CURRENT TREND IN RESEARCH ABOUT INVISALIGN® SUCCESS OUTCOME: A SYSTEMATIC REVIEW USING PRISMA GUIDELINE

Akhyar Dyni Zakyah*, Avi Laviana**

* Dentistry Division Faculty of Medicine, Universitas Sriwijaya
** Department of Orthodontic, Faculty of Dentistry, Universitas Padjadjaran

Correspondence: akhyardyni@fk.unsri.ac.id

Keywords: ABO-MGS; Invisalign®; orthodontic treatment; removable clear aligner.

ABSTRACT

Background: While some researches have claimed that Invisalign® is successful in creating meaningful tooth movement, there is still no comprehensive literature review that summarizes them using one valid and reliable parameter. American Board of Orthodontics-Model Grading System (ABO-MGS) is an objective measurement for the success of orthodontic treatment. This systematic review is determined to provide scientific pieces of evidence to prove the truth of the claim using ABO-MGS.

Method: English-based health sciences journal databases were searched using "Invisalign" and "Model Grading System". The databases included in this study were Pubmed, Ebscohost, Google Scholar, ScienceDirect, and MEDLINE. Inclusion criteria were clinical cross-sectional, cohort, or case-control study using human subjects who finished Invisalign treatment. Two independent authors summarized the data from the obtained articles using predefined data fields and discussed the data result together.

Result: Out of 60 studies found during the identification process, only 6 studies were included in the qualitative analysis. Only one of them studied the efficacy of Invisalign® by using ABO-MGS to measure the pre- and post-treatment model. Invisalign® mostly successful in creating change for alignment, overjet, and interproximal contact. Other categories were not that successful to be treated with Invisalign.

Conclusion: Invisalign®’s success is mostly pursued by correcting the malocclusion in the anterior region. Given the scarce amount of reliable evidence available, it is suggested that more studies are required to be able to draw a further conclusion.

INTRODUCTION

The ideal alignment of teeth is desired to improve dental function, eliminate traumatic occlusion, and create a better appearance. One of the treatments of choice to achieve that is by orthodontic treatment. However, the traditional orthodontic appliance makes people spend one or two years with a decrease in their smile’s
cosmetic appeal. Owing to the appearance of the treatment appliance itself. That’s one of the main reasons clear aligners are in high demand.¹

Since the launch of Invisalign® in 1999, the system (Align Technology, Santa Clara, Calif) has become a popular treatment choice for clinicians because of the esthetics and comfort of the removable clear aligners compared with traditional appliances. The system utilizes 3-dimensional graphic imaging and computer-aided design/computer-aided modeling (CAD-CAM) techniques to fabricate aligners for the patient.² It can accurately fabricate numerous aligners to move teeth with relative precision to provide comprehensive orthodontic treatment. However, concern regarding the success of the treatment is still disputed among clinicians.³

Invisalign® was developed to be used as an orthodontic treatment alternative for adults with a Class I malocclusion with mild-to-moderate crowding. Caution should be taken when dealing with malocclusions that have more than 5 millimeters of spacing and crowding, skeletal anteroposterior discrepancies of greater than 2 mm, centric relation and occlusion discrepancies, teeth rotations of greater than 20 degrees, anterior and posterior open bites, teeth extrusion, teeth tipping of greater than 45 degrees, teeth with short clinical crowns and arches missing multiple teeth.⁴ As can be seen from these references, there is controversy about the complexity of orthodontic cases that can be treated successfully with Invisalign®.

The American Board of Orthodontics has established a Model Grading System (ABO-MGS) to evaluate the final dental casts and panoramic radiographs. This was done to assists orthodontists with a tool to assess the adequacy of their finished orthodontic results. It had been used to grade the treatment outcome of orthodontic appliances since 1999. The casts are scored in 7 categories (alignment, marginal ridges, buccolingual inclinations, occlusal relationships, occlusal contacts, overjet, and interproximal contacts), and panoramic radiographs are scored for root angulation.⁵ Since the parameters are highly accessible and globally accepted, clinical trials of the Invisalign® treatment outcome should be measured using the ABO-MGS.

Although many studies have been published that assess the efficacy or the outcome of the treatment, it was a disadvantage that most of them have different aspects measured in the study. Thus, making it hard to understand the success definitive of the treatment or to compare them with one another. This study aims to translate the researches using the parameters in the ABO model grading system to understand the current trend on research regarding Invisalign® treatment outcome. To give insight into what might be considered a successful treatment that should be pursued in clinical practice.

LITERATURE REVIEW

Method

This systematic review was conducted using the PRISMA statement. One focused population, intervention, comparison, and outcome question were delivered according to the PRISMA guideline. The question addressed the outcome of Invisalign which was measured using the ABO-MGS. The question is “Can
orthodontic treatment using Invisalign® achieve a successful treatment outcome based on ABO-MGS standard?"

Health sciences journal databases were searched using "Invisalign" AND "ABO-MGS" or "Model Grading System". The databases included in this study were PubMed, Ebscohost, Google Scholar, ScienceDirect, MEDLINE, and Cochrane Database of Systematic Reviews. Studies were also searched from the reference list of the obtained papers. Because of the scarcity of literature available, no publication year limitation was applied.

Inclusion criteria were English-based clinical cross-sectional, cohort, or case-control study using human subjects with full-text available published on reputable scientific journals and used the ABO-MGS guideline to measure the treatment success in a minimum of 7 out of 8 aspects. Exclusion criteria were reviews, case reports, commentaries, letters to the editor, and studies that only used partial aspects of ABO-MGS without using the scoring guideline were excluded. Populations were patients treated with Invisalign and finished the treatment within the allocated time. Extracted and non-extracted cases were welcomed. The cast must be taken on the day the treatment was considered done so as not to confuse the result with the outcome of the retention period.

The authors of this review developed an extraction sheet to collect data from the obtained papers. One extracted the data and another one checked the data to prevent numerical mistakes. If the data was unclear, contacts were made to the author of the original paper to settle the confusion. Two independent authors summarized the data from the obtained articles using predefined data fields and discussed the quality assessment data result together. Disagreements were resolved by discussion between the two authors.

Included on the data extraction sheet were; 1) Study characteristics (age, sex, socioeconomic background, how long the Invisalign treatment was going, and whether the patient was subjected to extraction or not), 2) Inclusion and exclusion criteria, and 3) ABO-MGS score from each criterion (alignment, marginal ridges, buccolingual inclinations, occlusal relationships, occlusal contacts, overjet, interproximal contacts, and root angulation). Data were presented in Tables 1, 3, and 4.

The Joanna Briggs Institute Prevalence Critical Appraisal Tool (JBI) for Studies Reporting Prevalence Data was used to assess the quality of the obtained studies. This tool was chosen because the obtained studies were epidemiological studies without any intervention was given or the control group presented. The quality assessment was presented in Table 2. A study was considered to have a low-quality assessment if 0-5 criteria were met, and high-quality assessment if studies met 5-10 criteria. Summary measures revolved around the score of ABO-MGS of Invisalign treatment outcome. Data were analyzed qualitatively and quantitatively.

Result

Of all 6 databases searched by both authors, 60 studies were yielded using "Invisalign" AND "ABO-MGS" or "Model Grading System" as keywords. Twenty-three papers were excluded because they had duplicates. Thirty-
seven studies were screened by both authors and 28 of them were then excluded. The exclusion, because the paper was written in a language other than English and study methodology, was a case report or review. For the remaining 9 papers, full-text was obtained and assessed to ensure that the study fulfills the inclusion and exclusion criteria. The total of studies included for qualitative and quantitative analysis was 6.\textsuperscript{6-11} The flow diagram of the PRISMA protocol was presented in Figure 1. Studies measuring the treatment outcome of Invisalign using ABO-MGS was done as early as in 2002. From 6 journals that were analyzed, only one study measured the efficacy of Invisalign\textsuperscript{®} by using ABO-MGS to measure the pre- and post-treatment model. Three studies compared the treatment outcome of Invisalign\textsuperscript{®} and braces with ABO-MGS. One study compared the ABO-MGS score from the last day of treatment and the last day of retention period between Invisalign and braces. Another one compared the treatment outcome of Invisalign between the predicted model using ClinCheck and the achieved model. Most of the studies were conducted in the USA (n=5, 83\%) and the rest were conducted in China (n=1, 17\%). Subjects were mostly female (n=93, 50.5\%) with two studies that did not disclose the female and male subject ratio. The youngest age of subjects was 32.5 y.o. and the oldest was 37.05 y.o. ± 9.2. The period of Invisalign treatment was around 12 ± 3.5 months to 36 months. Data were presented in Table 1.
Figure 1. PRISMA flow diagram for the review process
The main finding in this study is the ABO-MGS score of the Invisalign treatment outcome. The lowest recorded score was 23 and the highest was 45.36. Since the score of root angulation was not present in some studies, the highest ABO-MGS score would yield a result of 41.8 if the root angulation score were to be excluded from the calculation. The lowest score for alignment, marginal ridge, and buccolingual inclination was 4, 1.38, and 2.38 respectively. The highest score for them was 7.56, 5.45, and 6.26. The lowest score for occlusal contact, occlusal relationship, overjet, and interproximal contact were 3, 4, 2.56, and 0. The highest score for them was 10.46, 10.26, 7, and 0.77. Only 4 studies measured the score for root angulation, of that the lowest score was 0.58 and the highest score was 3.56.

A total of 214 cases were studied. Around that number 27.6% receive the passing grade, 3.7% as borderline, and 39.3% were decided as failed. Since not every author mentioned the decision of the measurement, 29.4% could not be determined as pass or fail. The lowest passing percentage was 3.22% and the highest was 66.7%. The detailed ABO-MGS score as explained in Table 2.
Table 2. ABO-MGS score

<table>
<thead>
<tr>
<th>No</th>
<th>Author(s), et al.</th>
<th>Sample size</th>
<th>Alignment</th>
<th>Marginal ridge</th>
<th>Buccal lingual inclination</th>
<th>Occlusal contact</th>
<th>Occlusal relationship</th>
<th>Overjet</th>
<th>Interproximal contact</th>
<th>Root angulation</th>
<th>Case decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Kassa, et al.</td>
<td>31</td>
<td>6.0 ± 2.0</td>
<td>6.26 ± 3.78</td>
<td>6.71 ± 1.51</td>
<td>10.26</td>
<td>4.90 ± 3.58</td>
<td>0.06 ± 0.36</td>
<td>10.26 ± 6.84</td>
<td>4.90 ± 3.72</td>
<td>P = 1, B = 8, F = 22</td>
</tr>
<tr>
<td>2</td>
<td>Li, et al.</td>
<td>72</td>
<td>4.35 ± 1.81</td>
<td>3.70 ± 1.51</td>
<td>4.25 ± 1.01</td>
<td>4.35 ± 1.44</td>
<td>3.83 ± 0.15</td>
<td>0.15 ± 0.15</td>
<td>2.05 ± 2.05</td>
<td>2.05 ± 1.58</td>
<td>P = 48, B = - , F = 24</td>
</tr>
<tr>
<td>3</td>
<td>Buschang, et al.</td>
<td>27</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>5.0</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Djeu, et al.</td>
<td>48</td>
<td>7.56 ± 3.36</td>
<td>4.9 ± 2.55</td>
<td>4.19 ± 2.73</td>
<td>10.46 ± 7.06</td>
<td>7.71 ± 4.76</td>
<td>6.21 ± 4.64</td>
<td>0.77 ± 1.39</td>
<td>3.56 ± 2.35</td>
<td>P = 10, B = - , F = 38</td>
</tr>
<tr>
<td>5</td>
<td>Kuncio, et al.</td>
<td>11</td>
<td>5.91 ± 4.09</td>
<td>5.45 ± 2.50</td>
<td>3.45 ± 2.07</td>
<td>8.27 ± 4.24</td>
<td>6.73 ± 4.64</td>
<td>7.00 ± 3.79</td>
<td>0.55 ± 1.21</td>
<td>2.00 ± 1.48</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Robins, et al.</td>
<td>25</td>
<td>4.36 ± 2.26</td>
<td>1.38 ± 1.28</td>
<td>2.38 ± 1.75</td>
<td>8.76 ± 5.10</td>
<td>4.52 ± 4.62</td>
<td>2.56 ± 2.89</td>
<td>0.58 ± 1.48</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

P = pass, B = borderline, F = fail

The risk of bias in selected studies was measured using JBI-CAT. Five of them were deemed of having high-quality study and one has low quality. The detailed criteria were presented in Table 3. Most of the potential bias resulted from the failure to identify confounding factors such as sex, age, and socioeconomic background. Most of the studies also did not mention the drop-out rate or the percentage of sample who did not meet the inclusion or exclusion criteria. Invisalign might give a slightly different result between extracted and non-extracted cases. Some studies did not specify this matter. Nor was the detailed time interval between aligner and how many aligners were used by subjects during treatment. Also, the use of additional retention or technique was not mentioned in the studies. This could result in a different number of forces generated by the appliance, thus affecting the result of treatment. This might create a reporting bias within studies.

Since the authors of this study can only process articles written in English, language publication bias might occur here. There were
some unobtainable studies due to the nature of the journal which the authors found might give a meaningful contribution to this study. This might create bias where some related papers might not have been identified.

Table 3. Risk of bias within selected studies assessed using JBI-CAT

<table>
<thead>
<tr>
<th>No</th>
<th>Author</th>
<th>Was the sample representative of the target population?</th>
<th>Were the participants recruited appropriately?</th>
<th>Was the sample size adequate?</th>
<th>Was the study subjects and setting described in detail?</th>
<th>Was the data analysis conducted with sufficient coverage of the identified sample?</th>
<th>Was the objective, standard criteria used for the measurement of the condition?</th>
<th>Was there an appropriate statistical analysis?</th>
<th>Are all important confounding factors/subgroups/differences identified and accounted for?</th>
<th>Were subpopulations identified using objective criteria?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Kassas, et al.</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>2</td>
<td>Li, et al.</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>3</td>
<td>Busch, et al.</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>4</td>
<td>Djeu, et al.</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>5</td>
<td>Kuncio, et al.</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>6</td>
<td>Robins on WL</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

Y = yes, N = no, U = unclear

DISCUSSION
The proper way to research the Invisalign success outcome is by following the guideline of ABO-MGS. The pre-treatment model must be measured using Discrepancy Index and the post-treatment model must be measured using ABO-MGS. Two studies that went through this procedure were Kassas, et al. (2013) and Djeu, et al. (2005). Kassas et al. (2013) reported a study done on 31 subjects in New York. Before the treatment, the discrepancy index is 13.03 (SD
2.46) points. The post-treatment measurement showed a statistically significant result in alignment, buccolingual inclination, and total MGS score. There was also an increase in occlusal contact and occlusal relationship score. Interproximal contact didn't change. Djeu et al. (2005) conducted a study on 48 subjects. The mean DI score obtained was 18.67 while the OGS was 45.35 from 8 categories. The treatment improved alignment, interproximal contact, marginal ridges, and root angulation.8

Alignment can be determined by first setting the guidance. On the upper arch, alignment was measured at the lingual aspect of the anterior teeth and the incisal portion. For the posterior region alignment was measured by the line created by the mesiodistal central groove of the premolars and molars. On the contrary, on the lower arch alignment was measured at the labial aspect of the anterior teeth and buccal cusps of the lower premolars and molars.12

To determine whether the teeth are in proper vertical position marginal ridges are used and measured. The proper vertical position of teeth in the dentition is considered adequate when the marginal ridges of adjacent teeth are at the same height. This will result in better occlusion since marginal ridges provide contact areas for the cusps of the counterpart teeth. The condition will establish proper occlusal contacts for the dentition. To achieve ideal marginal ridges height, the orthodontic appliance must have vertical control during teeth alignment.13

The buccolingual inclination is a parameter used to measure the angulation of posterior teeth in both arches. If the inclination is adequate, proper occlusion may be achieved during intercuspation and there will be a decreased risk of occlusal interferences. An adequate buccolingual inclination is when there is a balanced height of the buccal and lingual/palatal cusps of the premolars and molars of upper and lower arch.14

To determine whether the posterior teeth are in proper occlusion, occlusal contacts are used and measured. The epitome of orthodontic treatment is the ideal maximum intercuspation when teeth are in centric occlusion. To achieve that, there should be no space between occluding posterior teeth during centric occlusion.1

An occlusal relationship is a parameter to measure the sagittal relation of teeth in the upper and lower arch. The relationship is based on Angle’s classification. The mesiobuccal cusp of the maxillary first permanent molar must align within 1 mm of the buccal groove of the mandibular first permanent molar. This is also one of six keys of ideal occlusion as mentioned by Andrew (year).3

To determine the lateral relation of posterior teeth and the sagittal relation of anterior teeth, overjet was used. In the anterior region, overjet is measured between the incisal tip of the maxillary central incisor and the labial aspect of the mandibular central incisor. In the posterior region, the mandibular buccal cusps and maxillary lingual cusps are used to determine proper position within the fossae of the opposing arch.12

Interproximal contacts are used to assess the sagittal relation of the posterior teeth and the transversal relation of the anterior teeth. By the end of the orthodontic treatment, there should be no space left between teeth. This was to ensure
that the surrounding periodontal tissues will be healthy after the treatment was ended.\textsuperscript{15}

The last parameter to be measured is the root angulation. It is used to determine the adequate position of teeth roots in the dentition. The measurement can only be done using radiographic aid. Usually, the panoramic radiograph is used for this purpose. The adequate angulation is measured by the amount of bone present between tooth roots. The importance of alveolar bone for orthodontic treatment lies in the need for bone support during the retention period.\textsuperscript{13}

From 6 journals that were analyzed, only one journal studied the efficacy of Invisalign\textsuperscript{®} by using ABO-MGS to measure the pre and post-treatment model.\textsuperscript{6,10} Two journals compared the treatment outcome of Invisalign\textsuperscript{®} and braces with ABO-MGS.\textsuperscript{8,9} The rest of them use some parameters from ABO-MGS to measure the outcome of the treatment. This proves the background of the study, stating that there was a lack of clinical study measuring the outcome of the Invisalign\textsuperscript{®} treatment using a standardized measurement such as ABO-MGS (Table 1).

Subjects included in each study varied from 11-50 people. Too little sample might be because of the decision of the authors to collect data from a single dental practitioner office. This was done to prevent bias from the difference in the practitioner's skill when treating the subject. However, a bigger sample may give a more accurate conclusion to the study.\textsuperscript{3} Female to male ratio was uneven, with subjects mostly are female. This was because female patients were more attracted to the cosmetic appeal of orthodontic appliances.\textsuperscript{2} Age ranging from 15-63 years old by the time subjects started treatment. This diverse range affected the result because patients still in puberty may show a more progressive change in teeth movement than older patients.\textsuperscript{16}

Of all 8 categories mentioned in ABO-OGS, most of the author measured the overjet and overbite for post-treatment results (6 journals, 100%) as mention in Table 3. Next was alignment, marginal ridge, occlusal contact, and interproximal contact (4 journals, 67%). Most patients sought orthodontic treatment because of apparent malocclusion they saw, mostly in anterior teeth. That was the reason behind most clinicians prefer measuring the change in the anterior region to measure the successful outcome of their treatment.\textsuperscript{17}

The least category measured was root angulation (2 journals, 33%). This might be because most of the studies only use the dental study model and didn't use panoramic radiography in assessing the treatment outcome. Even though a radiograph is probably the best practical means to assess the health of the periodontal tissue of the patient. If roots are properly angulated, then sufficient bone will be present between adjacent roots, which could be important if the patient were susceptible to periodontal bone loss at some point in time.\textsuperscript{18}

Of all the categories measured, and only a minor portion of them showed a statistically significant change that correlates with the success of Invisalign\textsuperscript{®} treatment outcome. Invisalign\textsuperscript{®} mostly successful in creating change for alignment, overjet, and interproximal contact. The score for each is 67%, 33%, and 33% respectively. The result is parallel with the
outcome said by Invisalign® manufacturer, Align Technology. The removable aligners are known to consistently produce adequate space closure of up to 6 mm by progressively tipping teeth into spaces in small increments. In terms of alignment, Invisalign has also had success with straightening arches by derotating teeth, especially when composite attachments are bonded to premolars.

Other categories were not that successful to be treated with Invisalign. Certain types of tooth movement, such as extrusion, may be difficult with Invisalign, which probably makes adequate occlusal contacts difficult to achieve using aligners. Besides, the thickness of the aligners over the occlusal surfaces of the teeth might interfere with the settling of the occlusion.

However, patients may still prefer Invisalign® treatment, regardless of treatment outcome, due to improved aesthetics, reduced treatment time, and the ability to remove the appliance during meals and while performing oral hygiene. Because scientific evidence alone should not automatically dictate the selection of the treatment by the health professional, those making health care decisions should consider the values of not only the health care professional but also the patient. All these factors should be evaluated to determine whether the intervention benefits are worth the associated costs. The fact that each clinician is responsible for the treatment results achieved in every patient makes it important to conduct more clinical studies.

Conclusion

The current trend in research about the achievement of Invisalign® is done in retrospective study and success is mostly pursued by correcting the alignment, interproximal contact, and overjet in the anterior region. Overall, the provided data suggests using Invisalign® to correct complaints in the anterior region. Given the scarce amount of reliable evidence available, it is suggested that more studies are required to be able to draw a further conclusion. Scientific evidence alone should not automatically determine the selection of the treatment option. Many factors should be evaluated to determine whether the intervention benefits are worth the associated costs. Such factors, such as the cost of the treatment and the aesthetic concerns of the patients, should be factored into the treatment considerations to provide comprehensive patient-centered and evidence-based care.

ACKNOWLEDGEMENT

None.

CONFLICT OF INTEREST

None.

References
11. Robinson WL. The Effectiveness of Invisalign treatment as assessed by the Par Index and the American Board of Orthodontics standards. 2002;