Predicted and actual end-of-treatment occlusion produced with aligner therapy

Peter H. Buschang; Mike Ross; Steven G. Shaw; Doug Crosby; Phillip M. Campbell

ABSTRACT
Objective: To compare three-dimensional (3D) ClinCheck™ models with the subjects' actual 3D posttreatment models using the American Board of Orthodontics Objective Grading System (OGS).

Materials and Methods: This prospective, within-subject study included 27 consecutive cases treated with aligner therapy. The posttreatment plaster models taken immediately after treatment were scanned and converted to stereolithography (STL) files; the ClinCheck models were also converted to STL format. MeshLab software was used to measure the seven components of the OGS, including alignment, marginal ridges, buccolingual inclinations, occlusal contacts, occlusal relationships, overjet and interproximal contacts. An overall OGS deduction score was also calculated.

Results: Compared with the posttreatment models, the ClinCheck models showed significantly ($P < .016$) fewer overall OGS point deductions (24 vs 15). These overall differences were due to significantly ($P < .05$) more deductions among the posttreatment models than the ClinCheck models for alignment (4.0 vs 1.0 deductions), buccolingual inclinations (4.0 vs 3.0 deductions), occlusal contacts (3.0 vs 2.0 deductions), and occlusal relations (4.0 vs 2.0 deductions).

Conclusion: The ClinCheck models do not accurately reflect the patients' final occlusion, as measured by the OGS, at the end of active treatment. (Angle Orthod. 2015;85:723–727.)

KEY WORDS: Aligners; Occlusion; Prediction; Objective Grading System

INTRODUCTION
In order for aligner treatments to be valid and effective, the predicted and actual outcomes should be comparable. Align Technology provides orthodontists with ClinCheck™ (Align Technology Inc, Santa Clara, Calif) models, which purportedly reflect the treatment outcomes. The orthodontist sends the patient's initial records, plus his or her verbal expectations of the

treatment outcome, to the company. The company then fabricates a series of custom-made aligners that sequentially reposition the teeth. The aligners incrementally shift the teeth into place based on the outcome the orthodontist expects to achieve. The final end-of-treatment aligner is based on the three-dimensional (3D) ClinCheck model, which provides the patient's visual treatment outcome.

For clinical applications, it remains unclear how closely the occlusion of the ClinCheck model compares with the actual occlusion produced during treatment. Based on superimpositions of the ClinCheck and final treatment models of 37 patients, it has been shown that extrusion of the central incisors and tipping of the mandibular canines were the most difficult to accurately predict. Krieger et al., who compared final ClinCheck and posttreatment models of 50 patients, reported differences <0.4 mm for arch width and lengths and a 0.9 mm difference for overbite. Because the existing comparisons are limited to the anterior six teeth, they provide only partial information about the occlusal characteristics thought to be important for evaluating orthodontic treatments.
Due to the lack of previous studies, the purpose of the present study was to compare the ClinCheck model with the actual treatment outcome to determine whether overall occlusion and the various aspects of occlusion were comparable. This study endeavors to establish the relative validity of ClinCheck models by determining whether the 3D treatment outcome of aligner therapy can be accurately predicted.

**MATERIALS AND METHODS**

Twenty-seven consecutive posttreatment patient models were included in this study and compared with their ClinCheck models provided by Invisalign™. The impressions for the posttreatment plaster models were taken immediately after treatment. The models were scanned using the Ortho Insight 3D scanner and converted to a stereolithography (STL) format using the MotionView software (OrthoInsight 3D version 5.5.5002 (MotionView Software, Chattanooga, TN). An STL file was created for each set of models in occlusion as well as one for each maxillary and mandibular arch separately. The digital models were compared with the digital ClinCheck models. The accuracy and reliability of scanned 3D models has been previously established. The study was approved by the Texas A&M University Baylor College of Dentistry Internal Review Board.

The ClinCheck model files were also converted to the STL format to allow them to be read and assessed using the same software. The MeshLab V1.30 software (ISTI–CNR, Pisa, Italy) program was used to make digital measurements and derive the American Board of Orthodontics (ABO) Objective Grading System (OGS) scores on both the posttreatment and ClinCheck models. The measurements were all made by one calibrated investigator (SGS).

The components of the OGS scores included alignment, marginal ridges, buccolingual inclinations, occlusal contacts, occlusal relationships, overjet, and interproximal contacts. Measurements for each component were made from the digital models using the instructions for measurements from an ABO article explaining the OGS scoring system. The MeshLab software allowed the measurements to be made using a digital ruler instead of a standard ruler. After drawing a line on the screen representing the occlusal plane, from which measurements were made to the respective cusps, the zoom function was used to enlarge images 150% in order to measure buccolingual inclinations. The occlusal contacts and marginal ridges were also measured after enlarging the images. Root parallelism was not measured because it is not available for the ClinCheck models. The total number of deductions was measured for each component. The sums of the differences were used to determine the patients’ OGS score for their posttreatment model and their ClinCheck model. The difference between the scores for the ClinCheck model and the corresponding posttreatment model were calculated for each component of the OGS score and for the total score.

Replicate analyses of 12 randomly selected patient records were performed to establish reliability. Wilcoxon signed rank tests show no statistically significant difference between the individual component scores or total scores. Method errors of the component score deductions ranged from 0.29 (interproximal contacts) to 0.82 (occlusal contacts). The method error for the total component score was 1.21 deduction.

**Statistics**

The skewness and kurtosis statistics showed that the measurements were not normally distributed. As such, the data were summarized using medians (50th percentile) and interquartiles (25th and 75th percentiles). The models were statistically compared using Mann-Whitney U-tests. Statistical significance was set at ≤.05.

**RESULTS**

Median scores for each of the components, as well as for the total score, were higher for the actual than for the ClinCheck models (Table 1). The posttreatment models had more points lost in each component of the

---

Table 1. Total Points Lost for American Board of Orthodontics Objective Grading System Score Components and Total Score for the ClinCheck™ Models and Actual Posttreatment Models

<table>
<thead>
<tr>
<th>Component</th>
<th>ClinCheck Model</th>
<th>Actual Model</th>
<th>Group Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alignment</td>
<td>Median</td>
<td>1.0</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>25th</td>
<td>1.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>75th</td>
<td>2.0</td>
<td>4.0</td>
</tr>
<tr>
<td>Marginal ridges</td>
<td>1.0</td>
<td>0.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>25th</td>
<td>3.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>75th</td>
<td>4.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Inclination</td>
<td>2.0</td>
<td>2.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Occlusal contacts</td>
<td>2.0</td>
<td>3.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>25th</td>
<td>3.0</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>75th</td>
<td>4.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Occlusal relations</td>
<td>2.0</td>
<td>1.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>25th</td>
<td>7.0</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>75th</td>
<td>12.0</td>
<td>7.0</td>
</tr>
<tr>
<td>Overjet</td>
<td>2.0</td>
<td>1.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>75th</td>
<td>4.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Interproximal contacts</td>
<td>3.0</td>
<td>0.0</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>75th</td>
<td>5.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Total score</td>
<td>14.0</td>
<td>0.0</td>
<td>17.0</td>
</tr>
</tbody>
</table>

724 BUSCHANG, ROSS, SHAW, CROSBY, CAMPBELL

Angle Orthodontist, Vol 85, No 5, 2015
OGS score. Differences were greatest \((P < .001)\) for alignment, marginal ridges, and occlusal contacts. Differences for occlusal relations were also highly significant. The smallest differences were for inclinations and overjet. No statistically significant difference was found between the interproximal contacts of the ClinCheck and posttreatment models.

The median total score for the ClinCheck models was 27 points with a range of 14–43. The median total score for the posttreatment models was 24 points with a range of 12–43. The median difference between the total scores was 9 points, which was highly significant \((P < .001)\).

**DISCUSSION**

The actual posttreatment models showed more deductions than the ClinCheck models for every component of the OGS. Although the OGS has not been previously used for comparison, it has been shown that overjet measured on the posttreatment models is slightly, but not significantly, larger than overjet on the final ClinCheck models.\(^\text{10}\) It is certainly possible that additional refinement aligners could have been used to improve the scores. However, the orthodontist performing the treatment (DC) purposely decided to take final records immediately after the end of the last aligner (ie, to give no time for occlusion to settle).

The vertical components of the OGS system, including marginal ridges and occlusal contacts, showed the largest differences between predicted treatment outcome and the posttreatment model. For both of these components, the posttreatment models lost 2 points more on average than the respective ClinCheck model. This might be expected if moving teeth with aligners is more difficult in the vertical than the sagittal plane, as previously suggested.\(^2,10,11\) For example, Krieger et al.\(^\text{10}\) showed that there was only a 14.3% concordance in overbite between the predicted and actual treatment results. This would also explain why Kravitz et al.\(^2\) reported that the predicted treatment results were less accurate when teeth were being extruded. Future studies need to be performed to determine whether similar differences are produced after treating subjects who have different malocclusions or were treated with different techniques.

On average, the posttreatment models lost twice as many points for alignment than the respective ClinCheck models. This may have been due to the fact that aligner treatments do not typically incorporate the same level of detailing as conventional braces. In other words a full finishing/detailing phase of treatment may be needed to achieve the results indicated in the ClinCheck model.

Importantly, the actual treatment outcomes attained with aligner therapy compared well with those attained with conventional treatments. The overall OGS scores for the present study were similar to or lower than most overall scores previously reported (Table 2). Variability

### Table 2. Mean Points Lost for American Board of Orthodontics Objective Grading System Scores for the Clincheck Models and Actual Posttreatment Models Compared with Previously Reported Research Results, Along with the Sum of the First Six Components (Minus Interproximal Contacts) and All Seven Components\(^a\)

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Alignment</th>
<th>Marginal Ridges</th>
<th>Buccolingual Inclination</th>
<th>Overjet</th>
<th>Occlusal Contacts</th>
<th>Occlusal Relationship</th>
<th>Interproximal Contacts</th>
<th>Overall (6 Components)</th>
<th>Overall (7 Components)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abei et al.(^\text{14})</td>
<td>126</td>
<td>5.4</td>
<td>3.9</td>
<td>4.5</td>
<td>4.2</td>
<td>4.2</td>
<td>3.1</td>
<td>0.9</td>
<td>25.3</td>
<td>26.1</td>
</tr>
<tr>
<td>Cook et al.(^\text{15})</td>
<td>62</td>
<td>5.4</td>
<td>2.2</td>
<td>1.8</td>
<td>5.8</td>
<td>4.7</td>
<td>3.7</td>
<td>1.4</td>
<td>23.6</td>
<td>25.0</td>
</tr>
<tr>
<td>Cook(^\text{16})</td>
<td>115</td>
<td>6.0</td>
<td>3.7</td>
<td>2.4</td>
<td>3.7</td>
<td>3.5</td>
<td>3.6</td>
<td>NA</td>
<td>22.9</td>
<td>NA</td>
</tr>
<tr>
<td>Costalos et al.(^\text{17})</td>
<td>24</td>
<td>7.8</td>
<td>4.0</td>
<td>6.7</td>
<td>4.7</td>
<td>5.3</td>
<td>2.2</td>
<td>0.3</td>
<td>30.7</td>
<td>30.9</td>
</tr>
<tr>
<td>Deguchi et al.(^\text{18})</td>
<td>54</td>
<td>6.1</td>
<td>2.9</td>
<td>5.6</td>
<td>4.5</td>
<td>5.8</td>
<td>3.7</td>
<td>0.4</td>
<td>28.6</td>
<td>29.0</td>
</tr>
<tr>
<td>Djeu et al.(^\text{19})</td>
<td>48</td>
<td>6.8</td>
<td>4.4</td>
<td>2.8</td>
<td>3.6</td>
<td>5.7</td>
<td>5.5</td>
<td>0.8</td>
<td>28.8</td>
<td>29.5</td>
</tr>
<tr>
<td>Nett and Huang(^\text{20})</td>
<td>100</td>
<td>5.0</td>
<td>3.6</td>
<td>3.0</td>
<td>4.1</td>
<td>3.9</td>
<td>1.8</td>
<td>NA</td>
<td>21.4</td>
<td>NA</td>
</tr>
<tr>
<td>Yang-Powers et al.(^\text{21})</td>
<td>32</td>
<td>7.3</td>
<td>5.1</td>
<td>7.9</td>
<td>2.6</td>
<td>2.5</td>
<td>3.2</td>
<td>NA</td>
<td>28.6</td>
<td>NA</td>
</tr>
<tr>
<td>Knierim et al.(^\text{22})</td>
<td>437</td>
<td>3.5</td>
<td>3.7</td>
<td>4.1</td>
<td>2.8</td>
<td>4.6</td>
<td>3.3</td>
<td>0.8</td>
<td>22.0</td>
<td>22.8</td>
</tr>
<tr>
<td>Wes Fleming et al.(^\text{23})</td>
<td>138</td>
<td>5.2</td>
<td>4.5</td>
<td>4.6</td>
<td>2.6</td>
<td>6.3</td>
<td>1.7</td>
<td>NA</td>
<td>24.9</td>
<td>NA</td>
</tr>
</tbody>
</table>

\(\text{Current study: actual posttreatment models} \quad 27\)

\(\text{Current study: ClinCheck models} \quad 27\)

\(\text{Average} \quad 14.0 \quad 14.0\)

\(\text{Median} \quad 24.0 \quad 24.0\)

\(\text{Max} \quad 28.8 \quad 29.5\)

\(\text{Min} \quad 14.0 \quad 14.0\)

\(\text{Range} \quad 4.0 \quad 4.5\)

\(\text{NA indicates not applicable.}\)
CONCLUSIONS

Within the limits of this study, it can be concluded that:

- ClinCheck models do not accurately reflect the patients’ final occlusion immediately at the end of active treatment.
- Compared with the patients’ models taken immediately after treatment, the ClinCheck models overestimated alignment, buccolingual inclinations, occlusal contacts, and occlusal relations.

ACKNOWLEDGMENTS

This research was partially funded by Align Technology, Inc, San Jose, CA, USA.

REFERENCES


16. Cook MK. Evaluation of Board-certified Orthodontist’s Sequential Finished Cases with the ABO Objective Grading System *Orthodontics*. St Louis, Mo: St Louis University; 2003.


