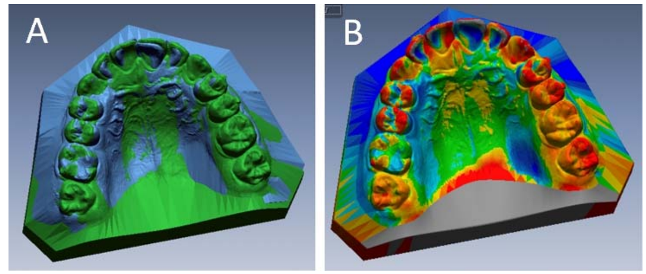

Figure 2. Three-dimensional superimposition (A) and color map (B) image.

short time period. Gu et al.[10] found that aligner treatment requires 5.7 months less treatment time than does a conventional appliance in similar cases, controlling for initial malocclusion measure in PAR (Peer Assessment Rating) score in 96 patients. On the other hand, Lin et al.[11] demonstrated that the conventional appliance group finished treatment 4.8 months earlier than did the aligner group, while the treatment outcomes rated by the American Board of Orthodontics Objective Grading System scores were not significantly different. However, neither study compared the types and amount of tooth movement achieved.

To determine factors that cause an extension of treatment duration, analysis regarding changes in individual misalignment types (eg, in-out, vertical, mesiodistal, rotation, tipping, and torque) with increased sample sizes is necessary. However, these measures are labor-intensive and can easily suffer from evaluator recognition errors. Automated methods using deep learning technology, a subset of artificial intelligence, were developed to reduce human error and save time.[12] This approach allows precise assessment of each tooth movement in a three-dimensional manner and substantial numbers of measurements, a measurement strategy that is impossible to accomplish using manual methods.

This study aimed to identify how the following affected total aligner treatment duration based on digital scan data: the type of alignment strategy and the degree of malalignment severity. The null hypothesis was that there is no correlation between the characteristic of malalignment and the length of aligner treatment time.

MATERIALS AND METHODS

The protocol for this study was reviewed and approved by The Ohio State University Institutional Review Board (No. 2020H0459).

This retrospective clinical cohort included 116 consecutive patients treated with the Invisalign appliance at The Ohio State University graduate orthodontic clinic, with treatment finished between 2016 and 2019. The treatment plan for these subjects was determined by one faculty member. Inclusion criteria for recruiting samples were patients who were 16 years of age or older with full permanent dentition with crowding and spacing problems, patients with Class I molar relationships, and patients with pre- and posttreatment records of digital models and photos. Subjects using additional appliances other than elastics, extraction cases, orthognathic surgery cases, syndromic cases, and patients with early termination of treatment due to poor compliance were excluded.

Pre- and posttreatment intraoral digital models scanned by Trios (3Shape, Copenhagen, Denmark) were collected. An orthodontic resident who was trained to be the calibrated investigator using 50 randomized models (following Richmond et al.[13] guidelines) calculated the PAR index of all digital intraoral scanned models using the OrthoAnalyzer PAR scoring module. The linear displacement of adjacent anatomic contact points of upper incisors and canines was determined in order to measure the pre- and posttreatment anterior PAR indices. The summation of the five measurements between mesial of the canine on one side to the contralateral canine represented the upper and lower anterior irregularity values. All cases selected in this study reflected 10 or fewer points for the posttreatment unweighted PAR score.

Initial and final models exported in STL format (Figure 1) were superimposed using Geomagic Design X software (Rock Hill, SC) by registering at least four landmarks on the medial two-thirds of the rugae area as a reference (Figure 2).[14]

For quantitative assessment of tooth movement, specific landmarks were defined to construct a local coordinate system on a single tooth. For central and lateral incisors, four landmarks were located: the most mesial point on the incisal edge, the most distal point on the incisal edge, the most gingival point on the palatal surface indicating the long axis, and the most gingival point on the labial surface indicating the gingival zenith. For the canine, three landmarks were located: the most prominent point on the mesial side,

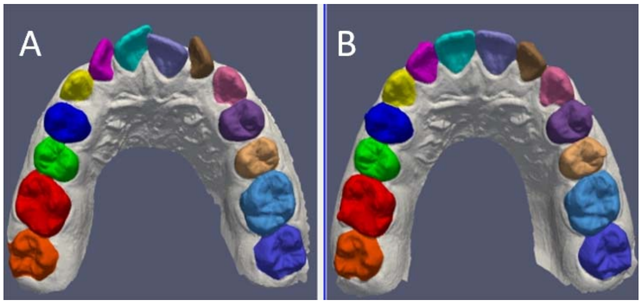

Figure 3. Automatic tooth segmentation of pretreatment (A) and posttreatment (B) model via TS-MDL.

the most prominent point on the distal surface, and the cusp tip.

A deep learning–based system called Two-Stage Mesh Deep Learning (TS-MDL)[15] was implemented for automatic tooth segmentation and landmark identification on all scans of 116 subjects after preliminarily training with an additional data set of 36 samples (Figures 3 and 4). Although the deep learning system was validated to achieve an accuracy of mean absolute error of 0.597 $\pm$ 0.761 mm for landmarking, two experienced orthodontic residents still refined these landmarks manually to ensure the accuracy of landmark location. In addition, inter-evaluator reliability between two examiners for manual landmark identification was assessed using 30 randomly selected samples.

By locating these landmarks formed on the concept of tooth anatomy, an individual three-dimensional coordinate system was defined on each tooth (Figure 5). Based on coordinate transformation, six degrees of freedom (DOF), representing six types of malalignment (such as buccal-lingual translation, mesial-distal translation, extrusion-intrusion, tipping, rotation, and torque) were identified, and the quantitative changes in all dimensions were precisely measured for each tooth (Figure 6). Based on the current literature,[16] all linear values in millimeters were converted into equivalent angular values in degrees for comparison; 1 mm was defined as 2.63°. The total number of trays, including all refinement sets, was used to calculate the total

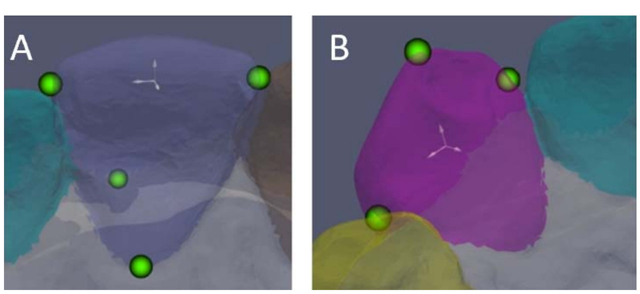

Figure 4. Landmark identification of incisors (A) and canine (B) via TS-MDL.

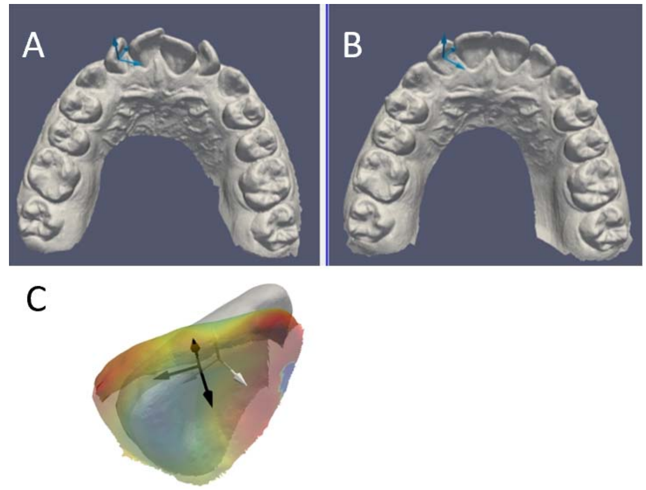

Figure 5. Coordinate system of pretreatment (A) and posttreatment (B) and its transformation (C).

treatment time, and one set of trays corresponded to 1 week of treatment time, based on the activation frequency suggested by the manufacturer.

Statistical Analysis

Regression analysis was performed to find the factor contributing most to prolonging the orthodontic treatment and the correlation between 6 DOF and the total number of trays. Signed values in mesial-distal, buccal-lingual, extrusive-intrusive, and mesial in-mesial out rotation were investigated separately. Because there was no significant difference between both directions, absolute values were used for the statistical analysis. According to Boyd,[17] the lead tooth that needs the most correction would dictate how many stages are required. Therefore, the maximum absolute value of

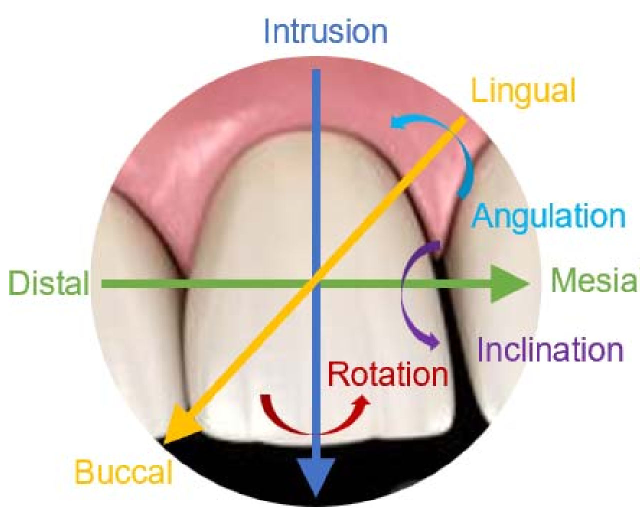

Figure 6. Six degrees of freedom (DOF) representing six malalignment factors.

Table 1. Basic Demographic Data[a]

Sample size, n	116
Total number of trays, mean (SD)	38.73 (13.87)
Treatment start age, y, mean (SD)	34.05 (14.23)
Treatment finish age, y, mean (SD)	35.26 (14.23)
Sex, No. (%)	
Female	88 (75.9)
Male	28 (24.1)
Pretreatment unweighted PAR, mean (SD)	8.64 (4.12)
Pretreatment weighted PAR, mean (SD)	12.99 (7.52)
Pretreatment upper-segment PAR, mean (SD)	3.02 (1.88)
Pretreatment lower-segment PAR, mean (SD)	2.56 (2.03)
Posttreatment unweighted PAR, mean (SD)	2.42 (2.07)
Posttreatment weighted PAR, mean (SD)	4.03 (4.16)
Unweighted PAR difference, mean (SD)	6.21 (3.76)
Weighted PAR difference, mean (SD)	8.97 (6.77)

[a] PAR indicates Peer Assessment Rating Index; SD, standard deviation.

Table 2. Absolute Maximum Value of Each Tooth Movement[a]

Variable	
absMax_buccal_lingual translation, mm, mean (SD)	0.86 (0.50)
absMax_mesial_distal translation, mm, mean (SD)	0.57 (0.57)
absMax_extrusion_intrusion, mm, mean (SD)	0.55 (0.30)
absMax_tipping, °, mean (SD)	4.00 (1.81)
absMax_rotation, °, mean (SD)	7.17 (4.23)
absMax_torque, °, mean (SD)	5.79 (3.46)

[a] absMax_ indicates the absolute maximum value of each tooth movement among six teeth; SD, standard deviation.

each movement among six teeth was considered a representative value for each type of tooth movement.

The equation for a composite score was derived from the linear regression model. The relationship between the total number of trays and a composite score, defined as the integrated value of 6 DOF, was also analyzed to see how the combined movements affected the total treatment duration. A correlation test was applied to find the interaction between each type of malocclusion corrected and the treatment duration. Regression analysis was also performed to demonstrate the relationship between the initial anterior segment PAR score, sex, and the total number of trays.

RESULTS

The basic demographic data are presented in Table 1. The sample consisted of 116 subjects (88 females and 28 males), with a mean age of 34.05 years (range, 16–67 years). The mean pre- and posttreatment unweighted PAR scores were 8.64 and 2.42, respectively. Mean pretreatment upper and lower anterior malalignment PAR scores were 3.02 and 2.56, respectively. The mean total number of trays, including trays for refinement, was 38.73. There was no significant difference between females and males in pretreatment PAR score and each tooth movement (Figure 7).

Interevaluator reliability between two independently calibrated evaluators for manual landmark identification showed a high level of agreement, with an intraclass correlation coefficient ranging from 0.975 to 0.999. The mean absolute difference was 0.0359 mm (range, 0.0172–0.0833 mm).

The mean maximum amounts of buccal-lingual translation, mesial-distal translation, extrusion-intrusion, tipping, rotation, and torque among maxillary incisors and canines were 0.86 mm, 0.57 mm, 0.55 mm, 4.00°, 7.17°, and 5.79°, respectively (Table 2).

There was a significant difference in the total number of trays used between males and females (Figure 8), with males using 8.79 more trays than females ($P = .0015$). Since the values were significantly different

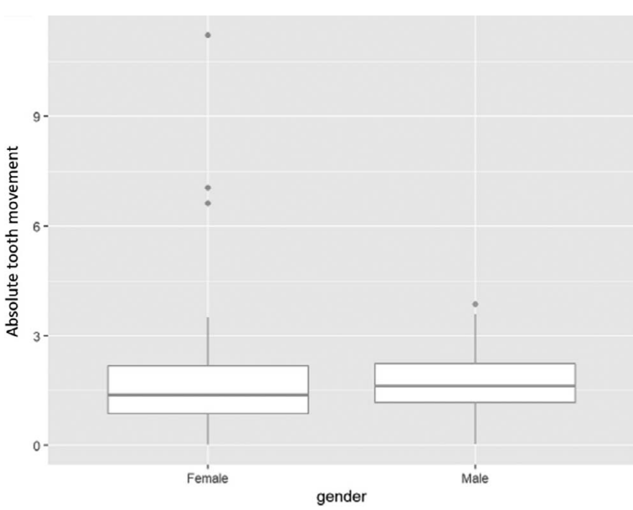

Figure 7. Comparison of maximum absolute tooth movement between males and females.

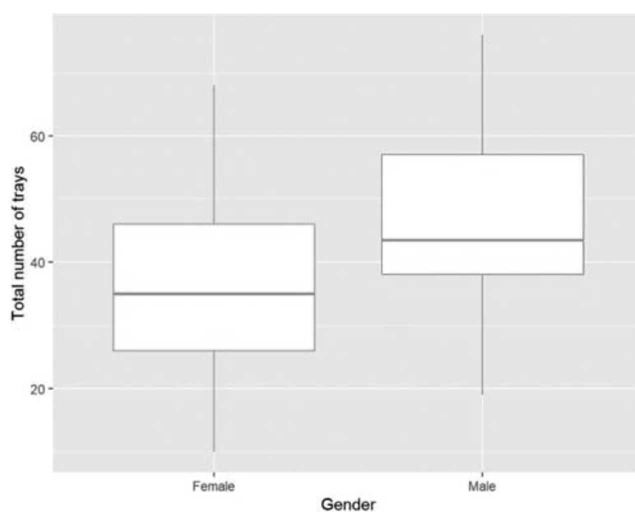

Figure 8. Comparison of the total number of trays between males and females.

Table 3. The Effect of Orthodontic Tooth Movement on the Total Number of Trays With 95% Confidence Intervals for the Coefficient Estimates

Variable	Estimate	Standard Error	*P* Value	Lower Bound of 95% CI	Upper Bound of 95% CI
absMax_buccal_lingual	−2.7534	1.6039	.0889	−5.9325	0.4259
absMax_mesial_distal	2.2759	1.5643	.1486	−0.8249	5.3767
absMax_extrusion_intrusion	−1.1306	2.0180	.5765	−5.1306	2.8694
absMax_tipping	0.0044	0.8583	.9960	−1.6969	1.7056
absMax_rotation	0.2567	0.3428	.4556	−0.4227	0.8361
absMax_torque	1.1299	0.5745	.0518	−0.0089	2.2687
sex	9.8023	3.0285	.0016	3.7994	15.8052
R^2: 0.125					

[a] absMax_ indicates the absolute maximum value of each tooth movement among six teeth; CI, confidence interval.

between women and men, a regression model with adjustment for sex was formulated (Table 3). The effect of the type of movement on the total number of trays demonstrated that no movement was significantly associated with the total number of trays. The absolute maximum torque was the only movement that was marginally associated with the total number of trays ($P = .052$). In addition, the percentages of the number of teeth for which the maximum absolute value was obtained (among the six teeth) are given in Table 4.

The composite score of the six movements based on the model was significantly associated with the total number of trays ($P = .045$), suggesting that the composite score has a higher correlation with the total number of trays than do individual movements (Figure 9). In addition, the correlation coefficient showed that each movement type was not highly correlated.

Additionally, the initial PAR score influenced total treatment duration. Pretreatment upper and lower anterior segment PAR scores were associated with the total number of trays ($P < .001$).

DISCUSSION

There was a significant difference in the total number of trays used between males and females, while both had similar baseline characteristics regarding age and each tooth movement. Patient level of compliance affects the total treatment time and may be related to sex. The percentage of planned movement has been attained, and the number of refinements can be a clue to tracking the contribution of compliance.

Previous studies have reported that aligners can perform torque movement effectively.[9] The amount of tooth movement was first analyzed by differentiating lingual and buccal crown torque as a signed value. However, there was no significant difference between both directions. In this study, absolute torque was marginally associated with the total number of trays ($P = .0518$). Clinicians might spend more time correcting the torque as a result of paying greater attention to the difference in anterior torque.

On the other hand, there was no significant correlation between mesial-distal tipping and the number of trays ($P = .9960$). Drake et al.[18] assessed the type of tooth movement in both aligners and conventional orthodontic appliances and demonstrated that the majority of tooth movement was tipping with the aligner system. As tipping has not been considered challenging to attain with aligners, planned tipping movement can be achieved in a relatively short time.

It is known that bodily movement, which requires a broader area of alveolar remodeling, takes a much more extended period than does tipping. Thus, it was assumed that there would be a positive correlation between translational movement and the number of trays. However, a significant relationship was not found between the amount of buccal-lingual translation and total treatment time ($P = .0889$), nor was this the case in the relationship between mesial-distal translation and total treatment time ($P = .1486$). It may have been because only movement on the crown portion, and not the actual amount of root movement, was measured, given the limitations of the digital scan data. Additionally, extraction cases were not included, and they would be expected to have more translational movement involved.

In addition, this study did not find a significant relationship between extrusive-intrusive movement and the total number of trays ($P = .5765$). Extrusion and intrusion were analyzed separately in this study, with the expectation of prolonged treatment time with

Table 4. The Percentages of the Number of Teeth for Which the Maximum Absolute Value Was Obtained Among the Six Teeth

	Buccal-Lingual Translation, %	Mesial-Distal Translation, %	Extrusion Intrusion, %	Tipping, %	Rotation, %	Torque, %	Average, %
Central incisors	37.06	31.89	25.86	15.51	30.17	40.51	30.17
Lateral incisors	37.06	32.75	35.34	43.10	37.93	39.65	37.63
Canines	25.86	35.34	38.79	41.37	31.89	19.82	32.18

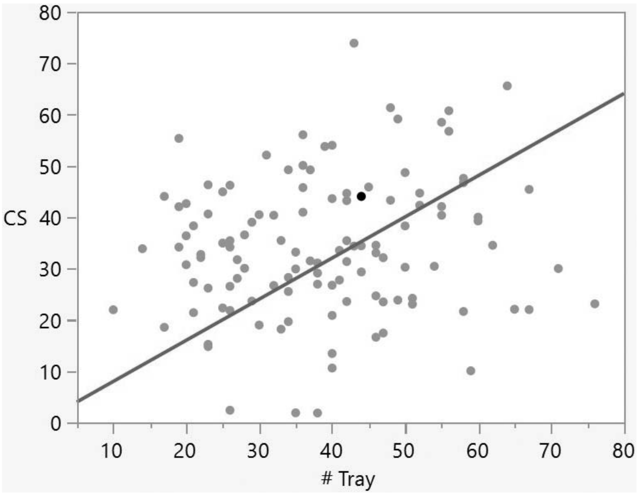

Figure 9. The correlation between composite score (CS) and the total number of trays.

extrusive movement. However, there were no significant differences between signed and absolute values regarding the treatment duration. This may have been because the extrusive and intrusive movements in this study were too subtle, with an average movement of 0.318 mm of extrusion, compared to a previous study.[19] In addition, the subjects were mainly crowded or had spacing for which vertical correction was not critical.

In this study, the amount of rotation was not related to total treatment time ($P = .4556$). The total amount of rotation was too small, with 4.46° of movement, as complicated cases requiring a large amount of rotational movement were excluded on the basis of clinical judgment during the screening process.

The composite score showed a significant correlation ($P < .001$) with the total number of trays. The fitted model combining 6 DOF demonstrated superior predictability to individual tooth movements. From a clinical standpoint, this makes sense, as more malalignment factors come together and take more time to correct.

There was a correlation between the anterior segment PAR score and the total number of trays ($P < .001$), implying that the total treatment time can be influenced by the amount of upper and lower anterior crowding. In this study, because of the lack of a stable registration method for superimposition of the mandibular arch, only upper incisors and canines were evaluated, and there was no assessment of the lower dentition, except for the pretreatment PAR score.

This study had some limitations. There is the possibility that the subjects in the study were not representative of the whole aligner patient pool, since this study was retrospective and observational in nature, with a risk of selection bias. In addition, only patients with mild to moderate crowding or spacing were included without considering sagittal and vertical skeletal relationships.

Measuring the amount and direction of tooth movement has been considered challenging because of the need to evaluate in three different planes of space. This study was distinctive and robust because deep learning technologies were applied to perform numerous steps more efficiently and precisely. For a more precise quantitative evaluation of the orthodontic changes in all dimensions for each tooth, three-dimensional superimposition of initial and final models, tooth segmentation, and landmark registration are required. It is hoped that these extensive data would be used as training resources to improve performance contributing to future studies.

CONCLUSIONS

- Deep learning technologies allowed massive calculation of DOF.
- The evidence to conclude that certain types of tooth movement affect total aligner treatment time is insufficient.
- In aligner treatment, a combined degree of freedom seems to be a better predictor for total treatment time than are individual malalignment factors.
- Upper and lower anterior malalignment factors have a significant effect on the total treatment duration.

REFERENCES

1. Shalish M, Cooper-Kazaz R, Ivgi I, et al. Adult patients' adjustability to orthodontic appliances. Part I: a comparison between labial, lingual, and Invisalign. *Eur J Orthod*. 2012; 34:724–730.
2. Robertson L, Kaur H, Fagundes NCF, Romanyk D, Major P, Flores Mir C. Effectiveness of clear aligner therapy for orthodontic treatment: a systematic review. *Orthod Craniofac Res*. 2020;23:133–142.
3. Rossini G, Parrini S, Castroflorio T, Deregibus A, Debernardi CL. Efficacy of clear aligners in controlling orthodontic tooth movement: a systematic review. *Angle Orthod*. 2015;85: 881–889.
4. Djeu G, Shelton C, Maganzini A. Outcome assessment of Invisalign and traditional orthodontic treatment compared with the American Board of Orthodontics objective grading system. *Am J Orthod Dentofacial Orthop*. 2005;128:292–298; discussion 298.
5. Graf I, Puppe C, Schwarze J, Hofer K, Christ H, Braumann B. Evaluation of effectiveness and stability of aligner treatments using the Peer Assessment Rating Index. *J Orofac Orthop*. 2021;82:23–31.
6. Kravitz ND, Kusnoto B, BeGole E, Obrez A, Agran B. How well does Invisalign work? A prospective clinical study evaluating the efficacy of tooth movement with Invisalign. *Am J Orthod Dentofacial Orthop*. 2009;135:27–35.
7. Simon M, Keilig L, Schwarze J, Jung BA, Bourauel C. Treatment outcome and efficacy of an aligner technique—

regarding incisor torque, premolar derotation and molar distalization. *BMC Oral Health*. 2014;14:68.
8. Grunheid T, Loh C, Larson BE. How accurate is Invisalign in nonextraction cases? Are predicted tooth positions achieved? *Angle Orthod*. 2017;87:809–815.
9. Haouili N, Kravitz ND, Vaid NR, Ferguson DJ, Makki L. Has Invisalign improved? A prospective follow-up study on the efficacy of tooth movement with Invisalign. *Am J Orthod Dentofacial Orthop*. 2020;158:420–425.
10. Gu J, Tang JS, Skulski B, et al. Evaluation of Invisalign treatment effectiveness and efficiency compared with conventional fixed appliances using the Peer Assessment Rating Index. *Am J Orthod Dentofacial Orthop*. 2017;151: 259–266.
11. Lin E, Julien K, Kesterke M, Buschang PH. Differences in finished case quality between Invisalign and traditional fixed appliances. *Angle Orthod*. 2022;92:173–179.
12. Ko CC, Tanikawa C, Wu TH, et al. Machine learning in orthodontics: application review. Embracing novel technologies in dentistry and orthodontics. Paper presented at: 46th Annual Moyers Symposium and the 44th Annual International Conference on Craniofacial Research; March 1–3, 2019; Ann Arbor, Mich.
13. 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.
14. Chen G, Chen S, Zhang XY, et al. Stable region for maxillary dental cast superimposition in adults, studied with the aid of stable miniscrews. *Orthod Craniofac Res*. 2011;14:70–79.
15. Wu T-H, Lian C, Lee S, Pastewait M, Piers C, Liu J, Wang F, Li Wang, Chiu C-Y, Wang W, Jackson C, Chao W-L, Shen D, Ko CC. Two-stage mesh deep learning for automated tooth segmentation and landmark localization on 3D intraoral scans. *IEEE Trans Med Imaging*. 2022. https://doi.org/10.48550/arXiv.2109.11942022.
16. Mutinelli S, Manfredi M, Cozzani M. A mathematic-geometric model to calculate variation in mandibular arch form. *Eur J Orthod*. 2000;22:113–125.
17. Boyd RL. Complex orthodontic treatment using a new protocol for the Invisalign appliance. *J Clin Orthod*. 2007; 41:525–547; quiz 523.
18. Drake CT, McGorray SP, Dolce C, Nair M, Wheeler TT. Orthodontic tooth movement with clear aligners. *ISRN Dent*. 2012;2012:657973.
19. Harris K, Ojima K, Dan C, et al. Evaluation of open bite closure using clear aligners: a retrospective study. *Prog Orthod*. 2020;21:23.