

Does initial malocclusion predict the outcome of orthodontic treatment?

Keyvan Sohrabi

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science in Dentistry

University of Washington

2016

Committee:

Geoffrey Greenlee

David Turpin

Charles Spiekerman

Christopher Riolo

Program Authorized to Offer Degree:

Orthodontics

©Copyright 2016
Keyvan Sohrabi

University of Washington

Abstract

Does initial malocclusion predict the outcome of orthodontic treatment?

Keyvan Sohrabi

Chair of the Supervisory Committee:

Dr. Geoffrey M. Greenlee

Department of Orthodontics

Introduction: The assessment of orthodontic treatment outcomes and clinical performance are important goals in orthodontics. This study assessed the improvement of malocclusion in patients treated in University of Washington Graduate Orthodontic Clinic. The relationships between initial severity of the malocclusion, duration of treatment, and final treatment outcomes were also determined.

Methods: Patients who completed treatment during years 2012-2013 were included in this study. Exclusion criteria included complex multidisciplinary patients, phase 1 treatment procedures, and patients with craniofacial problems. A calibrated examiner performed PAR index scoring on initial and final models and ABO Cast-Radiographic evaluation on final records. Patient data collected included treatment duration, Angle classification, type of appliances used, treatment strategy, and cephalometric analysis. Data were analyzed with t-test, ANOVA, and multiple linear regression models.

Results: Records from 102 different patients were included in this study. The mean age \pm SD of subjects was 16.4 yrs \pm 9.9 mos. The following types of malocclusion were included: 43% Class I, 33% Class II, and 24% Class III. Average treatment duration for these patients was 23.2 \pm 5.5 months. Results showed an overall 80 \pm 23% improvement in PAR. Of all patients treated, 48% were greatly improved, 46% improved, and 6% of the patients did not improve. The analysis showed that the average ABO score for patients after completion of treatment was 19.5 \pm 8.8. The ABO component garnering the greatest points was occlusal relationship (3.8 \pm 2.9) and lowest scoring area was found to be interproximal contacts (0.5 \pm 1.2). A positive and significant ($P<0.0001$) relationship was found between initial and final PAR score and the ABO score. A trend was noted for the lower mean final PAR and ABO scores in the extraction group, but this difference between extraction and non-extraction treatments did not reach statistical significance. The mean ABO score for Class I patients was the lowest compared to other categories. The regression model indicated that the initial PAR score and duration of treatment were associated with the ABO score however, age, ANB and SN-MP were not significantly associated with the outcome of treatment.

Conclusion: The majority of patients were treated satisfactorily in the graduate orthodontic clinic. Initial severity of patients and duration of treatment were significantly associated with for the final outcome of treatment.

Introduction

Assessment of treatment outcome in a measurable and objective way should be a goal in orthodontic treatment. It is useful for educational purposes, defines standards, and can help guide clinicians to identify and improve specific areas when shortcomings are noted. The Peer Assessment Rating Index (PAR) is an objective and established method to assess outcomes as well as improvement in orthodontic treatment. PAR has the advantage of quantifying percentage of improvement in treatment (Richmond et al. 1992 a,b) from start to finish. It has also been used extensively throughout the world to quantify treatment need and complexity as well as improvement in public health clinics, teaching hospitals and schools. The American Board of Orthodontics (ABO) developed a detailed grading system in 1998 that evaluates the quality of finish in treated patients in a more detailed way compared to other indices (Casko et al, 1998).

Some teaching clinics have performed treatment outcome studies in the past. At the graduate orthodontic clinic at the University of Illinois at Chicago, Yang-Powers et al. (2002) used the ABO score to evaluate 92 consecutively completed patients. Their findings concluded that 19.6% of their patients would acquire a passing score for ABO. Orthodontists were good at correcting spaces in the arch and deficient in placing adequate torque in the buccal segments (Yang-Powers et al. 2002). Deguchi et al. compared 72 consecutively finished patients from the graduate orthodontic clinic of Okayama University, Japan to 54 patients randomly selected from Indiana University. They used PAR and objective grading system (OGS) to compare the results between the two schools. Based on the PAR index, 89.3% of OU and 89.5% of IU patients were considered to have acceptable results. The percentage of patients with less than 30 points (potentially passing the ABO) was 45.1% for OU and 46.6% for IU. The authors found no correlation between the post-treatment PAR and the OGS (Deguchi et al., 2005).

Difficulty in achieving an ideal occlusion might arise from the pre-treatment occlusion (Richmond et al, 1997; Llewellyn et al, 2007). It is not clear in the literature whether duration of treatment is significantly affected by the initial complexity of the malocclusion being treated. It might seem obvious that patients with more complex malocclusions would take additional time to treat. However, this may not be the case. In addition, the relationship between the pre- and post-treatment PAR score and ABO score has not been established. The aim of the current study was to quantify the clinical improvement of patients treated in the UW Graduate Orthodontic Clinic by PAR and ABO scores and to determine the relationship between the initial severity of malocclusion, duration of treatment and final outcome of patients treated in the UW Graduate Orthodontic Clinic.

Materials and Methods

This study was a retrospective investigation of consecutive cases completed during 2012-2013 at the UW Graduate Orthodontic Clinic. Patients who were recorded as having their orthodontic treatment completed were selected for this investigation. Patient demographic information such as age, clinical diagnosis, treatment plan, and treatment duration were abstracted from the patient charts. Other data captured include the clinical team and the clinical session assigned to each patient. Exclusion criteria were as follows: 1.) Patients with significant craniofacial anomalies or syndromes 2.) Mixed-dentition limited treatment (Phase 1 treatment) 3) Adult patients that underwent multidisciplinary treatment. 4) Patients who were missing pre or post-treatment models or post-treatment panoramic x-rays (required for calculating ABO score).

PAR measurements were completed on pre and post-treatment casts of the patients. Pre and post-treatment PAR scores were recorded by calibrated examiners. After completing different PAR training modules, each examiner was fully calibrated against the gold standard materials (20 pre-treatment and post-treatment casts sent by Steve Richmond, the developer of PAR and director of index calibration courses). ABO Cast-Radiographic evaluations were performed on the post-treatment models and panoramic X-rays. For calibration purposes, the ABO example models were used as well as training sessions by a ABO-certified, experienced faculty member. The overall mean ABO score was acquired as well as scores for each component. All patients were placed in one of three categories based upon their final score: (1) less than 20 (passing score), (2) 20 to 27 (borderline passing) and (3) over 27 (failing).

Data collected included patient demographic information, such as age at start of treatment (reported in years) and the Angle's molar classification. The following cephalometric values were recorded for each patient: ANB (difference between SNA and SNB measurements) and SN-MP (SN to mandibular plane angle GO-GN). In addition to numeric value collected, an ANB of $\geq 6^\circ$ or $\leq -2^\circ$ and SN-MP of $\geq 38^\circ$ or $\leq 26^\circ$ was used to denote patients with skeletal problems. To assess lower incisor proclination, the lower incisor to mandibular plane angle (L1-MP) was recorded, with measurements $\geq 99^\circ$ considered to be excessively proclined. These cephalometric references were adopted from the American Board of Orthodontics (Cangialosi 2004). Cephalometric values were extracted from previously recorded data.

Duration of active treatment was recorded in months. Patients were considered to have started on the banding/bonding date with active treatment continuing until all fixed appliances were removed. The removal of any tooth for orthodontic reasons prior to or during active treatment meant the treatment was considered an 'extraction' case. Each

patient was categorized based on a chart review as follows: 1) Orthognathic Surgery, 2) Exposure, 3) Headgear, 4) Fixed Functional Appliances, 5) Removable Functional Appliances, 6) Expansion with appliance (Hyrax, quad helix, etc), 7) TADs, 8) Elastics, 9) Clear Aligners.

Descriptive statistics were reported as medians, means, and standard deviations. Data were analyzed with T-test, ANOVA, and Spearman Rank Correlation test. A regression model was developed to test the hypothesis of whether the final treatment outcome could be predicted by the initial complexity after adjusting for co-variates.

Results

A total of 182 patients were finished during 2012 and 2013. Among these, 31 patients were multidisciplinary and 37 were a mix of limited treatment patients, Phase I or transfer in from another location (lacking initial records). Three patient records could not be accessed due to an administrative hold on their account. Two patients were further excluded because their models were not stable in maximum intercuspation and the occlusion was unclear, rendering the calculation of PAR and ABO unreliable. Seven patients were missing an initial or final model.

Ultimately, the records of 102 clinical patients were included in this study for statistical analysis. This patient sample had a mean age of 16.4 ± 9.9 years (median 15 years, range 10-75) at the time they started treatment. Of all the patients, 43% had Class I malocclusion, 33% were Class II and 24% were Class III.

After further classifying the malocclusions to their respective subdivisions, 18.6% of patients were categorized as Class II division 1, 3.9% were Class II division 2 and 17.6% were Class II subdivision patients (7% of patients were classified as both Class II division 2 and Class II subdivision). Class III patients were 17% and Class III subdivisions were 7% (Figure 1).

The cephalometric data from the patients included in this study had a median ANB of 3.5° (mean \pm SD of $3.1 \pm 3.2^\circ$). The ANB ranged from -6.2 to 9.7° . Our findings indicate that 20 (19.6%) patients in the current sample had an ANB $\geq 6^\circ$ or $\leq -2^\circ$ (indicating an anterior-posterior skeletal issue). The median SN-MP was 30.2° (mean, $29.6 \pm 7.2^\circ$). SN-MP ranged from 14.9 to 47.7° . Our findings indicate that 36 (35.2%) patients had SN-MP $\geq 38^\circ$ or $\leq 26^\circ$ (indicating a vertical skeletal issue). The median inclination of the lower incisors in relation to lower mandibular border (L1-MP) was 94.4° (mean, $94.5 \pm 8.3^\circ$). L1-MP ranged from 68 to 115° . Our findings indicate that 28 patients had L1-MP $\geq 99^\circ$ (proclined lower incisors).

Average treatment duration for the patients in this sample was 23.2 ± 5.5 months (median 24 months). The range of treatment duration was 10-39 months. Based on our findings, 83% of the patients were treated non-extraction vs. 17% with at least one tooth extracted for orthodontic reasons. Inter-arch elastics were among the most popular reported mechanics used (45%) followed by expansion appliances (14%) and extraoral anchorage (9%). Fixed functional appliances and clear aligners were each used in 8% of the sample. 6% of patients underwent orthognathic surgery (Figure 2).

Dahlberg's formula (Dahlberg, 1940) was used to compute the intra-examiner reliability for ABO scores. Sixteen post-treatment models were measured at baseline and then 3 weeks later. Dalberg error was 2.3, indicating that the amount of error was small enough between the two measurements for the ABO scoring to be considered reliable. To explore if our measurements suffered from systematic vs random error or to see if our measurements suffered from drift, we computed the paired t-test comparing all our first measurements to the second measurements. The result was not significant ($p > 0.05$).

Our analysis shows that the average PAR score before treatment was 26.7 ± 11.7 (median 25). After treatment was complete, average PAR score was 4.3 ± 4.6 (median 3). Percentage Improvement in PAR was $80 \pm 23\%$. Based on percentage improvement of patients, 48% of the patients were greatly improved, 46% improved and 6% did not improve. The average ABO score for patients after completion of treatment was 19.5 ± 8.8 , with a median of 17. ABO scores ranged from 7 to 57. The least amount of points were contributed by the interproximal contacts (0.5 ± 1.2) and root angulations (1 ± 1.2), indicating that these components were treated more ideally. The components contributing more points to the final score were occlusal relationships (3.9 ± 2.9), followed by marginal ridges (3.4 ± 1.8), alignment (3.1 ± 2.4), buccolingual inclination (2.8 ± 2.4), overjet (2.6 ± 2.7), and occlusal contacts (2.3 ± 2.5), indicating the more challenging aspects to achieve an ideal finish. 66% of the treatment results acquired a board score of 20 or less (passing score). 22% of the treatment results acquired a board score between 20 and 27 (borderline score) and 12% acquire a score above 27 (most likely a failing score).

The differences between duration of treatment, ABO score and initial and final PAR score was not significant between clinical sessions and faculty ($p > 0.05$).

The association between initial PAR score and final ABO score was found to be significant ($p = 0.0018$, Spearman's $\rho = 0.31$). This statistical analysis indicated a positive moderate association between these variables: the more complex the patient pre-treatment, the more difficult it would be to treat ideally (Figure 3). There was a weak

positive relationship between initial PAR score (complexity) and duration of treatment but this association was not significant ($p>0.05$, Spearman's $\rho=0.18$). The association between final PAR score and total ABO score was also explored with Spearman rank correlation. The association was significant ($p<0.001$) and the relationship was moderate ($r=0.31$) (Figure 4).

The mean final PAR for the non-extraction group was higher compared to the extraction group (4.4 ± 4.9 vs. 3.8 ± 3.4); the extraction group had a better improvement in PAR, but the difference between the two groups did not reach a significant level ($p>0.05$) (Figure 5). Similarly, to test whether the group of patients that had extractions had better total ABO score, an independent sample t test was conducted. The mean total score for the non-extraction group was 3 points higher when compared to the extraction group and standard deviation was smaller in the extraction group (20 ± 9.2 vs. 17.4 ± 5.8). The extraction group had a better ABO outcome with smaller variation, but the difference did not reach a significant level ($p>0.05$) (Figure 6).

An ANOVA test demonstrated that Angle classification had no impact on the duration of treatment ($p=0.42$). The mean ABO score for Class I patients was the lowest compared to other categories (17.7 ± 6.1). Class III patients had the highest mean ABO score (24.0 ± 11.0). The difference between groups did not reach a significant level, however ($p=0.10$). Similarly, while Class I patients had the lowest mean final PAR score, one way ANOVA didn't indicate a significant difference in PAR outcome across various Angle classifications ($p=0.37$).

The average duration of treatment for non-extraction patients was shorter vs. the extraction patients. The mean non-extraction duration of treatment was 22 ± 5.6 months and 25 ± 4.8 months for those having extractions (mean difference of 3 months in duration of treatment). The independent sample t test for the difference in treatment time was not significantly different however ($p=0.08$) (Figure 7).

To explore variables that might predict the treatment outcome, i.e. ABO score, a regression model was constructed. In the final regression model, total ABO score was the dependent variable (outcome) and its predictors were initial PAR score, duration of treatment, age, ANB and SN-MP (independent variables). The regression model had an F test of $p<0.01$, indicating that the model was significant ($R^2=0.17$). The initial PAR score, a measure of the pre-treatment complexity of the patient was statistically significant ($p=0.002$). The coefficient ($b= 0.25$) was positive indicating that higher initial PAR score is related to difficulty in getting to an ideal finish. Next, the effect of treatment duration ($b=-0.35$, $p=0.047$) was significant and its coefficient negative indicating that the longer the duration of treatment, the better the final ABO scores were. Age of the

patients at start ($b=-0.13$, $p=0.18$) seems to be unrelated to the final treatment finish. ANB angle ($b=-0.10$, $p=0.72$) and the mandibular plane angle ($b=.06$, $p=0.57$) do not seem to predict the final ABO score.

Discussion

Graduate orthodontic programs accredited by the American Dental Association are a good source of data with complete records for research. However, utilizing records from orthodontic programs with several residents and attending faculty creates great variability compared to private practice. Assessing the quality of treatment outcomes can be used to make educational changes in the curriculum.

In this study, the study casts and radiographs of 102 patients were evaluated. For this sample, the mean total ABO score was 19.5 ± 8.8 . Similar to Yang-Powers et al study, this sample had the lowest score on the interproximal section (0.54 ± 1.2). The sample's highest score was in the occlusal relationship (3.9 ± 2.9) which is related to the anterior-posterior position of the dentition. The deficit in the AP correction may be explained by the fact that this component is heavily dependent on the patients' compliance to wear elastics. Overall, the results from both studies show that orthodontists were good at correcting spaces in the arch. Indeed, open interproximal contacts were more easily visualized and treated by clinicians when compared to the occlusal problems. The overall OGS score lost in Yang-Powers study is higher (33.9) (Yang-Powers et al. 2002). This can be partly explained by the major modifications in measurements and scoring in the ABO scoring system. In the new ABO grading, several areas of scoring have been removed, thus it was expected that the new scores would be lower compared to the original OGS (The American Board of Orthodontics Grading System for Dental Casts and Panoramic Radiographs 2012, Casco 1998). A recent retrospective study on 1693 patient records treated at a university setting had a mean ABO score of 18 (Cansunar & Uysal 2014). In another study completed in the University of Washington on retrospective and prospectively selected patients, a mean ABO-OGS score of 16.2 and 23.1 were reported respectively (Struble & Huang, 2010).

After introduction of the PAR index by Richmond et al. (1992), several studies were performed with this index in the UK. Feghali et al. (1996) completed some of the earliest school based studies in the US. In one of their studies conducted in Case Western Reserve University, patient records from 100 subjects were evaluated from start to finish with the PAR index. In this study the mean age of the patients at start was 13 years with an average treatment duration of 32 months. The average pre and post-treatment PAR scores were 34 ± 10 and 11 ± 5 respectively. The authors concluded that 55% of the patients treated at Case Western Reserve were greatly improved, 41% were

improved, and only 4% showed no improvement (Feghali et al., 1995). In the University of Washington sample, 48% were greatly improved, 46% were improved, and 6% showed no improvement. Mean duration of treatment for the sample from Washington was shorter (23 months) compared to the sample from Case Western Reserve. This might be related to improvements in orthodontic materials as well as more efficient treatment mechanics compared to the norm for treatment in the 1990s.

The association between post-treatment PAR and ABO scores has been controversial due to the fact that these two indices measure outcomes differently. In our study, a significant association was found between post-treatment PAR and ABO Cast-Radiographic evaluation via Spearman test ($r=0.31$). As expected, the ABO score was more stringent in assessing treatment outcomes compared to PAR. Eighty-eight percent of the treatment outcomes were considered as acceptable by PAR compared to 66% by ABO score. In a study by Onyeaso and Begole (2007), a moderately positive association was also found between the post-treatment PAR and ABO scores on a sample of 100 patients ($P<0.01$, Spearman $r=0.27$). In that study, the ABO score was also the most stringent in assessing treatment outcomes; post-treatment PAR categorized 97% of the results as acceptable, and ABO categorized 86% as such (Onyeaso and Begole, 2007).

One of the aims of the current study was to determine whether or not there is any difference between the time it takes to treat a malocclusion when the initial malocclusions differ in severity. The timely completion of orthodontic treatment is an expectation of patients and practitioners; however reports in the literature remain elusive whether duration of treatment is significantly affected by the initial complexity of the malocclusion. One would logically think that patients with more challenging malocclusions would take additional time to treat, but this might not be case. This study shows that the increased complexity of care needed by some patients did not lead to significantly longer treatment duration (Spearman's $\rho=0.18$). Since poor patient compliance, missed appointments, broken appliances, and poor oral hygiene can contribute to longer duration of treatment (Mavreas & Athanasiou, 2008) this result should be interpreted with caution. These factors were not assessed in our study. Similarly, although Parish et al. studied a sample of 732 patient records, they found a weak positive correlation between duration of treatment and the discrepancy index (DI) ($r=0.09$) (Parish et al., 2011).

A regression model showed that treatment duration was a significant predictor of the final outcome of treatment, i.e, ABO score. We think this result could be partly due to the time faculty and residents spend on the finishing and detailing stages of treatment to produce board quality patients. Another explanation is that providers discontinue

treatment earlier on patients who are non-compliant or have less optimal outcomes. Treatment is sometimes discontinued when an optimal result is not possible, where the initial treatment objective was limited or due to poor oral hygiene, poor patient compliance or negative sequela such as root resorption, periodontal issues and decalcification.

The relationship between the initial complexity of all patients and the final outcome of treatment was explored in this study. Results showed that a significant positive association existed between the initial complexity of treatment and the final outcome. This association held for both final PAR and ABO scores, indicating that more complex patients are more difficult to treat ideally. The multiple regression model showed that the initial PAR score, together with the treatment duration, are significant predictors of the ABO score. However ANB and SN-MP measurements were not significant predictors of the ABO score. Therefore it seems ANB and vertical skeletal issue may not be great indicators for complexity of the patient's orthodontic treatment. This result may also be due to the fact that ABO score is an occlusal index, derived from model analysis and not directly related to the cephalometric radiographs. Indeed, other relevant studies have also shown that when the cephalometric values for ANB and SN-MP are deleted from the discrepancy index, the association between DI and ABO score is stronger. In the Pulfer et al. (2009) study, ABO OGS scores showed only a weak positive association with DI scores (correlation 0.17), however DI without cephalometric values (ANB and SN-MP) was strongly associated with ABO ($r=0.67$, $P<.0001$).

The results of this study showed that the average ABO score for patients post-treatment was 19.5. Sixty-six percent of the patients got a passing score and 22% achieved a borderline passing score. While this result indicates that the majority of patients were finished with close to ideal results, it should be noted that not all the treatments with passing score are qualified to be presented to the board for two reasons: 1) The current analysis does not take the DI into account, so it is possible that some of the patients might not have a sufficient DI to be qualified. 2) 42% of these patients were transferred between residents, thus disqualifying them from board examination.

In this study, treatment duration was longer in the extraction group by an average of 3 months. This is consistent with other studies (Vig et al. 1998) and is potentially related to the extra time needed to close all spaces. A systematic review states that treatments involving extractions generally take 2-6 months longer than non-extraction treatment (Mavreas & Athanasiou, 2008). A trend was noted for reduced mean final PAR and ABO scores in the extraction group of patients. The extraction group had a 3-point advantage on ABO score over the non-extraction group with reduced overall variation in

outcome. It should be noted that extraction patients are inclined to achieve better results vs. non-extraction patients due to elimination of some tooth structure from the ABO scoring, thus scores are expected to be slightly better.

The current retrospective study has some limitations. Due to the design, blinding of the examiners was not possible. Although the examiners were calibrated, only one examiner ultimately did the measurements. Lack of blinding and a having a single examiner might have introduced systematic measurement bias in the study, favoring lower (improved) PAR and ABO scores in this study. Also, the appliances and techniques information reported in this study were collected from treatment plan sheets and chart notes. This method of data collection might have led to an underreporting of the application of some commonly-employed techniques such as elastics. Due to the consecutive nature of the case selection, the potential for selection bias is low; however the exclusion of patients with improperly trimmed models or those missing models might have biased the results of the study.

Conclusions

- Based on PAR and ABO scores, the majority of patients were treated satisfactorily in the graduate orthodontic clinic.
- Initial severity of patient and duration of treatment were significantly associated with the final outcome of treatment, however age, mandibular plane angle, and ANB were not associated with the final outcome of treatment.
- We observed a trend for extraction strategies to have better occlusal outcomes.
- ABO Cast-Radiographic evaluation was significantly associated with pre and post-treatment PAR scores.

References

Cangialosi TJ, Riolo ML, Owens SE Jr, et al (2004). The ABO discrepancy index: a measure of case complexity. *American Journal of Orthodontics and Dentofacial Orthopedics*;125(3):270-8.

Cansunar HA, Uysal T (2014). Relationship between pretreatment case complexity and orthodontic clinical outcomes determined by the American Board of Orthodontics criteria. *Angle Orthodontist* ;84(6):974-9.

Casko JS, Vaden JL, Kokich VJ, et al (1998). Objective grading system for dental casts

and panoramic radiographs. *American Journal of Orthodontics and Dentofacial Orthopedics*, 114(5):589-99.

Dahlberg G (1940). *Statistical methods for medical and biological students*. London: George Allen and Unwin; pp. 122–132.

Daniels C, Richmond S (2000). The development of the index of complexity, outcome and need (ICON). *Journal of Orthodontics*, 27(2):149-62.

Deguchi T, Honjo T, Fukunaga T, et al (2005). Clinical assessment of orthodontic outcomes with the peer assessment rating, discrepancy index, objective grading system, and comprehensive clinical assessment. *American Journal of Orthodontics and Dentofacial Orthopedics*, 127 (4): 434-443

Feghali R, Nelson S, Afsharpanah A, Hans M (1995). Assessment of an orthodontic clinic sample using the PAR index [abstract]. *J Dent Res*, 74:139.

Fox NA, Daniels C, Gilgrass T (2002). A comparison of the index of complexity outcome and need (ICON) with the peer assessment rating (PAR) and the index of orthodontic treatment need (IOTN). *British Dental Journal*, 193: 225–230.

Llewellyn SK, Hamdan AM, Rock WP (2007). An index of orthodontic treatment complexity. *European Journal of Orthodontics*, 29: 186-92.

Mavreas D, Athanasiou AE (2008). Factors affecting the duration of orthodontic treatment: a systematic review. *European Journal of Orthodontics*, 30:386–395.

Onyeaso CO, Begole EA (2007). Relationship between index of complexity, outcome and need, dental aesthetic index, peer assessment rating index, and American Board of Orthodontics objective grading system. *American Journal of Orthodontics and Dentofacial Orthopedics*, 131: 248-252.

Parrish LD, Roberts WE, Maupome G, et al (2011). The relationship between the ABO discrepancy index and treatment duration in a graduate orthodontic clinic. *The Angle Orthodontist*, 81(2):192-7.

Pulfer RM, et al (2009). The Association of Malocclusion Complexity and Orthodontic Treatment Outcomes. *The Angle Orthodontist*, 79 (3): 468-472.

Richmond S, Shaw WC, Roberts CT, Andrews M (1992). The PAR Index (Peer Assessment Rating): methods to determine outcome of orthodontic treatment in terms of improvement and standards. *European Journal of Orthodontics*, 14 (3): 180-187.

Richmond S, Shaw WC, O'Brien KD, et al. (1992). The development of the PAR Index (Peer Assessment Rating): reliability and validity. *European Journal of Orthodontics*, 14: 125- 139.

Richmond S, Daniels CP, Fox NA, Wright J (1997). The professional perception of orthodontic treatment complexity. *British Dental Journal*, 183: 365-370.

Richmond S, Aylott NAS, Panahei ME, et al (2001). A 2-centre comparison of orthodontists' perceptions of orthodontic treatment difficulty. *Angle Orthodontist*, 71: 404-410.

Struble BH, Huang GJ (2010). Comparison of prospectively and retrospectively selected American Board of Orthodontics cases. *American Journal of Orthodontics and Dentofacial Orthopedics*, 137(1):6.e1-8; discussion 6-8.

Yang-Powers LC, et al (2002). Treatment outcome in a graduate orthodontic clinic using the American Board of Orthodontics grading system. *American Journal of Orthodontics and Dentofacial Orthopedics* ;122(5):451-5.

Vig K W L , Weyant R , Vayda D , et al (1998). Orthodontic process and outcome: efficacy studies-strategies for developing process and outcome measures: a new era in orthodontics. *Clinical Orthodontics and Research* 1:147-155.

<https://www.americanboardortho.com/media/1191/grading-system-casts-radiographs.pdf>

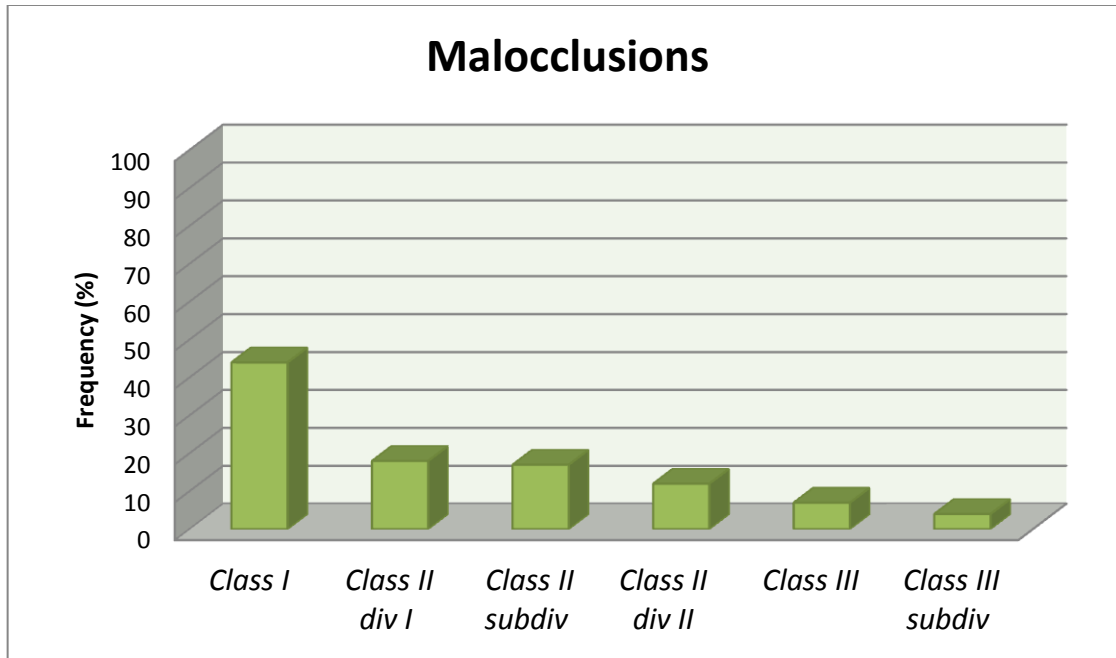


Figure 1. Frequency of malocclusion among included patients

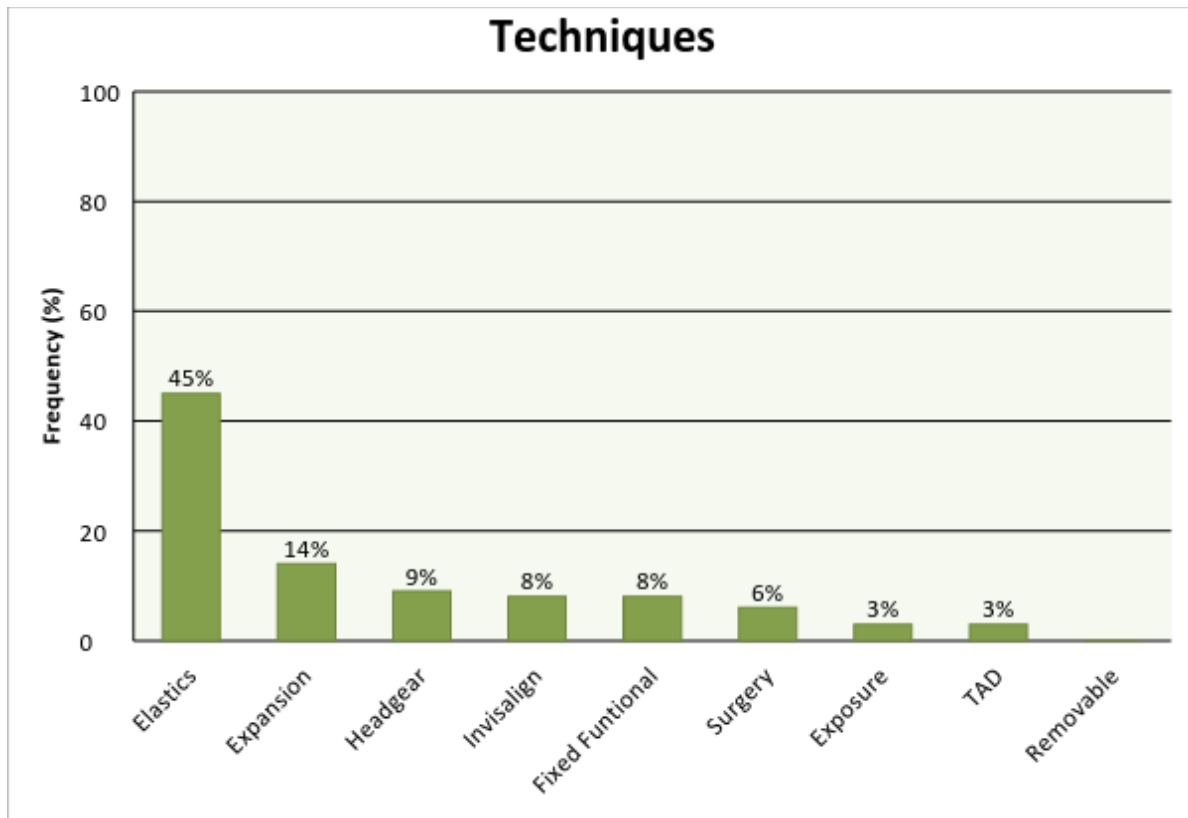


Figure 2. Frequency of techniques and appliances

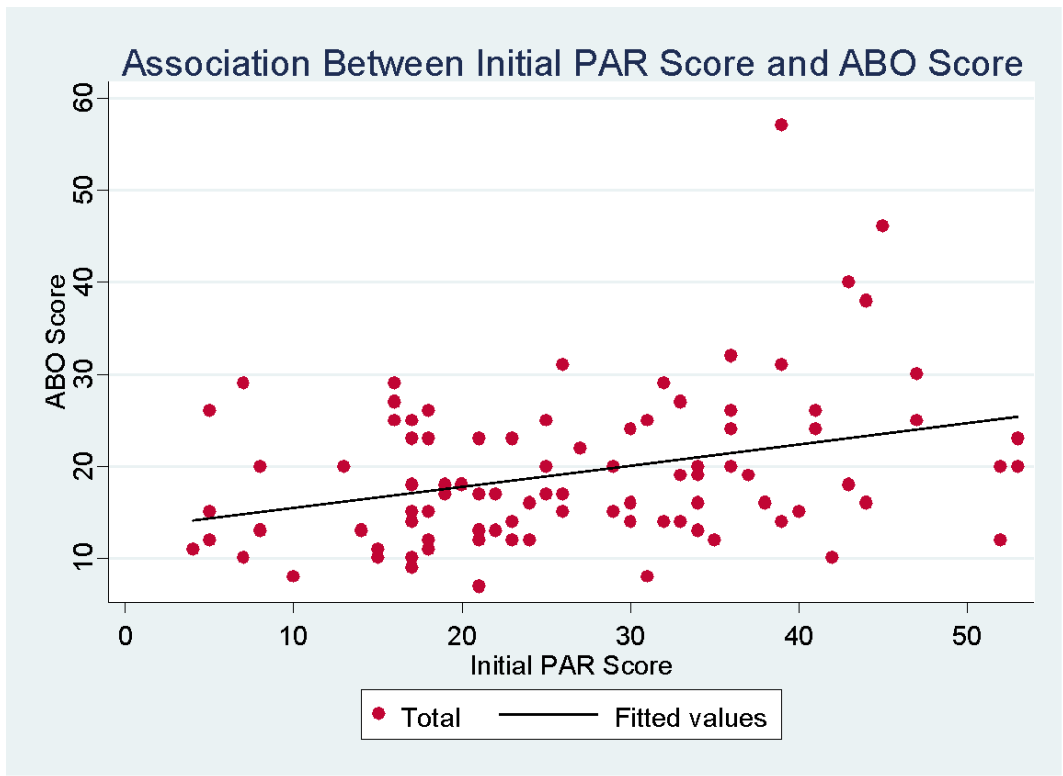


Figure 3. Association between Initial Par Score and ABO score

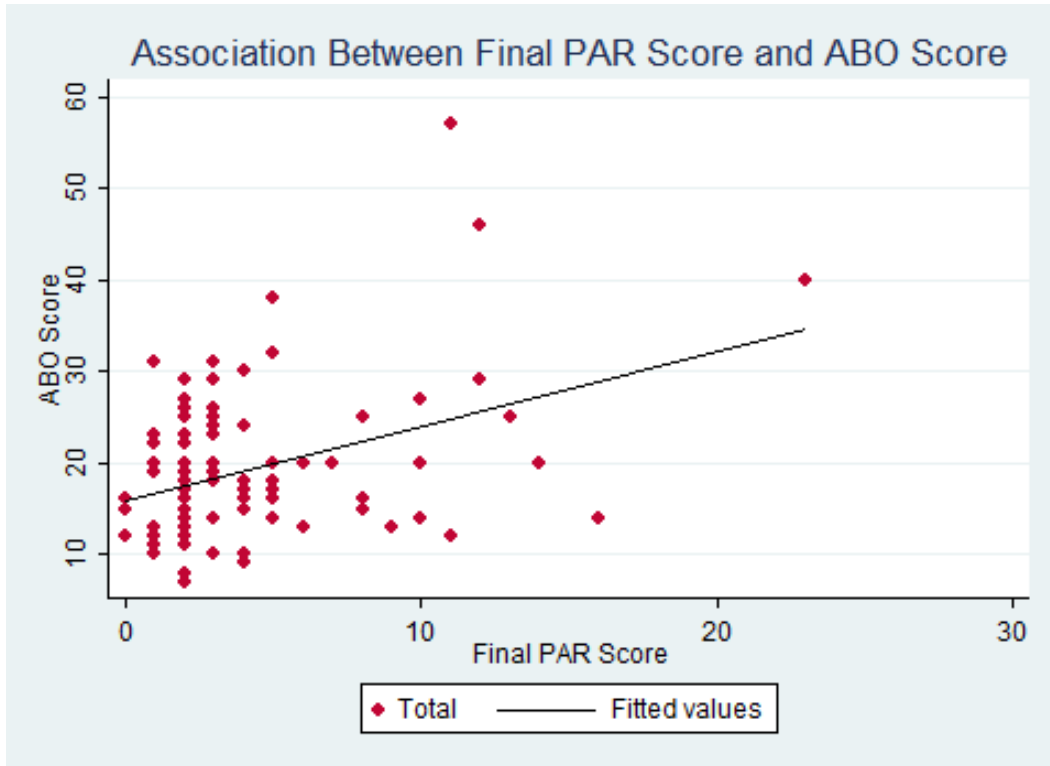


Figure 4. Association between Final Par Score and ABO score

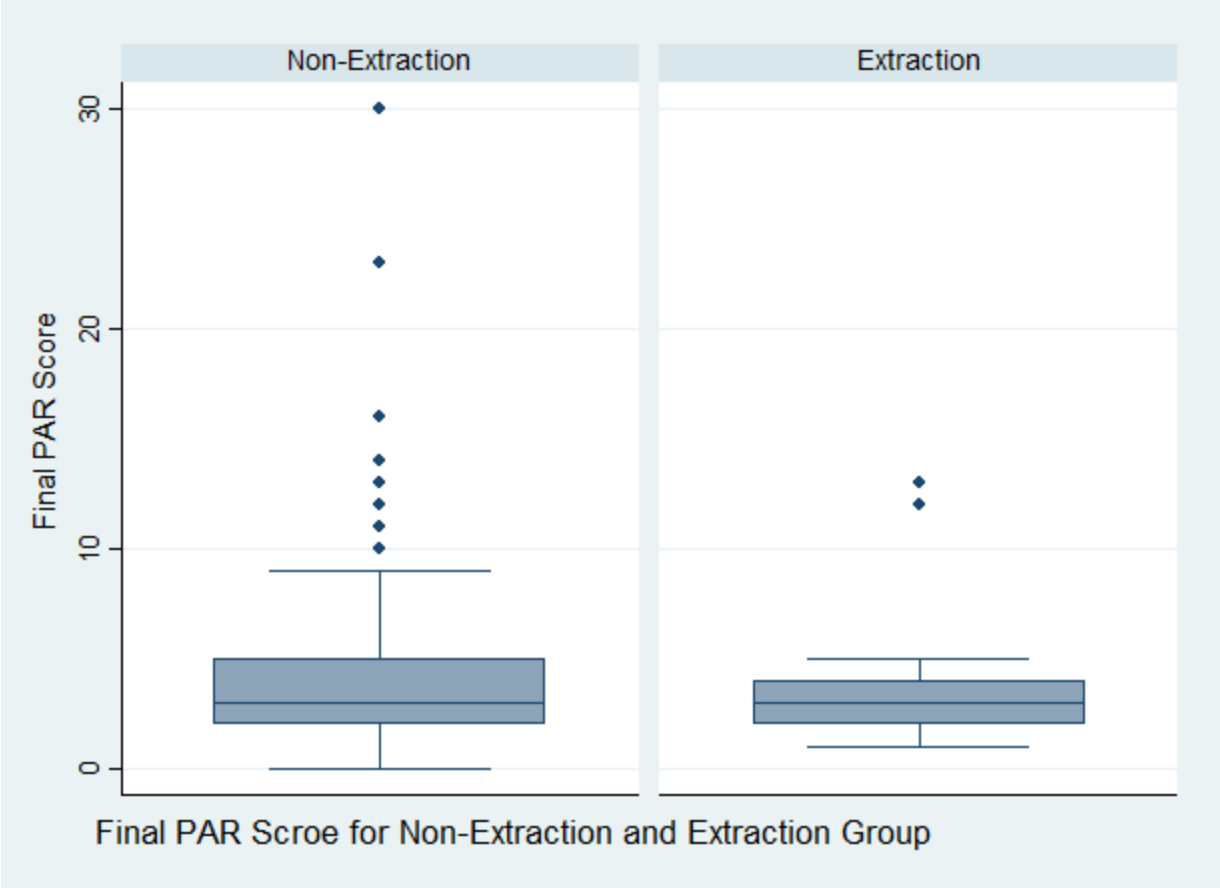


Figure 5. Final Score for Non-Extraction and Extraction Groups

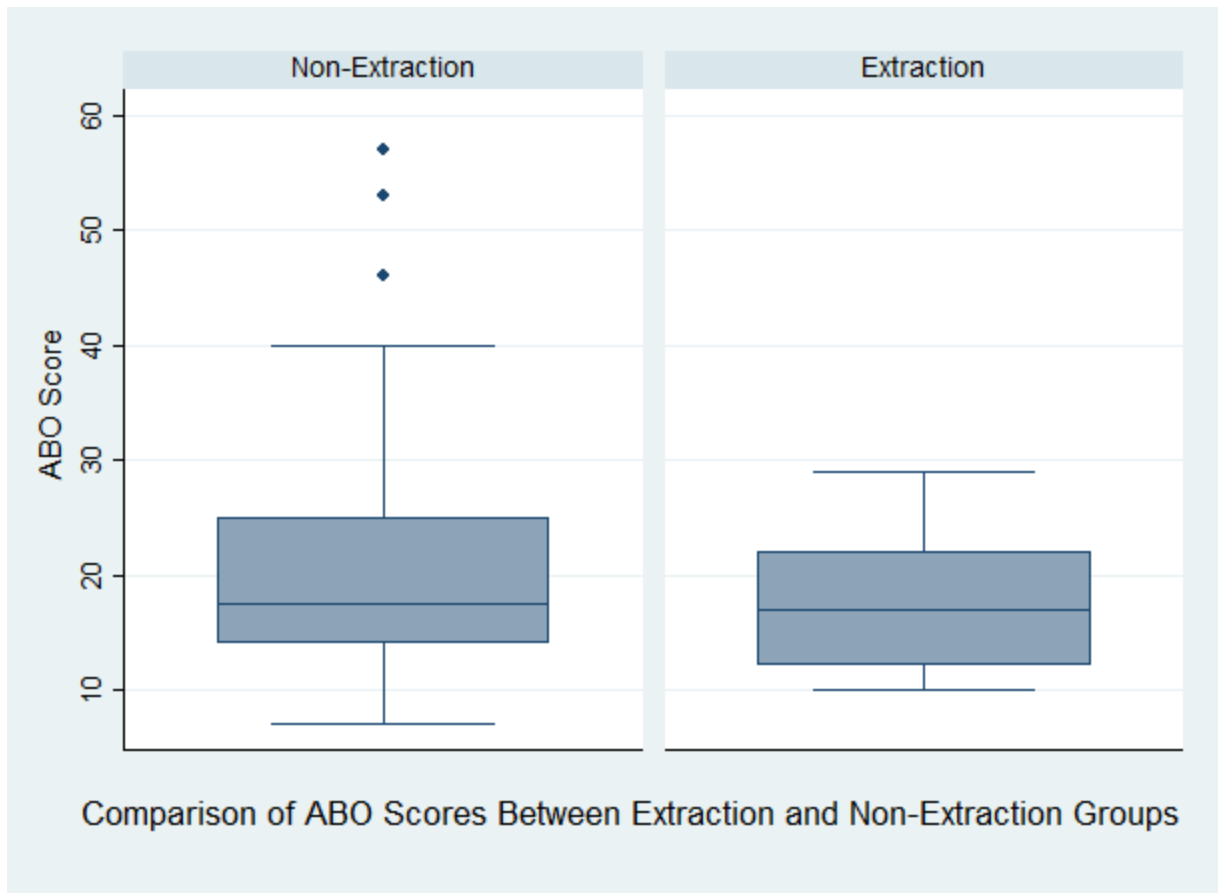


Figure 6. Comparison of ABO scores between extraction and no-extraction groups

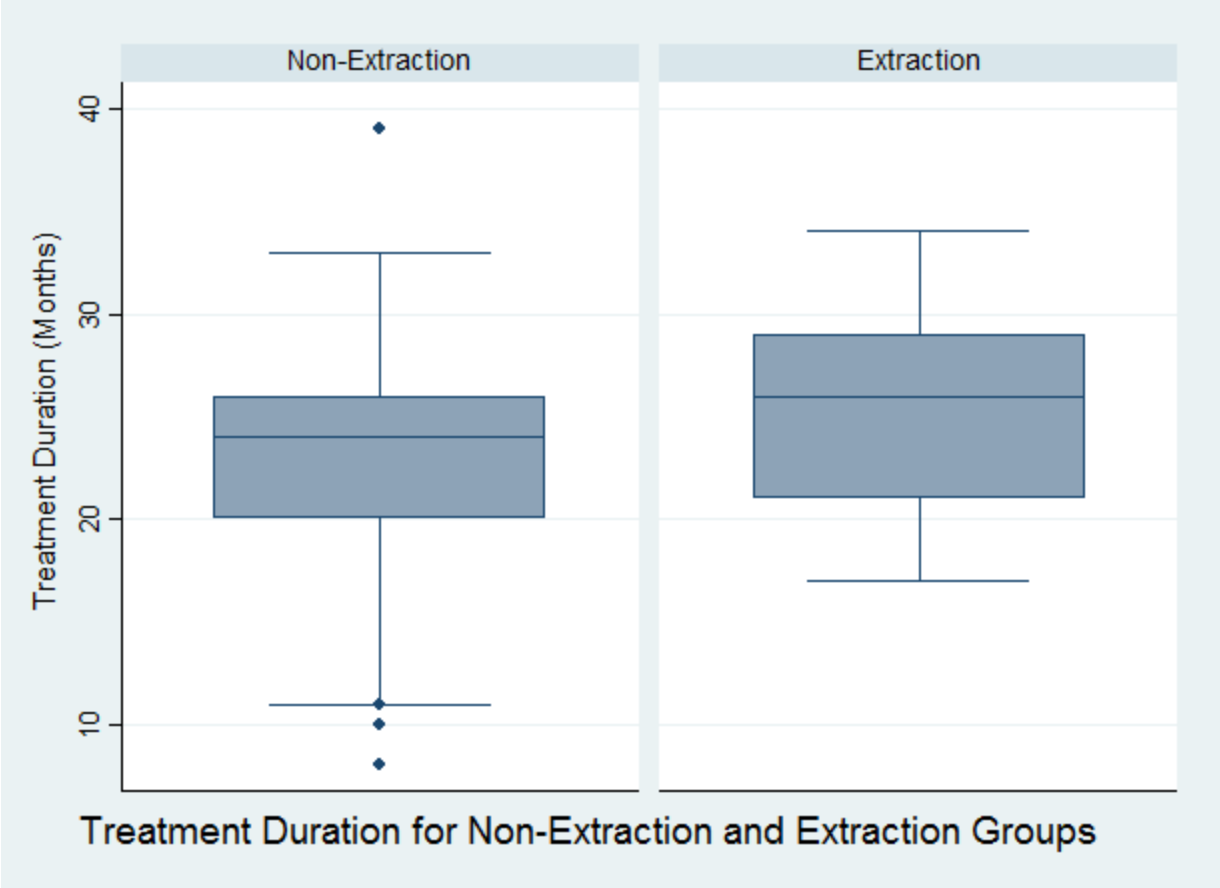


Figure 7. Treatment Duration for non-extraction and extraction groups