## Original Article

# How accurate is Invisalign in nonextraction cases? Are predicted tooth positions achieved?

## Thorsten Grünheid<sup>a</sup>; Charlene Loh<sup>B</sup>; Brent E. Larson<sup>c</sup>

## ABSTRACT

Objective: To evaluate the accuracy of Invisalign technology in achieving predicted tooth positions with respect to tooth type and direction of tooth movement.

Materials and Methods: The posttreatment models of 30 patients who had nonextraction Invisalign treatment were digitally superimposed on their corresponding virtual treatment plan models using best-fit surface-based registration. The differences between actual treatment outcome and predicted outcome were computed and tested for statistical significance for each tooth type in mesial-distal, facial-lingual, and occlusal-gingival directions, as well as for tip, torque, and rotation. Differences larger than  $0.5$  mm for linear measurements and  $2^{\circ}$  for angular measurements were considered clinically relevant.

**Results:** Statistically significant differences ( $P < .05$ ) between predicted and achieved tooth positions were found for all teeth except maxillary lateral incisors, canines, and first premolars. In general, anterior teeth were positioned more occlusally than predicted, rotation of rounded teeth was incomplete, and movement of posterior teeth in all dimensions was not fully achieved. However, except for excess posttreatment facial crown torque of maxillary second molars, these differences were not large enough to be clinically relevant.

Conclusions: Although Invisalign is generally able to achieve predicted tooth positions with high accuracy in nonextraction cases, some of the actual outcomes may differ from the predicted outcomes. Knowledge of dimensions in which the final tooth position is less consistent with the predicted position enables clinicians to build necessary compensations into the virtual treatment plan. (Angle Orthod. 2017;87:809–815.)

KEY WORDS: Invisalign; aligner; treatment outcome

## INTRODUCTION

In the past two decades, the field of orthodontics has been revolutionized by technological advancements. Three-dimensional imaging has expanded diagnostic and treatment planning abilities,<sup>1</sup> intraoral scanners now provide an alternative to traditional impressions, and digital models can replace plaster models for both treatment planning and appliance fabrication.<sup>2,3</sup> Combined with increasing patient demand for esthetic treatment options and the drive toward personalized treatment, these developments have given rise to a number of clear aligner systems now serving as alternatives to conventional bracket-and-archwire systems.<sup>4</sup>

In 1999, Align Technology (Santa Clara, Calif) introduced Invisalign as the pioneer clear aligner system for comprehensive orthodontic treatment. Invisalign has continually evolved through the development of new aligner materials, attachments on teeth, staging of tooth movement, and incorporation of interproximal reduction and interarch elastics to address a wider range of malocclusions.<sup>5,6</sup>

According to the company's internal data, more than 3 million patients have been treated with the Invisalign system in more than 90 countries worldwide. Despite its widespread use, relatively few studies have quantified the system's efficacy. This is significant because it has been suggested that aligners have limitations when it

<sup>a</sup> Assistant Professor, Division of Orthodontics, School of Dentistry, University of Minnesota, Minneapolis, Minn.

**Private Practice, Richmond, British Columbia, Canada** 

<sup>c</sup> Professor and Director, Division of Orthodontics, School of Dentistry, University of Minnesota, Minneapolis, Minn.

Corresponding author: Dr Thorsten Grünheid, Division of Orthodontics, School of Dentistry, University of Minnesota, 6-320 Moos Health Science Tower, 515 Delaware Street S.E., Minneapolis, MN 55455

<sup>(</sup>e-mail: tgruenhe@umn.edu)

Accepted: May 2017. Submitted: February 2017.

Published Online: July 7, 2017

<sup>-</sup> 2017 by The EH Angle Education and Research Foundation, Inc.



Figure 1. Superimposition of digital models to compute differences between predicted and achieved tooth positions. (A) Matching points for initial registration. (B) Global alignment of posttreatment model (orange) and virtual treatment plan model (white). (C) Placement of coordinates at the center of resistance of each tooth. (D) Superimposition of individual teeth of posttreatment model (green) and virtual treatment plan model (white) using best-fit surface-based registration.

comes to producing certain tooth movements.<sup>5</sup> For instance, questions have been raised regarding the extent to which aligners can control extrusion, rotation, bodily movement, and torque.<sup>5</sup> Some authors even doubt that bodily movements or torque can be accomplished at all by aligners.<sup>7</sup> Therefore, the aim of this study was to evaluate the efficacy of Invisalign technology to achieve predicted tooth positions with respect to tooth type and direction of movement.

## MATERIALS AND METHODS

Approval for this retrospective cohort study was granted by the institutional review board at the University of Minnesota (Study 1411M56201). A total of 30 consecutive patients (13 male, 17 female; age 21.6  $\pm$  9.8 years) were selected based on the following inclusion criteria: full permanent dentition including second molars in both arches, nonextraction Invisalign treatment with no deviation from the default amounts of tooth movement embedded in each aligner stage, aligners changed every 2 weeks following the manufacturer's protocol, no midcourse corrections or additional aligners, and no combined treatment with fixed appliances, intraoral distalizers, or other auxiliary appliances. Patients were excluded if they required oral surgery or received dental restorations during treatment. Treatment was provided by 12 orthodontic residents and 10 orthodontists certified in the use of the Invisalign system. The residents provided care under the supervision of the orthodontists, and each virtual treatment plan was reviewed and approved by an orthodontist to ensure that no unrealistic goals were set. The average treatment time was  $11 \pm 4$  months. Of the 30 patients, 22 had class I molar occlusions, 7 had class II molar occlusions, and 1 had a class III molar occlusion (all less than 2 mm). The average amount of crowding in each arch was  $2 \pm 2$  mm. Interproximal enamel reduction was performed as prescribed in each patient's virtual treatment plan.

To obtain posttreatment digital models, all patients had alginate impressions taken and poured into plaster casts, which were then digitized using an R700 orthodontic model scanner (3Shape A/S, Copenhagen, Denmark). To obtain digital models of the predicted outcome (virtual treatment plan models), the final stage of each patient's virtual treatment plan was exported through Align Technology's ClinCheck program and converted to digital models using e-model 9.0 software (GeoDigm Corporation, Falcon Heights, Minn). All digital models were de-identified, and soft tissue and bonded retainers were digitally removed to ensure that evaluation was based solely on tooth-surface features. The posttreatment models were segmented to isolate each tooth as a separate object and compared with the unsegmented virtual treatment plan models using emodel Compare 8.1 software (GeoDigm). This software compares individual tooth positions between two digital models. Corresponding dental arches are first aligned globally, and then individual teeth from a segmented model are superimposed on analogous teeth of an unsegmented model using a best-fit algorithm so that differences between tooth positions can be computed (Figure 1).

#### HOW ACCURATE IS INVISALIGN? 811

Table 1. Pearson's Correlation Coefficients  $(r)$  for Intraoperator Agreement

<b>Direction</b>	
Mesial-Distal	0.9360
Facial-Lingual	0.9864
Occlusal-Gingival	0.9635
Tip	0.9397
Torque	0.9858
Rotation	0.9481

For global alignment, the mesial-buccal cusps of the first molars and the mesial-incisal point of the right central incisor in each arch were used as matching points for initial registration. This initial registration was then refined by 50 iterations of a closest-point algorithm to achieve best fit of the occlusal surfaces. After that, a single operator placed a reference coordinate system with the origin of the axes at each tooth's approximate center of resistance for each tooth of the posttreatment model. Because the center of resistance depends on many local factors such as tooth morphology, root length, attachment levels, and the direction of force application,<sup>8,9</sup> which could not be determined for each individual tooth, the coordinates were placed at a point in the center of each tooth, 8 mm apical to the cemento-enamel junction. This was based on the assumption that the center of resistance was situated halfway between the alveolar crest and the root apex, the average root length of all tooth types, and a biologic width of 2 mm.<sup>10-12</sup> Once the axes were placed in the posttreatment model, the software automatically generated analogous axes for each corresponding tooth in the virtual treatment model. The software then superimposed individual teeth from the segmented posttreatment model on the corresponding teeth in the unsegmented virtual treatment model using best-fit surface-based registration. Based on the transformation of axes required to fit each tooth, the software quantified the differences between achieved and predicted position for each tooth in the following six directions: mesial-distal, facial-lingual, occlusal-gingival, tip, torque, and rotation. The differences were expressed with respect to the center of resistance.

Because the software allowed for the detection of differences that were too small to be clinically relevant, threshold values were chosen in reference to the American Board of Orthodontics (ABO) model grading system for case evaluation.<sup>13</sup> According to the model grading system criteria, discrepancies of 0.5 mm or greater in the alignment of contact points and marginal ridges will result in the deduction of points. A marginal ridge discrepancy of 0.5 mm equates to a crown-tip deviation of  $2^\circ$  for an average-sized molar. Therefore, differences of 0.5 mm or more in the mesial-distal,

Table 2. Results of Bland-Altman Analyses for Intraoperator Agreement

<b>Direction</b>	<b>Mean Difference</b>	Limits of Agreement
Mesial-Distal (mm)	0.0041	$-0.3109, 0.3191$
Facial-Lingual (mm)	0.0055	$-0.1864, 0.1974$
Occlusal-Gingival (mm)	0.0039	$-0.1043, 0.1122$
Tip $(^\circ)$	$-0.0480$	$-1.5444, 1.4484$
Torque $(^\circ)$	0.0303	$-0.8646, 0.9252$
Rotation $(°)$	0.0719	$-1.3154, 1.4592$

facial-lingual, and occlusal-gingival directions and differences of  $2^\circ$  or more in tip, torque, and rotation were considered clinically relevant.

## Statistical Analysis

Data from each patient's left and right analogous teeth were pooled for analysis after a linear mixed model had been used to verify that there were no significant side differences. Descriptive statistics were computed for the differences between predicted and achieved tooth positions in each of the six directions. A linear mixed model was used to calculate the corresponding 95% confidence interval for each mean difference. To assess whether the differences were statistically significant, P values were calculated using a false discovery rate method to adjust for the multiple comparisons performed.

Equivalence testing was used to assess whether the differences between predicted and achieved tooth positions were large enough to be clinically relevant. Two one-sided *t*-tests were used to test for differences above 0.5 mm and below  $-$ 0.5 mm, and above 2 $^{\circ}$  and below –2°. Mean differences that fell within –0.5 mm to  $+0.5$  mm for linear measurements and within  $-2^{\circ}$  and  $+2^{\circ}$  for angular measurements were practically equivalent and therefore considered too small to be clinically relevant.

A post-hoc power analysis was performed based on 30 independent samples to estimate the power of the study to detect differences that were small enough to fall within the equivalent region between -0.5 mm to  $+0.5$  mm or  $-2^{\circ}$  to  $+2^{\circ}$ .

Because the calculated differences between predicted and achieved tooth positions included both positive and negative values, an additional analysis was performed on the absolute values of the mean differences to eliminate the possibility of positive and negative values averaging in a mean close to zero and giving the false impression of clinical accuracy. For this, the data based on absolute values were log transformed to normalize the distributions, and a onesided test of equivalence was applied to the log transformed values.

Table 3. Results of Power Analysis

<b>Direction</b>	Power
Mesial-Distal	0.999
Facial-Lingual	0.986
Occlusal-Gingival	>0.999
Tip	0.977
Torque	0.761
Rotation	0.970

Placement of coordinates and superimposition were repeated for 10 randomly selected patients, and Pearson correlation coefficients and Bland-Altman analyses were used to assess intraoperator agreement for each of the six directions of tooth movement. Statistical analyses were performed with SAS 9.4 for Windows (SAS Institute Inc, Cary, N.C.), and P values of less than 0.05 were considered statistically significant.

## RESULTS

Intraoperator agreement was generally very high for all measurements. The Pearson correlation coefficients and the results of the Bland-Altman analyses are reported in Tables 1 and 2, respectively.

The power to detect differences small enough to fall below the thresholds set for clinical relevance in the various directions is reported in Table 3. Because the power calculations were based on 30 patients as independent samples, the actual power is even higher because each of the 30 patients had 28 teeth.

The differences between predicted and achieved tooth positions are reported in Table 4. In the maxillary

arch, several tooth types showed statistically significant differences between predicted and achieved tooth positions: The central incisor was positioned more facially and occlusally and had more lingual crown torque. The second premolar was more distal and lingual relative to the predicted position and had more facial crown torque. The first molar was aberrant in these same three directions and had more mesial crown tip than predicted. The second molar had more facial crown torque compared to the prediction and was positioned more lingually and occlusally. This difference in maxillary second molar torque between the predicted and achieved position exceeded  $2^\circ$  and was therefore considered clinically relevant.

In the mandibular arch, all tooth types showed statistically significant differences. Both central and lateral incisors were positioned more occlusally than predicted. The lateral incisor also had more mesial rotation. The canines were more lingual and had more facial crown torque and distal rotation than predicted. Both first and second premolars had more mesial rotation than predicted. Finally, both the first and second molars had more facial crown torque than predicted. The second molar also had more distal crown tip. Although statistically significant, none of these differences in the mandibular arch were considered clinically relevant.

Statistical analysis performed on the absolute value of each discrepancy measurement did not reveal any additional differences that had not been detected previously.





<sup>a</sup> Results are mean values  $\pm$  standard deviations (with 95% confidence intervals in parentheses). Positive values indicate an achieved tooth position more distal, lingual or gingival, or with more distal crown tip, more lingual crown torque, or more distal rotation than the predicted tooth position.

\* Statistically significant difference between predicted and achieved tooth position ( $P < 0.05$ ).

## **DISCUSSION**

Using mathematical superimposition of digital models, it has become possible to quantify treatment changes and, as in the present study, discrepancies between virtual treatment plans and actual treatment outcomes. Previous studies that evaluated the effectiveness of clear aligner therapy used the ABO model grading system,<sup>14-16</sup> Tooth Measure (Align Technology)<sup>17-20</sup> or Surfacer (Imageware, Plano, Tex) software.<sup>6</sup> Although these tools can provide a general assessment of accuracy, the software used in the current study was uniquely able to quantify differences between objects with respect to six degrees of freedom.21 The software calculates differences automatically, that is, not influenced by potential operator bias, and has been previously used in other projects directed at quality control and outcomes assessment.<sup>22–24</sup> However, the multistep process used to create the digital models constituted a potential source of error. The dimensional accuracy of the plaster casts from which the digital models were derived may have been affected by possible distortion or shrinkage of the impression material, and the accuracy of the digitization may have been limited by the resolution of the model scanner (20  $\mu$ m).

Predicted and achieved tooth positions differed in all tooth types. Anterior teeth were often positioned too far occlusally, rounded teeth such as mandibular canines and premolars were not fully rotated, and posterior teeth had discrepancies in all directions. The largest difference was found for maxillary second molar torque, which exceeded  $2^\circ$  and was therefore considered clinically relevant. This difference in facial-lingual inclination of second molars has also been described following treatment with traditional fixed appliances.<sup>25</sup> This may be related to the decreasing amount of force exerted by the end of an archwire as interbracket distance and flexibility of the wire increase. Moreover, molars have larger root surface areas and require greater forces for tooth movement.<sup>10</sup> The same concept may apply for clear aligner therapy; there seems to be greater flexibility and less force exerted by the posterior segments of aligners.

In addition, maxillary posterior teeth were positioned more lingual with more facial crown torque than predicted. It is likely that maxillary arch expansion was not fully achieved and the molars tipped rather than moved bodily during the process, both of which could have resulted from flexing of the aligners. This notion is supported by a recent study that found the mean accuracy of maxillary expansion with Invisalign to be 72.8% with more tipping observed than predicted in the virtual treatment plan.<sup>26</sup> The mandibular molars also had more facial crown torque than predicted. This, too, could be the consequence of an inability of the aligners to fully express the torque specified in the virtual treatment plan and may have been compounded by biological limitations such as proximity of the molar roots to the cortical plate of the mandible.

With regard to incisors, the results of the current study resemble those of others that found movements of anterior teeth to have relatively poor accuracy. For instance, both Kravitz et al.<sup>17</sup> and Krieger et al.<sup>18</sup> reported deficiencies in vertical incisor movement with only 44% to 46% of the predicted intrusion achieved for central incisors. Thus, significant correction of a deep





overbite with Invisalign appears difficult. Similarly, the difference in maxillary central incisor torque found in the current sample was consistent with other studies that observed tipping of incisors rather than bodily movement.19,27 Possible reasons for the more upright positions include potential torque loss during space closure or the use of class II interarch elastics. Surprisingly, no significant differences in the occlusalgingival position or torque of the maxillary lateral incisors were found in the patients in this study. Anecdotally, these teeth are often challenging to treat and, because of their shape, sometimes require the use of auxiliaries such as bonded buttons combined with intraoral elastics.<sup>17</sup>

Interestingly, there were statistically significant differences in mandibular canine facial-lingual position, torque, and rotation, but not tip. Although the discrepancies in mandibular canine torque and rotation in the current data coincided with reports of minimal torque changes and low accuracy of canine rotation produced by clear aligners previously,<sup>28,29</sup> the minimal discrepancies in angulation differed from studies that found low accuracy of mesiodistal tipping of mandibular canines.17 In contrast to mandibular canine position, maxillary canine position did not differ significantly. This is of interest because the accuracy of rotation and tip of maxillary canines has been described as notoriously low,17,28 likely because these teeth have the longest roots in the dentition with large root surface areas, requiring greater force to produce orthodontic tooth movement.<sup>10</sup>

Although the present results suggest that the intraarch tooth position predicted by the virtual treatment plan is not consistently achieved by Invisalign aligners and some limitations in the appliance system remain, they in no way suggest unsatisfactory treatment results. In fact, the difference in maxillary second molar torque between predicted and achieved position was the only difference considered clinically relevant. When planning clear aligner therapy, clinicians may utilize Align Technology's ClinCheck program to design their biomechanics rather than merely as a tool for visualization of predicted treatment outcomes. Knowledge of the dimensions in which the final tooth position is less consistent with the predicted position allows them to pay extra attention to more challenging movements, modify attachments, and build specific overcorrections into their virtual treatment plans to increase efficiency and achieve better treatment outcomes.

## **CONCLUSIONS**

In general, Invisalign is able to achieve predicted tooth positions with high accuracy in nonextraction cases. Clinicians may consider the following when planning treatment with Invisalign:

- Maxillary arch expansion may not be fully achieved.
- Mandibular incisors tend to be positioned more occlusally than predicted.
- Rotation of rounded teeth may be incomplete.
- Molar torque may not be fully achieved, with maxillary second molars often having a clinically relevant magnitude of more facial crown torque than predicted.

#### REFERENCES

- 1. Kau CH, Richmond S, Palomo JM, Hans MG. Threedimensional cone beam computerized tomography in orthodontics. J Orthod. 2005;32:282–293.
- 2. Al Mortadi N, Jones Q, Eggbeer D, Lewis J, Williams RJ. Fabrication of a resin appliance with alloy components using digital technology without an analog impression. Am J Orthod Dentofacial Orthop. 2015;148:862–867.
- 3. Rossini G, Parrini S, Castroflorio T, Deregibus A, Debernardi CL. Diagnostic accuracy and measurement sensitivity of digital models for orthodontic purposes: a systematic review. Am J Orthod Dentofacial Orthop. 2016;149:161–170.
- 4. Malik OH, McMullin A, Waring DT. Invisible orthodontics part I: Invisalign. Dent Update. 2013;40:203–215.
- 5. Phan X, Ling PH. Clinical limitations of Invisalign. J Can Dent Assoc. 2007;73:263–266.
- 6. 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;14:1–7.
- 7. Brezniak N. The clear plastic appliance: a biomechanical point of view. Angle Orthod. 2008;78:381–382.
- 8. Smith RJ, Burstone CJ. Mechanics of tooth movement. Am J Orthod. 1984;85:294–307.
- 9. Meyer BN, Chen J, Katona TR. Does the center of resistance depend on the direction of tooth movement? Am J Orthod Dentofacial Orthop. 2010;137:354–361.
- 10. Proffit WR, Fields H, Sarver D. Chapter 8 The biologic basis of orthodontic therapy, pp 278–311. Contemporary Orthodontics. 5th ed. St. Louis, Mo: Mosby; 2012. Chapter 9 Mechanical principles in orthodontic force control, pp 312– 346.
- 11. Nelson SJ, Ash MM/ Appendix B, Tooth traits of the permanent dentition, pp 327–334. Wheeler's Dental Anatomy, Physiology, and Occlusion. 9th ed. St. Louis, Mo: Saunders; 2010.
- 12. Gargiulo AW, Wentz FM, Orban B. Dimensions are relations of the dentogingival junction in humans. Periodontol. 1961;32:261–267.
- 13. American Board of Orthodontics. Grading system for dental casts and panoramic radiographs. Available at: https://www. americanboardortho.com/media/1191/grading-system-castsradiographs.pdf. Accessed December 14, 2016.
- 14. 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.
- 15. Kassass W, Al-Jewair T, Preston CB, Tabbaa S. Assessment of Invisalign treatment outcomes using the ABO Model Grading System. J World Fed Orthod. 2013;e61–e64.
- 16. Li WH, Wang SM, Zhang YZ. The effectiveness of the Invisalign appliance in extraction cases using the ABO model grading system: a multicenter randomized controlled trial. Int J Clin Exp Med. 2015;8:8276–8282.
- 17. 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.
- 18. Krieger E, Seiferth J, Marinello I, et al. Invisalign treatment in the anterior region: were the predicted tooth movements achieved? J Orofac Orthop. 2012;73:365–376.
- 19. Drake CT, McGorray SP, Dolce C, Nair M, Wheeler TT. Orthodontic tooth movement with clear aligners. ISRN Dent. 2012:657973.
- 20. Chisari JR, McGorray SP, Nair M, Wheeler TT. Variables affecting orthodontic tooth movement with clear aligners. Am J Orthod Dentofacial Orthop. 2014;145(4 suppl):S82–S91.
- 21. Larson BE, Vaubel CJ, Grünheid T. Effectiveness of computer-assisted orthodontic treatment technology to achieve predicted outcomes. Angle Orthod. 2013;83:557– 562.
- 22. Grünheid T, McCarthy SD, Larson BE. Clinical use of a direct chairside oral scanner: an assessment of accuracy, time, and patient acceptance. Am J Orthod Dentofacial Orthop. 2014;146:673–682.
- 23. Grünheid T, Lee MS, Larson BE. Transfer accuracy of vinyl polysiloxane trays for indirect bonding. Angle Orthod. 2016;86:468–474.
- 24. Lieber WS, Carlson SK, Baumrind S, Poulton DR. Clinical use of the ABO-Scoring Index: reliability and subtraction frequency. Angle Orthod. 2003;73:556–564.
- 25. Houle JP, Piedade L, Todescan R Jr, Pinheiro FH. The predictability of transverse changes with Invisalign. Angle Orthod. 2017;87:19–24.
- 26. Zhang XJ, He L, Guo HM, Tian J, Bai YX, Li S. Integrated three-dimensional digital assessment of accuracy of anterior tooth movement using clear aligners. Korean J Orthod. 2015;45:275–281.
- 27. Grünheid T, Gaalaas S, Hamdan H, Larson BE. Effect of clear aligner therapy on the buccolingual inclination of mandibular canines and the intercanine distance. Angle Orthod. 2016;86:10–16.
- 28. Kravitz ND, Kusnoto B, Agran B, Viana G. Influence of attachments and interproximal reduction on the accuracy of canine rotation with Invisalign: a prospective clinical study. Angle Orthod. 2008;78:682–687.