OUTCOMES OF ORTHODONTIC TREATMENT PERFORMED BY INDIVIDUAL ORTHODONTISTS VERSUS TWO ORTHODONTISTS COLLABORATING ON TREATMENT

by

Suliman Abdulkarim Alsaeed

B.D.S., King Saud University, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate and Postdoctoral Studies

(CRANIOFACIAL SCIENCES)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2018

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, a thesis/dissertation entitled:

OUTCOMES OF ORTHODONTIC TREATMENT PERFORMED BY INDIVIDUAL ORTHODONTISTS VERSUS TWO ORTHODONTISTS COLLABORATING ON TREATMENT

submitted by Suliman Alsaeed in partial fulfilment of the requirements for

the degree of Master of Sciences

in Craniofacial sciences

Examining Committee:

Dr. David Kennedy
Supervisor

Dr. Edwin Yen
Supervisory Committee Member

Dr. Jolanta Aleksejuniene
Supervisory Committee Member

Dr. Benjamin Pliska
Supervisory Committee Member

Dr. Ya Shen
External Examiner

Abstract

Objectives: The aim of this study was to evaluate orthodontic treatment quality, length and efficiency when two orthodontists collaborated on treatment compared to cases treated by either orthodontist.

Methods: The sample consisted of 150 consecutively treated subjects gathered from three groups of patients (A, B and C), each group included 50 patients. Group A patients were treated by orthodontist A, group B by orthodontist B, and group C by both orthodontists. PAR index, ICON, ABO-DI and ABO-CRE assessed the pre- and post-treatment status. Variables including age, gender, type of malocclusion, extraction versus non-extraction, orthognathic surgery, treatment length, number of visits, frequency of missed, cancelled and emergency appointments were collected for statistical analysis. Treatment efficiency Index (TEI) was also assessed.

Results: There was no statistical significant difference in the pre-treatment status, age, gender, type of malocclusion or number of extractions between the three groups. Post-treatment PAR and ICON indices showed excellent results in all three groups, with no statistical significant difference between groups. ABO-CRE was significantly higher in group C (25.3 points) than either group A (21.5 points) or group B (22.0 points) ($P=0.014$). Group A cases, on average, had significantly less treatment time (23 months) than either group B or C (26 months) ($P=0.011$). Group C patients required more appointments (27 visits) than either group A or B (23 and 25 visits, respectively). The treatment efficiency index showed no statistical significant differences between the three groups ($P=0.113$).

Conclusions: Good outcomes were achieved in all three groups as assessed by PAR index and ICON, with no difference between providers. Cases treated by a collaboration of both orthodontists required 2 to 4 more visits and had higher ABO-CRE scores than those treated by a single orthodontist.

Lay Summary

Traditionally, orthodontic treatment is performed by one orthodontist. However, it is not uncommon that two or more orthodontists collaborate during a patient's orthodontic treatment. This research aimed to measure the quality and length of treatment for patients who were treated by one orthodontist, compared to those who were treated by a collaboration of two orthodontists.

We collected data from a private dental group. Three different indices were used to assess the quality of treatment. We found that treatment quality was good for all cases, whether treated by one or two orthodontists. However, one of the indices showed that the outcome of joint cases was not as good as cases treated by one orthodontist. Joint cases required two to four more visits compared to cases treated by one orthodontist.

Preface

Dr. David Kennedy suggested the research topic of this project. The research sample was gathered from Dr. David Kennedy's private orthodontic office. The methodology was formulated by agreement between committee members Dr. Edwin Yen, Dr. David Kennedy, Dr. Jolanta Aleksejuniene and Dr. Benjamin Pliska. Data collection, entry and analysis was performed by Dr. Suliman Alsaeed. Statistical analysis was guided by Dr. Jolanta Aleksejuniene. The study was approved by the University of British Columbia Office of Research Services, Humans Research Ethics Board (Certificate number: H16-01867)

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ABO: American Board of Orthodontics.

ABO-CRE: American Board of Orthodontics – Cast and Radiographic Evaluation.

ABO-DI: American Board of Orthodontics – Discrepancy Index.

ICON: Index of Complexity, Outcomes and Needs.

PAR: Peer Assessment Rating.

TEI: Treatment Efficiency Index.

Acknowledgements

First and foremost, a special thanks to my research supervisor Dr. David Kennedy for giving me the chance to work with him. His support and guidance gave me invaluable research, clinical and teaching skills that helped me to think deeply and critically. He has been an inspiration to work with, a true mentor and a great role model.

Big thanks to my committee members, Dr.Yen, Dr. Pliska, and Dr. Aleksejuniene for their help, support and guidance. Their valuable comments helped me develop this thesis.

I would like also to thank Dr. Greg Huang and Dr. Stephen Richmond for their help during the calibration process.

Ms. Molly Furlong, thank you for your help searching for models and arranging charts.

Big thanks to my country, The Kingdom of Saudi Arabia, for their academic and financial support during my studies at UBC.

This work would not have been possible without the support of my family. I owe everything in my life to my parents. My father Abdulkarim Alsaeed and my mother Huda Alehaideb have devoted their lives to making me happy and because of them I am the man that I am today.

Dedication

To my wonderful family who has provided me with unconditional love and support. I am eternally grateful for all the words of wisdom, encouragement, and guidance that you have given me in life.

Chapter 1: Introduction

1.1 Overview

Excellent orthodontic treatment outcome is an ultimate shared goal. Many studies have discussed the different factors that might affect treatment quality[1-27]. These factors varied from treatment timing (early vs late)[1-6], method of ligation (conventional vs self-ligating)[7-11], treatment philosophy (extraction vs non-extraction)[12-17], practice setting (private vs educational)[18-20], involvement of orthognathic surgery[21-24], and many other contributing factors[25-27].

One variable that has not been extensively studied is the treatment provider[28,29]. McGuinness and McDonald evaluated whether a change in operator affected treatment length or quality in a postgraduate teaching environment[28]. Sixty pre- and post-treatment orthodontic models were assessed by Peer Assessment Rating (PAR) index scores with half treated by one operator, and the other half started by one operator and finished by another[28]. Cases treated by one operator were finished in a significantly shorter time as compared to those cases started by one clinician and finished by another[28]. However, the final PAR score was not significantly statistically different between the two groups[28]. This indicated that two operators working on the same patient may result in increased treatment time without any sacrifice in treatment quality[28].

Peppers et al. determined whether the primary attending (PA) doctor (clinical instructor) coverage in an educational environment affected treatment length or quality using the PAR index[29]. One hundred and ninety one patients were divided into three groups based on the supervisor coverage frequency, which was defined as high, medium and low[29]. The frequency of PA doctor coverage neither lengthened the treatment significantly nor altered treatment quality[29].

The assessment for orthodontic treatment need identified patients who should be provided with orthodontic treatment[30,31]. In the late 1960s, orthodontists worked to establish indices that provided a more objective assessment of orthodontic treatment need[32-36]. Draker proposed the Handicapping Labio-lingual Deviations Index (HLD Index) in 1960 [32], which was followed by Grainger's Treatment Priority Index (TPI) in 1967 [33], Salzmann’s Handicapping Malocclusion Assessment Record in 1968 [34] and Summer’s Occlusal Index (OI) in 1971 [35]. The American Association of Orthodontists (AAO) adopted Salzmann’s index in 1969 to identify patients who had a “seriously handicapping malocclusion”, as was required by the civilian (Medicaid) and military (Champus) medical programs[34,37]. The AAO officially reversed that action in 1985 and announced that they were against the use of any index to identify orthodontic treatment needs[37,38]. This resulted in a significant reduction in the use of orthodontic indices in North America[39]. The Dental Aesthetic Index (DAI) was subsequently developed in 1986 and then accepted by the World Health Organization (WHO) as a screening tool[40,41].

In contrast to the American position against indices, the Europeans established indices that were more efficient than the HLD Index or Grainger's TPI[30,31,36,42-45]. This was probably due to governmental pressures on orthodontists to identify patients who “need orthodontic treatment”, as these patients’ treatments were to be financially covered by the government[37,46]. This resulted in the development of many different indices[30,31,36,42-45], such as the Swedish Medical Board Index (SMBI)[36] and the Norwegian index of orthodontic treatment need[42]. In the United Kingdom (UK), five different orthodontic indices were established between 1989 and 2014 [30,31,43-45]. In 1989, Brook and Shaw proposed the Index of Orthodontic Treatment Need (IOTN), which was modified from the SMBI[30]. It was initially called the Index of Orthodontic Treatment Priority (IOTP) and consisted of two components: the Dental Health Component (DHC) and the Aesthetic Component (AC)[30]. Since 2006, the IOTN was routinely used in the National Health Service (NHS) Primary Care Clinics in the UK and in some secondary practice settings[47].

After the establishment of the IOTN, researchers in the UK worked to develop indices with a goal beyond treatment need assessment[31,43,44]. These efforts resulted in the development of the Peer Assessment Rating (PAR) Index in 1992 by Richmond et al[44]. The goal was not to measure the need for treatment, but rather to measure malocclusion at any stage[48]. This helped to assess both the pre-treatment status and the treatment outcome[48] .

In 2000, Daniels and Richmond developed the Index of Complexity, Outcome and Need (ICON) with 97 orthodontists from 9 different countries[31]. This index was easy, valid and reliable[49-52] . Llewellyn et al. criticized the high weighting for the aesthetic component of the ICON and proposed the Index of Treatment Complexity (IOTC)[45] . Recently, Ireland et al. developed the Index of Orthognathic Functional Treatment Need (IOFTN)[43]. This index prioritized cases with severe malocclusions that were not amenable to orthodontic treatment alone and required orthognathic surgery[43].

These European efforts coincided with a reduction in the American literature on orthodontic indices[37]. In 1998, the American Board of Orthodontics (ABO) established two systems: the American Board of Orthodontics – Discrepancy Index (ABO–DI) and the American Board of Orthodontics - Objective Grading System (ABO-OGS)[53-55]. These two systems were developed to objectively evaluate orthodontic cases presented as part of the ABO clinical examination process[53-55]. Since their introduction, many studies have used these two systems for research purposes[17,19,25-27,56-59].

In 2012, a review identified the most widely used indices in high-impact journals[60]. In descending order, the most frequently used indices were: IOTN, PAR Index, DAI and ICON[60]. The IOTN was established earlier than the PAR index or the ICON, which might have affected the frequency of its use[30]. In addition, the ABO-DI and the ABO-OGS were developed much later than the IOTN and the PAR index, which explained why neither the ABO-DI nor the ABO-OGS were listed in the most frequently used indices[60]. Most recently published studies have used the American Board of Orthodontics - Objective Grading System (ABO-OGS) to evaluate treatment outcomes and have come from the United Stated rather than Europe[17,19,20,25-27,59].

Recently, an eye-tracking device using the advanced technology measured the level of visual attention given by laypersons and was used to assess both treatment need and outcome from a lay perspective in a reliable and valid way[61,62].

1.2 Development of indices for pre-treatment assessment

1.2.1 <u>Peer Assessment Rating (PAR) Index</u>

The PAR index was introduced by Richmond et al. in 1992 in the United Kingdom[44]. Ten experienced orthodontists calibrated the index in 6 meetings[44]. It consisted of 11 components with two different weighting systems (UK and US weighting systems)[44,63] (See). It was also validated against the subjective opinions of a panel of 74 examiners[48]. The index was used to measure malocclusion at any stage and to assess orthodontic improvement[48]. However, there was no agreement on the use of the PAR index as a tool to assess treatment needs[64,65]. Cases with a score of zero indicated perfect alignment and occlusion, while high scores represented greater deviation from normal and associated malocclusion[44]. The index was validated for both digital and plaster models[66,67].

The PAR index has been criticized for the large weighting for overjet[68,69] . As shown in , a weighting of six was applied to overjet. Deguzman et al. proposed a US weighting system that involved decreasing overjet weight to 4.5 and increasing overbite weight to 3 [63]. The lower anterior alignment component was eliminated in the US weighting and the centerline was reduced to 3.5 from 4 [63].

McKnight et al. used UK and US weightings of the PAR index to measure the long term stability of 27 patients recalled on an average of 9 years after completion of a two-stage class II treatment[70]. A perfect occlusion at the end of treatment did not guarantee the absence of long-term relapse[70]. The amount of lower anterior relapse was not addressed on the US weighting system due to the elimination of the lower anterior component[70]. However, the difference between the two weighting systems (UK and US) was small[70].

Firestone et al. used UK and US weighting systems of the PAR index to assess orthodontic treatment needs[65]. Both weighting systems showed an excellent prediction for treatment need[65] with a high reliability[71,72] .

The PAR index was also used to evaluate orthodontists' perception of treatment need[73]. There was a high correlation between the PAR index score and orthodontists' perception of treatment need when the assessment was made without consideration of finances or patient desires[73].

1.2.2 Index of Complexity, Outcomes and Needs (ICON)

The ICON index was based on the average opinion of a large panel of international orthodontists in 9 countries (Germany, Greece, Hungary, Italy, Netherland, Norway, Spain, United Kingdom, United States)[31]. Five components were included: IOTN Aesthetic Component, crossbite, upper arch crowding/spacing, buccal segment antero-posterior relationships, and anterior vertical relationships (incisors overbite / open bite) (See , III and V)[31]. A high numerical score indicated a 'more urgent treatment need'[31]. There was 'no treatment needed' when total ICON score was less than 43, and 'treatment need' when ICON score was equal to or more than 43 points[31]. In addition to assessing treatment need, ICON evaluated case complexity[31]. The orthodontic treatment complexity was graded according to total ICON scores into easy (<29), mild (29-50), moderate (51-63), difficult (64-77) and very difficult (>77)[31] (See).

Firestone et al. found agreement between the ICON scores and a panel of US orthodontists who assessed treatment need[49]. The ICON was validated in other studies as a treatment need index in different countries[50,74-76]. However, some studies questioned the validity of ICON as a treatment need index and proposed a modified cut-off points guideline to improve its validity[77-79]. Moreover, the sensitivity of the ICON to detect treatment need was found to be high, but with a poor specificity[80]. However, the ICON index showed a good agreement with the IOTN and was considered to be a good possible substitute for the Dental Health Component (DHC) of the IOTN[81].

Many studies have shown the ICON to be a good measure of case complexity[49-52,76,82]. In addition, the ICON complexity category has a significant relationship with the severity of malocclusion as assessed by DAI[76]. Cases with higher pre-treatment ICON scores are more likely to be associated with longer treatment[83].

Onyeaso et al. evaluated four different indices (PAR , ABO-OGS, DAI and ICON) to determine whether 1 index could replace the other 3 to assess treatment needs and outcomes[84]. It was concluded that the ICON can be used in place of the DAI for assessing treatment need and in place of the PAR and the ABO-OGS for assessing treatment outcome[84].

One of the limitations of the ICON is the high weighting for the aesthetic component as it introduced a subjective judgment and reduced the required objectivity[45].

1.2.3 American Board of Orthodontics - Discrepancy Index (ABO-DI)

The ABO-DI was developed to measure case complexity, and considered both dental and skeletal findings[53]. The index included 12 categories: overjet, overbite, anterior open bite, lateral open bite, crowding, occlusal relationship, buccal crossbite, lingual crossbite, ANB angle , IMPA, Sn-GoGn, in addition to "others" category[53] (). The "others" category included: ankylosed , missing or supernumerary teeth, transposition, anomalies of tooth size and shape, skeletal asymmetry, midline discrepancy equal to or more than 3mm, diastema equal to or more than 2mm and treatment complexities (e.g. Cleft lip and palate) ()[53]. The ABO-DI was shown to be reliable if the evaluator was trained on the use of the ABO-DI[59,85,86].

Cases with higher DI scores took longer in treatment than those with a low DI scores[56,59,87]. A master thesis from Saint Louis University correlated the ABO-DI with treatment duration[87], cases with DI score of at least 15 points took longer than 22 months in treatment in 84.9% of the studied cases[87]. Vu et al. found that a 1-point increase in the ABO-DI was associated with an increased treatment time of 0.1month (3 days)[59]. Parrish et al. found a significant association between the ABO-DI and treatment duration, with 11 days increase in treatment duration for each increased point in total DI score[56]. A one-point increase in crowding, overjet, or overbite was associated with a 1 month increase in treatment duration[56]. In addition, a one-point increase in the occlusion category increased treatment duration by 3 weeks[56].

Cases with an ABO-DI of 7 or less finished 4.56 months sooner than cases with an ABO-DI of more than 17 points[88].

Given the ABO-DI scores for skeletal characteristics, consideration should be given to ethnic groups that have different skeletal norms than Caucasians because this might affect the DI score[89-92]. For example, on average, African Americans had a higher ANB angle than Caucasians[89,93]. Also, African Americans and Chinese individuals had higher lower incisor to mandibular plane angles than Caucasians (IMPA)[89,90,93]. Moreover, both Hispanic and Asian groups had steeper mandibular plane angles than Caucasians[90-92]. Therefore, the ABO-DI must be used with caution when applied to patients with mixed ethnicity.

When the ABO-DI was applied without scoring cephalometric values, it was found to be positively correlated with the PAR index for determining malocclusion severity[85].

1.3 Development of indices to measure post-treatment status (outcomes)

1.3.1 Peer Assessment Rating (PAR) Index

The PAR index was validated for measuring outcomes, in addition to assessing malocclusion severity[48]. After scoring pre- and post-treatment models, comparing scores revealed a percentage score reduction[48]. The degree of PAR improvement can be categorized into the following outcome categories[48] :

1. ‘‘Worse or no different’’ patients who showed less than a 30% reduction in the weighted PAR score.

2. ‘‘Improved’’ patients who showed greater than or equal to 30% reduction in the weighted PAR score.

3. ‘‘Greatly improved” patients who generally showed a reduction of 22 weighted PAR points, which is a greatly improved score compared to the initial malocclusion severity.

The PAR index has been criticized because it doesn’t evaluate precise tooth position and is insensitive to remaining extraction spaces[94,95]. Also, one of the disadvantages of the previous categorization is that cases with pre-treatment scores less than 22 points will never be classified as “greatly improved” even if the post-treatment outcome is zero. Recently, to overcome this limitation, Gu et al. proposed a new definition of “greatly improved” category[96]. They added cases that had an initial PAR score of less than 22 and scored zero points in the final PAR to the “greatly improved” category[96].

A further categorization of cases based on post-treatment PAR scores was suggested by Richmond et al.[48] and Tulloch et al.[3] . Cases with final PAR scores equal to or less than 5 were categorized as “almost ideal”, while cases with scores ranging from 6 to 10 were classified as “acceptable”[3,48]. Cases with final PAR scores of more than 10 were categorized as “less than acceptable”[3,48].

Dyken et al. evaluated the outcome of 54 cases which passed Phase III of the ABO certification process using the PAR index[97]. The PAR index scores of the Board cases were compared to 51 cases treated by orthodontic graduate students[97]. There was no statistically significant difference between the initial and final PAR scores of both groups[97]. It was concluded that the PAR index might not be the index of choice to compare excellent and good final occlusion[97].

Chalabi et al. compared treatment outcomes using US and UK weighted PAR indexes and the ABO-OGS on a sample of 50 patients. The authors found no statistically significant correlations between the US and UK weighted PAR index with the ABO-OGS[57].

Deguchi et al. used UK and US weightings of the PAR index and the ABO-OGS to compare the orthodontic outcomes of cases treated in postgraduate orthodontic clinics at Indiana University and Okayama University[85]. The authors didn’t find a significant correlation between the post-treatment PAR index and the ABO-OGS results[85].

To assess whether the initial malocclusion can predict the orthodontic treatment outcome, Sohrabi et al. performed a study on a sample of 102 subjects treated at the University of Washington Graduate Orthodontic Clinic[98]. The PAR index was applied to pre- and post-orthodontic records ,and the ABO-OGS was used on post-treatment records[98]. The authors found a significant positive relationship between initial and final PAR scores and ABO-OGS score, indicating that cases that started with more complex malocclusion were more difficult to be treated to higher standards[98].

1.3.3 Index of Complexity, Outcomes and Needs (ICON)

The ICON was developed to measure treatment outcomes as well as treatment complexity and needs[31]. To assess the degree of improvement, the post-treatment score was multiplied by 4 and the result was subtracted from the pretreatment score[31]. The improvement grade was ranked as greatly improved if the score was > –1 , substantially improved for –25 to –1 , moderately improved for –53 to –26 , minimally improved for –85 to –54 and not improved or worse if < –85[31] (See) . Savastano et al. concluded that the ICON is a valid index for assessing treatment outcome[82]. However, the degree of improvement component wasn't validated in their study due to low inter-rater reliability[82]. The intra-examiner agreement was slight for outcome and poor for degree of improvement, while inter-examiner agreement was moderate for outcome and fair for degree of improvement[82]. The ICON was found to be correlated with PAR index and can be used instead of the ABO-OGS to assess treatment outcomes[99].

1.3.4 American Board of Orthodontics - Objective Grading System (ABO-OGS)

The ABO-OGS was proposed in 1998 to evaluate dental models and panoramic radiographs of finished cases[54]. Unlike the PAR and the ICON, the ABO-OGS included an accurate and precise scoring for both occlusion and tooth position[54]. A total of eight criteria are examined to score each finished case (See)[54].A score of zero indicates excellent occlusion and teeth alignment. The highest possible score on the ABO-OGS is 280 points[55].

In 1998, the ABO published the guidelines for the use of the ABO-OGS and the “cut-off” scores for “pass/fail” categories. In order to pass the ABO Examination, a finished case has to score less than or equal to 20 points; cases scored above 30 points are considered to be “failed”[54]. Scores between 20 and 30 are subject to individual reassessment[54]. In 2012, the ABO published a revised version which stated that “scores for cases evaluated as “Complete/Passed” will vary from exam to exam, and may range from 27 or less”. This could mean that the “cut-off” for the “passed” category has changed from 20 to 27 points. In 2007, the term “Objective Grading System” was changed to “Cast and Radiographic Evaluation (CRE)”[55].

The ABO-CRE was studied in terms of its relationship to the ABO-DI[59,86,88,100,101]. Most of these studies were from Indiana University[59,86,100]. Both Campbell et al. and Pulfer et al. concluded that complex cases, based on the ABO-DI scores, are challenging to finish well[86,100]. Moreover, a third study from the same institution did not find any significance between the ABO-DI and the ABO-CRE and stated that “the initial DI scores were not significantly related to the final OGS score”[59]. Although patients with DI > 20 required more time to treat, a similar OGS outcome was achieved in comparison with less severe cases (DI <20)”[59].

A recent Master thesis concluded that cases with ABO-DI scores of 7 or less will probably score 4 points less on the ABO-CRE than cases with ABO-DI of more than 17 points[88]. The main areas that contributed to the 4-point difference were buccolingual inclination, occlusal contacts and occlusal relationships[88].

Asian populations have been found to have more difficulty in finishing by losing points on overjet and the buccolingual component of the ABO-CRE compared to Caucasian population[85]. This might be due to soft tissue factors or skeletal differences, as proposed by Deguchi et al.[85].

In 2017, Papageorgiou et al. conducted a systematic review and meta-analysis of published clinical studies to assess the level of orthodontic outcome and treatment duration[101]. Only studies that used full fixed appliances, assessed treatment duration and used the ABO-CRE for treatment outcomes were included[101]. Thirty-nine studies fulfilled the inclusion criteria, 27 of which were conducted at university settings. The average ABO-CRE score was 27.9 ± 2.6 [101]. The authors found no association between the ABO-CRE and treatment duration or ABO-DI scores, which is in agreement with some previous studies[102,103]. Based on the meta-analysis, cases that had teeth extraction were associated with a considerable reduction in the ABO-CRE score. This association was particularly strong for cases that required the extraction of 4 premolars, which had a mean reduction of 4.9 points, and a longer treatment time[101].

1.4 Aim of the study

To the best of our knowledge, there are no studies assessing the treatment outcomes of cases treated by a collaboration of two or more private practice orthodontists. The aim of the current study was to evaluate the quality of orthodontic treatment, treatment length and efficiency when two orthodontists collaborated on treatment and compare it to cases treated individually by either orthodontist.

1.4.1 Null hypothesis # 1

There is no difference in the orthodontic treatment outcome between cases treated by an individual orthodontist versus cases treated by two collaborating orthodontists.

1.4.2 Null hypothesis # 2

There is no difference in orthodontic treatment length between cases treated by an individual orthodontist versus cases treated by two collaborating orthodontists.

1.4.3 Null hypothesis # 3

There is no difference in orthodontic treatment efficiency between cases treated by an individual orthodontist versus cases treated by two collaborating orthodontists.

Chapter 2: Methodology

The protocol for this retrospective study was approved by the Ethical Board at the University of British Columbia, Vancouver, British Columbia, Canada (Certificate number: H16-01867).

2.1 Sample Selection

This study was conducted in three different settings:

1. Private orthodontic office in Richmond, British Columbia, Canada where all patients were only treated by orthodontist A (Group A).
2. Private orthodontic office in Delta, British Columbia, Canada where all patients were only treated by orthodontist B (Group B).
3. Private orthodontic office in Vancouver, British Columbia, Canada where both orthodontists (A and B) collaborated on treatment (Group C).

Both orthodontists are certified specialists in orthodontics, Diplomates of the American Board of Orthodontics, members of the Northwest Component of the Edward H. Angle Society and have a Fellowship in the Royal College of Dentists in Canada. Both clinicians also had at least ten years of experience as orthodontic specialists at the time of treatment of the study sample.

Consecutive sampling was used to review the treatment outcomes of all patients treated with full upper and lower edgewise appliance from January 2006 to December 2015. Exclusion criteria were patients with craniofacial syndromes or cleft lip and palate, phase I treatment, inter-disciplinary cases, partial orthodontic treatment cases, patients debonded prior to treatment completion, and patients with missing orthodontic records. A total of 50 subjects in each group were included. Dental charts were reviewed to confirm that all (100%) visits and progress notes were signed by the initials of orthodontist A for group A subjects and orthodontist B for group B subjects.

An arbitrary criterion for group C (joint cases) was that subjects needed to be seen by each orthodontist for at least 30% of the appointments, but not more than 70%. For example, if a patient was seen by orthodontist A for 25% of the appointments and by orthodontist B for 75% of the appointments, this patient was not eligible for group C, and was excluded from the study.

2.2 Data Collection

Clinical records of selected subjects were reviewed for data collection. This included a review of dental charts, pre-treatment orthodontic plaster models, intra-oral frontal photos, panoramic radiographs and lateral cephalometric radiographs (T1), post-treatment orthodontic plaster models, intra-oral frontal photos, panoramic radiographs and lateral cephalometric radiographs (T2).

Age, gender, treatment starting date (banding or bonding), date of debonding, number of visits, number of missed, cancelled and emergency appointments and type of malocclusion were also recorded. Treatment strategy (e.g. extraction or non-extraction) was recorded in addition to whether surgery was included. For group C, the number of appointments covered by orthodontist A/orthodontist B was recorded.

Study data were managed using REDCap electronic data capture tools hosted at The University of British Columbia[104]. The REDCap (Research Electronic Data Capture) is a secure, web-based application designed to support data capture for research studies that provides 1) an intuitive interface for validated data entry; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages, and; 4) procedures for importing data from external sources[104].

2.3 Pre-treatment scoring

All included cases were scored by an independent researcher (S.A.) using three different indices:

1. Peer Assessment Rating (PAR) Index

Pre-treatment orthodontic plaster models were scored using PAR index guidelines, including UK and US weightings. The PAR ruler was provided by Steven Richmond and was used for scoring (See). The scorer underwent and passed a calibration course on the PAR index in order to achieve a valid and reliable scoring.

2. Index of Complexity, outcomes and needs (ICON)

Pre-treatment orthodontic plaster models and intra-oral frontal photos were reviewed and scored using the ICON index guidelines (See , III and V).

3. American Board of Orthodontics - Discrepancy Index (ABO-DI)

Pre-treatment orthodontic plaster models, lateral cephalometric and panoramic radiographs were reviewed and scored using the ABO-DI worksheet (See Figure I). For calibration purposes, the scorer (S.A.) viewed the ABO website training video and underwent a training session by a previous ABO guest examiner (D.K.). The cases were divided into low, medium and high DI groups using ABO-DI scores as a reference. The low DI group included cases with a DI score of <10, the medium DI group consisted of cases with DI scores of ≥10 to <20, and the high DI group contained cases with DI scores of ≥20.

2.4 Post-treatment scoring

An independent researcher (S.A.) used three different indices to score all cases in order to measure orthodontic treatment outcomes. The indices were:

1. **Peer Assessment Rating (PAR) Index**

Post-treatment orthodontic plaster models were scored using PAR index guidelines for both UK and US weighting systems to evaluate treatment outcomes.

2. **Index of Complexity, Outcomes and Needs (ICON)**

Post-treatment orthodontic plaster models and intra-oral frontal photos were reviewed and scored using ICON index guidelines (See).

3. **American Board of Orthodontics - Cast Radiographic Evaluation (ABO-CRE)**

Post-treatment Orthodontic plaster models and panoramic radiographs were reviewed and scored using the ABO-CRE worksheet (See). Cases that scored equal to or less than 27 points were considered as "passed" cases, while cases with ABO-CRE scores of >27 to ≤30 were considered as "undetermined or borderline". Cases that scored more than 30 points were considered as "failed" cases. The examiner (S.A.) and a calibrated ABO previous guest examiner (D.K.) scored 40 cases independently and blindly as part of the calibration process for the ABO-CRE index. Three different statistical analyses were used to measure the level of agreement and variation between examiners. The first method assessed the variation between both examiners for every case by measuring the score range. For example, if one examiner scored 5 points and the other scored 7, then the scoring range was 2 points. A higher scoring range reflected more variation. This method was sensitive to outliers and showed the maximum and minimum scores. The second method assessed the agreement between both examiners using the Intraclass Correlation (ICC). The third method assessed the inter-examiner agreement employed the independent sample t-test.

2.5 Case Improvement and Treatment Efficiency

Pre- and post-treatment scores of both the PAR and ICON indices were calculated to assess the improvement level using the formula suggested by Richmond et al. and Gu et al. for the PAR Index[48,96] and by Daniels and Richmond for the ICON[31].

Treatment efficiency was measured using an equation defined as the percentage of the PAR index improvement divided by the treatment time in months, as suggested by Janson et al.[105].

2.6 Time Spent on Scoring Cases

The time spent to apply each index was recorded using an electronic timer. These data were collected for statistical analyses.

2.7 Measurement Error and Reliability

Twenty five randomly selected cases were scored twice, 3 weeks apart, using the PAR, ICON and ABO indices to measure the intra-examiner reliability.

2.8 Statistical Analysis

The data were entered using the REDCap Software and then transferred to the SPSS Software 22.0 (SPSS Inc., IL, USA) for statistical analysis.

Descriptive statistics were reported as means and standard deviations. An independent sample t-test was used during the ABO-CRE standardization process to compare the examiner's scores with the gold standard. One Way ANOVA compared means between the three groups. The chi square test and Fisher's exact test were used to compare the proportions between categorical variables. The significance threshold for all tests was set at 5%.

With a sample size of 150, the study had 92% power to detect a difference of 4 points on the ABO-CRE with a common standard deviation of 6.5 points; 96% power to detect a difference of 3.4 months in the length of treatment with a common standard deviation of 4.6 units, and; 97% power to detect a difference of 3 appointments in the number of visits with a common standard deviation of 4.5 units assuming a 2-sided type I error of 5%.

Chapter 3: Results

3.1 Inter-examiner reliability

The average variation between the examiner (S.A.) and the calibrated ABO previous guest examiner (D.K.) with regard to total ABO-CRE scores on 40 calibration cases was 1.6 points. The minimum range was zero, which indicated that absolute agreement was achieved in at least one case. The highest variation on the total ABO-CRE score was 9 points (See).

The second method to examine inter-examiner agreement was by ICC, which showed excellent agreement between examiners on the total ABO-CRE score (ICC = 0.88) as well as on all the parameters of the index, which ranged from 0.81 for buccolingual inclination to 0.98 for occlusal contacts (See). There was no statistically significant difference between both examiners on the total ABO-CRE score (P = 0.955) or on any of the ABO-CRE parameters (See).

3.2 Intra-examiner reliability

The ICC based on 25 cases showed an excellent intra-examiner reliability on all the occlusal indices used, and ranged from 0.86 for the ABO-CRE to 0.99 for the post-treatment ICON (See).

3.3 Sample description and characteristics

The distribution of the three groups with respect to the mean age at the start of treatment and gender is presented in . Age at start of treatment ranged from 10 years 2 months to 22 years 9 months with a mean age of 13.0 ± 1.8 years. Age of the oldest subject in group C was 16 years and 4 months compared to 22 years and 9 months for groups A and B. However, age difference was not statistically significant (P = 0.111) (See).

Seventy six of the patients in this study were females (50.7%) and 74 were male~~s~~ (49.3%). Groups A and B showed a similar gender distribution. Group C had more males (54%) than females (46%), but this difference was not statistically significant (*P* = 0.707) (See).

3.4 Distribution according to type of malocclusion, extraction treatment, surgical cases and year of bonding

The distribution of type of malocclusion, extraction treatment, surgical cases and year of bonding are shown in . Of the 150 cases, 58 patients (38.7%) had a Class I malocclusion, 59 patients (39.3%) had a Class II Division 1, 31 patients (20.7%) had a Class II Division 2 and only 2 patients (1.3%) had a Class III malocclusion. There was a small non-significant difference (*P* = 0.801) among the three groups (See).

Most patients (72.7%) had non-extraction orthodontic treatment, while 27.3% had extractions. All three groups showed a similar distribution of extraction and non-extraction cases (Chi square test, *P* = 0.791) (See). Two patients (1.3%) out of the 150 cases underwent jaw surgery, both of whom were from group A. There were no statistically significant proportional differences among the groups (Fisher Exact test, *P* = 0.132) (See).

In regards to the year of bonding, most of the cases were started in the period from January 2010 to December 2015 (58.0%) , with no statistically significant differences among the groups (Chi-square test, *P* = 0.563) (See).

3.5 Pre-treatment status

All pre-treatment scores are shown in .

3.5.1 PAR Index

The mean for the UK weighted PAR scores were: 22.1 ± 10.1 for group A, 24.2 ± 9.2 for group B and 23.3 ± 9.7 for group C, with no statistically significant differences among the groups (ANOVA, P = 0.559) (See).

Similar scores were shown using the US weighting. The mean US weighted PAR scores were: 23.3 ± 9.1 for group A, 24.2 ± 7.9 for group B and 23.7 ± 8.7 for group C, with no statistically significant differences among the groups (ANOVA, P = 0.871) (See).

3.5.2 ICON

The mean ICON scores were: 65.2 ± 18.6 for group A, 67.3 ± 16.5 for group B and 66.5 ± 16.3 for group C, with no statistically significant differences among the groups (ANOVA , P = 0.826) (See). Based on ICON difficulty categories, 30% of group B patients were classified as "very difficult" compared to 20% for group A and 28% for group C (See). However, 32% of group A were classified as "difficult" compared to 28% for group B and 26% for group C. Overall, all three groups had more than half of their patients classified as "difficult" or "very difficult" (52% for group A, 58% for group B and 54% for group C) , with no statistically significant differences among the groups (Chi-square test , P = 0.643) (See).

3.5.3 ABO-DI

The mean ABO-DI scores were: 16.3 ± 9.6 for group A, 16.3 ± 8.6 for group B and 18.5 ± 10.4 for group C, with no statistically significant differences among the groups (ANOVA, P = 0.424)

(See). The ABO-DI scores for most of the cases (46.7%) were on the range of ≥10 to <20 (See). Cases with scores of 20 or more were distributed almost evenly between the treatment groups. shows the distribution of cases based on the ABO-DI categories (<10, ≥10 to <20, ≥20), with no statistically significant differences among the groups (Chi-square test, $P = 0.430$).

3.6 Post-treatment status

All post-treatment scores are shown in .

3.6.1 PAR index

The mean UK weighted PAR scores were: 1.4 ± 1.1 for group A, 1.6 ± 1.4 for group B and 1.4 ± 1.6 for group C, with no statistically significant differences among the groups (ANOVA, $P = 0.805$) (See).

Higher points were scored using the US weighting. The mean US weighted PAR scores were: 2.7 ± 1.9 for group A, 2.8 ± 2.2 for group B and 2.7 ± 2.5 for group C, with no statistically significant differences among the groups (ANOVA, $P = 0.941$) (See).

Using the UK weighting, the mean overall post-treatment PAR score (1.5 ± 1.4) was significantly different from results obtained by using the US weighting (2.7 ± 2.7) (Independent sample t test, $P < 0.001$).

3.6.2 ICON

The mean post-treatment ICON scores were: 11.2 ± 2.9 for group A, 10.9 ± 3.3 for group B and 11.3 ± 4.0 for group C, with no statistically significant differences among the groups (ANOVA, $P = 0.825$) (See).

3.6.3 ABO-CRE

The mean ABO-CRE scores were: 21.5 ± 6.5 for group A, 22.0 ± 6.5 for group B and 25.3 ± 7.9 for group C (See). There was no statistically significant difference between groups A and B (See). The differences between groups A and C (*P* = 0.019) and groups B and C (*P* = 0.037) were statistically significant. Cases that had ABO-DI scores of 20 or more had a mean CRE score of 22.9 ± 7.4, with no statistically significant differences among the groups (ANOVA, *P* = 0.876) (See).

The majority of the sample patients (75.3%) were considered as "passed" cases (See). However, there was a significant difference in the distribution of passed, borderline and failed cases among the three groups, with less passed and more failed cases for group C compared to groups A and B (Chi-square test, *P* = 0.020) (See).

Groups A and B scored less points than group C on all parameters of the ABO-CRE, except for the "Alignment/Rotation" and "Occlusal Relationship", but these differences among the groups were not statistically significant (ANOVA, *P* = 0.906, *P* = 0.518, respectively) (See). The cases from groups A and B scored significantly less points on "Marginal Ridges" than the cases from group C (ANOVA with Post Hoc Bonferroni adjustment, *P* <0.001) (See). "Buccolingual Inclination" parameter scored significantly less points in group A than in group C (ANOVA with Post Hoc Bonferroni adjustment, *P* = 0.023) (See).

3.7 Treatment duration

On average, group A cases were treated in an average of 22.9 ± 4.6 months, while group B cases were treated in an average of 26.2 ± 6.5 months and group C was treated in average of 26.3 ± 7.0

months (See). Statistically significant differences were found between groups A and B (ANOVA with Post Hoc Bonferroni adjustment, *P* =0.027), and between groups A and C (ANOVA with Post Hoc Bonferroni adjustment, *P* =0.023) (See). There was no statistically significant difference in mean treatment time between groups B and C (*P* =0.998) (See).

Cases that had ABO-DI scores of 20 points or more had a mean treatment duration of 27.9 ± 5.7 months (See). These cases required significantly longer treatment time for group C (32.0 ± 7.5) than groups A and B (26.3 ± 5.1, 25.5 ± 5.7, respectively) (ANOVA with Post Hoc Bonferroni adjustment, *P* =0.032, *P* =0.014, respectively) (See).

3.8 Number of visits, cancellations, emergency visits and missed appointments

Total numbers of visits, cancellations, emergency visits and missed appointments were higher for group C than groups A and B. On average, group C cases required almost four more visits than group A cases and two more visits than group B cases. This difference was statistically significant only between groups A and C (ANOVA with Post Hoc Bonferroni adjustment, *P* =0.010) (See).

Cases that had ABO-DI scores of 20 or more required, on average, 27.7 ± 6.7 visits to be completed (See). These cases required significantly more visits in group C (32.0 ± 6.2), compared to groups A and B (26.2 ± 5.2, 24.8 ± 5.5, respectively) (ANOVA with Post Hoc Bonferroni adjustment, *P* =0.016, *P* =0.002, respectively) (See).

3.9 Case improvement and treatment efficiency

shows the improvement percentage for the studied sample using the UK and US weightings of the PAR index. All cases included in the study were classified as “improved” or “greatly improved”.

According to the original definition of "greatly improved" provided by Richmond et al.[48], 46% of the cases in groups A and B were classified as "greatly improved" compared to 38% for group C (Chi Square test, P = 0.653) (). Gu et al.[96] used a new definition of "greatly improved" category that included cases with a weighted PAR score reduction of 22 points or more or if the initial PAR score was less than 22 and a weighted PAR score after treatment equal to 0. Using this definition [96], 58% of group A cases, 60% of group B cases and 56% of group C cases were classified as "greatly improved". Similar results were shown using the US weighting (See).

Based on the ICON categories of case improvement, all cases were classified either as "greatly improved" or "substantially improved". Eighty six percent of group A cases, 88% of group B cases and 84% of group C cases were classified as greatly improved (See).

The Treatment Efficiency Index (TEI) showed no statistically significant differences among the three groups (See). However, the overall TEI results obtained by using UK weighting (3.9 ± 1.0) were significantly different from US weighting (3.7 ± 1.0) (Independent sample t-test, P = 0.039) (See).

3.10 Time spent on scoring the cases

shows the average time required for scoring the cases. For pre-treatment indices, on average, the ICON took significantly less time (ANOVA with Post Hoc Bonferroni adjustment, P <0.001) (1.5 minutes) to be applied than the ABO-DI or the PAR index (2.0 minutes). For the post treatment indices, the ABO-CRE took a significantly longer time (ANOVA with Post Hoc Bonferroni adjustment, P <0.001) (5.0 mins) to be applied that the ICON (0.7 minutes) or the PAR index scoring (0.8 minutes) (See).

Chapter 4: Discussion

This study was performed to assess whether cases treated by two collaborating orthodontists results in different treatment quality, length or efficiency compared to cases treated by a one orthodontist. To the best of our knowledge, this type of research has not been previously conducted.

Currently, corporate dental offices own 30-40 % of all dental offices in the US[106]. It is predicted that the total number of group practices will increase by 2020 and may come to constitute 50% of dental offices[107,108]. In regards to orthodontic specialty, reports found that the solo practice model in the US is declining among orthodontists, while professional partnerships are increasing[109]. Moreover, results from orthodontic resident surveys in 2014 showed that 56% of the residents planned to purchase a practice, while only 28% planned to work as associates[109]. However, orthodontic residents surveys in the 2015 survey showed the reverse: 54% planned to work as associates and 29% planned to purchase their own practice[109]. The findings of these orthodontic residents' surveys, in addition to shifting trends in orthodontic practices, highlight the importance of assessing whether multiple practitioners have similar, better, or worse treatment outcomes than sole practitioners. It is also assumed that corporate dental offices will aim for the most efficient way to provide treatments. The current study provided data on whether a collaboration of two orthodontists can affect treatment outcomes, length or efficiency.

It was essential to use a valid, and reliable method to classify the cases based on case complexity and treatment outcomes. Many orthodontic indices were proposed in the literature, and every index has its own advantages and disadvantages[30-34,36,37,40,43,44,48,53,55,60]. The PAR, ICON, ABO-DI and ABO-CRE were used in the present study. These indices were chosen as they quantify a variety of dental, skeletal and esthetic parameters[31,44,53,55]. In addition to measuring pre-treatment and post-treatment status, these indices allowed the classification of case difficulty and complexity (ICON,ABO-DI)[31,53], evaluated case improvement (PAR and ICON)[31,44], precisely examined treatment outcome (ABO-CRE)[55] and measured treatment efficiency (TEI-PAR)[105]. The present study provided a detailed evaluation of consecutively treated eligible cases.

Variables that can affect treatment difficulty, quality and/or duration were collected for each subject. These included age, gender, type of malocclusion, treatment modality, and the involvement of a surgical plan[87,110-114]. We also counted the total number of appointments and the number of missed, cancelled and emergency appointments, which may indirectly reflect patients' cooperation.

The PAR, ICON, ABO-DI and ABO-CRE showed excellent intra-examiner reliability. The examiner (S.A.) underwent a calibration course on the PAR index. The examiner (S.A.) was also trained on how to use the ABO-DI and was guided by the ABO website and the counsel of a previous guest ABO examiner (D.K.) in using the ABO-CRE prior to scoring the cases. Previous studies employing PAR[71,72,99], ICON[99,115] and ABO-DI[59,100] also indicated a similar level of excellent intra-examiner reliability. The ABO-CRE had the lowest intra-examiner reliability among all the indices used. Greco et al. reported a significant inter-examiner variations among the ABO board members on the use of ABO-CRE[116]. They reported a difference of 22 points in the total ABO-CRE score when two ABO board members scored the same case[116], compared to a maximum of 9 points difference in this study. The average score difference between the ABO board members on the total ABO-CRE scores for 40 cases was 6.1 points[116] , compared to 1.6 points in this study. Most of the variation in Greco et al. study was from the "Overjet" and

"Buccolingual Inclination" categories[116]. A year after publishing these results, the ABO made some modifications and explanations in the scoring of these two specific categories[55], possibly due to Greco et al findings[116]. The high inter-examiner reliability in this study during the standardization process can be attributed to the changes and clarifications made by the ABO on the application of the ABO-CRE and to the training sessions that the examiner underwent. It is important to mention that the scoring during the standardization process was done blindly by both examiners (S.A. and D.K.).

Many variables can affect treatment quality and duration[1-27]. A variety of variables were collected and analyzed to identify any differences in baseline status that might affect the study outcome. Age at the start of treatments was one variable we considered, although there is no consensus on whether age is an important determinant of treatment length or outcome[87,110,111]. The mean age of the three different groups in this study was not significantly different.

Gender is another possible confounding factor that can affect treatment length or outcome[87,112]. Studies have linked male patients to longer treatment times[112], possibly due to their expected poor cooperation during treatment which can affect treatment outcome[117]. In contrast, other studies did not find a difference between the overall treatment duration of males and females[87], while some found that female patients were more likely to miss appointments than males[118]. However, a recent study found that gender is not a significant predictor for adherence to orthodontic appointments[119]. In this study, gender distribution between groups A and B was similar, but both of these distributions were slightly different from group C. It is unlikely that differences in treatment length or outcome among the groups were due to gender differences, given the small difference and the conflicting findings regarding whether gender affects treatment length or outcome[101].

Almost 40% of our patients presented with a Cl II Division 1 type of malocclusion, followed by Cl I (38.7%), Cl II Division 2 (20.7%) and Cl III (1.3%). However, there were some differences among the study groups. Group A had equal numbers of Cl II Division 1 and Cl I cases, while group B had more Cl I than Cl II Division 1 cases and group C had more Cl II Division 1 than Cl I cases. We do not think that these differences affected the main outcomes of the study for a variety of reasons. First, these differences were not statistically significant. Second, when we combined the Cl II Division 1 and Cl II Division 2 cases into one category (Cl II malocclusion), all three groups had the same Cl II malocclusion distribution. Third, although the number of Cl II cases was higher in group C than group A or group B, the mean age of group C cases was less than the mean age of group A cases , also group C had more male patients than either group A or group B. These age and gender differences favored group C due to the potential for remaining growth which can improve treatment outcomes and shorten treatment duration[120].

In the current study, almost three quarters of the patients were non-extraction cases. There is evidence that extraction treatment takes longer than non-extraction treatment, with a difference of 7 months[113,114]. A study on the effect of extractions on treatment duration in five different private orthodontic offices found that the treatment duration was longer for extraction patients in each office[113]. However, when the data from the five offices were combined, the difference in treatment duration between extraction and non-extraction cases became almost similar and insignificant[113]. The authors concluded that different practitioners might contribute to different treatment length[113], which is why, in this study, we collected data from both orthodontists separately to address possible differences in treatment length or quality between orthodontist A and B.

The frequency of orthognathic surgery in this study was very small (1.2%). This is probably because the mean overall age for the patients in this study was 13.0 ± 1.8 years old.

Three occlusal indices (PAR index, ICON, ABO-DI) were used to measure pre-treatment status. When the PAR index was applied using the US and UK weighting systems, it showed a slightly higher non-significant score for group B, followed by group C and then group A. Similarly, the pre-treatment ICON scores were higher for group B, followed by group C and then group A. Unlike the PAR or ICON indices, ABO-DI scores were higher for group C than for group B. Group A scores were the lowest among the three study groups, as was shown by the PAR and ICON scores. These relatively small differences are not clinically important because the difference was one to two points at most. Although group C had the lowest number of Cl I patients, it had the highest ABO-DI scores. A future study should test the relationships among these indices and types of malocclusion, treatment duration and treatment outcomes.

Based on the ICON score, when “Very Difficult” and “Difficult” cases were combined into one category, a similar distribution of cases was noticed between the three groups. Based on the ABO-DI categories (low, medium and high), the distribution of cases was not significantly different. Also, the number of cases with a high ABO-DI (≥20 points) were almost evenly distributed among the three groups.

The similarity in baseline data (age, gender, type of malocclusion, extraction vs non-extraction orthognathic surgery, pre-treatment PAR, pre-treatment ICON and ABO-DI) allowed comparisons between cases treated by one versus two orthodontists with a minimum effect from possible confounding variables.

The main outcome of the study was the post-treatment status assessed by the PAR index, ICON and ABO-CRE. Different results were obtained using these three indices, but only the ABO-CRE mean scores showed statistically significant differences among the study groups. The PAR index and the ICON showed different results, but without statistically significant differences among the groups. Interestingly, the UK and US weighting of the post-treatment PAR index scores were significantly different, with higher scores for the US weighting which is similar to Dyken et al.'s findings[97], but different from the Deguchi et al.'s findings[85]. This could be attributed to the increased weight for "buccal occlusion" category on the US weighting because ABO-CRE, which is a more precise index, showed increased scores in "occlusal relationship" category in this study. However, both weighting systems showed a better outcome for group A, followed by group C and then group B. The overall orthodontic outcomes in this study were excellent based on the PAR scores, with a mean score of 1.5 on the UK weighting and 2.7 on the US weighting. Previous reports suggested that post-treatment PAR scores of 5 or less can be categorized as "almost ideal"[3,48]. The mean final PAR scores in this study fit in this category. The level of outcome (Post-PAR, Post-ICON, ABO-CRE) in the study was higher than in some previous studies[71,85,97,121].

Limited number of studies in the literature have assessed orthodontic outcomes in private orthodontic offices[111]. Most of the studies on orthodontic outcomes were performed on cases treated at graduate clinics[85,97,121]. Robb et al. used the PAR index on cases treated at a private orthodontic office, and the mean final PAR scores were 3.1 for adolescent and 3.7 for adult patients[111]. Although it was not clear whether the US or UK weighting was used, the final PAR scores were higher than our findings[111]. Dyken et al. assessed and compared the outcomes of 54 ABO-accepted cases treated by five different orthodontists to 51 cases treated at the University of Alabama School of Dentistry by graduate students[97] ; both the UK and US weightings were used in all the cases[97]. The mean UK-PAR scores for the Board cases were 3.1 compared to 4.0 for graduate students' cases[97]. The US weightings also showed a better outcome for the cases treated by orthodontists versus the cases treated by graduate students (5.4 versus 6.8 respectively)[97]. Our study, based on final PAR scores, reflected better outcomes than previous studies[71,85,97,121] , however some of those studies were conducted in educational setting[85,97,121].

Similar results of an excellent outcome were shown using the ICON index. The lowest possible ICON is 7 points, due to multiplying the score of the aesthetic component by 7[31]. Unlike PAR index, the ICON scores for group B cases, which had the highest scores on both weightings of the PAR index, showed better outcomes than the outcomes for group A or C. In this study, the mean final ICON score was 11.1 ± 3.4. Fox et al. evaluated 55 consecutively treated cases at the orthodontic clinics of Middlesbrough General Hospital and reported a mean final ICON score of 18.4 ± 7.9 [122].

Onyeaso et al. scored 100 post-treatment models of cases treated at a graduate orthodontic clinic at the University of Illinois at Chicago and obtained a final ICON score of 15.1 ± 2.9 [83]. There were fewer studies that used the ICON to measure treatment outcomes than the PAR index[60].

Although the ICON has the advantage of addressing esthetics when evaluating a case complexity or outcomes, it is not as precise as the ABO-CRE[31].

In order to precisely evaluate tooth position, we used the ABO-CRE. ABO-CRE results were different from the results of both the PAR index and the ICON. There was a significant difference between cases treated by one versus two orthodontists, without a significant difference between orthodontists A and B. Cases treated by both orthodontists did not result in as good outcome as cases treated by only one orthodontist. Concomitantly, the jointly treated cases still resulted in excellent quality based on the PAR and ICON indices.

A statistical analysis of the ABO-CRE parameters showed significant differences in the "marginal ridge" category between cases treated by one versus two orthodontists. Levelling marginal ridges needs precision during treatment, which might not be as likely if the case is treated by one solo orthodontist. The "buccolingual inclination" category was significantly better for group A cases compared to group B and C cases. Paying attention to premolar and molar torque is a routine for some orthodontists in order to avoid working and non-working interferences.

Most of the cases in this study were in the range of a passing grade based on the ABO-CRE scores ($\leq$27 points). However, cases treated jointly by the two orthodontists had significantly fewer passed cases and more failed cases. It is interesting that this significant difference was not observed when only complex cases (ABO-DI $\geq$ 20) were analyzed. This could mean that the main difference was in the low and medium ABO-DI groups.

Limited studies have used the ABO-CRE on cases treated in private practice[101]. A recent systematic review by Papageorgiou et al. found 34 studies that have used the ABO-CRE, and only 7 of these were conducted in private practice settings[101]. The total number of subjects in these studies were 6207 and all were included in a meta-analysis that presented an average ABO-CRE of 27.9 ± 2.6 [101].

There are different opinions on whether the cases treated in university setting are associated with different ABO-CRE scores than those treated in private practices[19,58,101,123]. Cook et al. assessed the outcome of 77 cases treated at three postgraduate orthodontic clinics and compared their ABO-CRE scores to 62 cases treated at three private orthodontic clinics[19]. They included only cases that had Cl II Division 1 malocclusion and required extractions of two or four premolars[19]. The mean ABO-CRE scores for cases treated at private practices (25.97 ± 9.70) was similar to the mean scores for cases treated at a university setting (25.14 ± 11.87)[19]. However, root angulation was not assessed[19]. Mislik et al. evaluated Cl I cases treated at a university setting and compared it to outcomes reported for private practice settings[58]. Sixty-six subjects were included, 32 of them were treated at a university setting and all variables of the ABO-CRE system were used[58]. There was no difference in the ABO-CRE scores for cases treated at private orthodontic clinics (25.94 ± 7.69) versus cases treated at a university setting (25.44 ± 9.81)[58]. In contrast, Yang-Powers et al. found a significant difference between ABO-CRE scores for cases treated by orthodontists (33.88 ± 9.69) versus cases treated by residents (45.54 ± 9.69)[123].

The second outcome of this study was treatment duration. On average, group A cases were treated in significantly less time than groups B and C cases, despite similar pre-treatment status. It was assumed that treatment duration would be similar between groups A and B since they were treated by individual orthodontists, but this was not the case. The reason for this could be the higher pre-treatment PAR and ICON scores for group B compared to group A. However, when cases with an ABO-DI of 20 points or more were analyzed separately, the treatment duration for groups A and B was found to be similar. The difference in treatment duration between groups A and C can be attributed to the fact that group C patients were treated by 2 orthodontists. Another explanation for this difference is that group C had a higher ABO-DI, which can increase treatment time[56]. Group C also had more Cl II Division 1 cases than either group A or B, which could have contributed to increased treatment time[124,125]. However, ABO-CRE scores for group C were significantly higher than the scores of groups A and B. If treatment of group C patients was continued until similar standards to groups A and B were achieved, the treatment time of group C patients might have been significantly longer than in group B.

Two recent systematic reviews assessed the treatment duration of orthodontic treatment with fixed appliances[101,126]. The first systematic review by Tsichlaki et al. included 22 studies, with an average treatment duration of 19.9 months[126]. In Papageorgiou et al.'s systematic review of 34 studies, the average treatment duration was 24.9 months[101]. This difference could be related to the fact that the Papageorgiou et al. reviewed only studies that used the ABO-CRE, which included board cases[101]. This might have required extra time to be spent during the final stages of treatment[101].

The number of appointments required is another important outcome from a practice management point of view. Group C “joint cases” required two more appointments than group B, and almost four more appointments than group A. Group B cases required significantly more treatment time than group A cases, but the number of appointments was increased by only one, which could indicate an increased inter-appointment interval in group B.

The number of missed and cancelled appointments was small. On average, patients in all groups missed almost two visits, cancelled two appointments, and scheduled two emergency visits. There was no significant difference among groups, although there was a trend of more cancellations, missed and emergency appointments in group C as compared to groups A and B.

It has been suggested that different pre-treatment statuses or post-treatment outcomes are associated with different treatment durations[56,100]. For example, difficult or complex cases need longer treatment time than easier cases[56]. In addition, cases with excellent outcomes are more likely to require increased treatment time than those with poorer outcomes[100]. These three variables (pre-treatment level, post-treatment outcome and length of treatment) were combined in one equation for every case to provide an objective measurement of treatment efficiency (as first used by Janson et al.)[105]. A higher score indicates a higher occlusal improvement and/or less treatment time. In this study, treatment efficiency was not significantly different among the three groups. An additional test of treatment efficiency was performed by dividing the percentage of occlusal improvement by the number of appointments instead of treatment duration, and similar results were found.

There is limited information on the time required to measure occlusal indices. Richmond et al. reported that six minutes were required to score pre- and post-treatment models using the PAR index[44], while the ICON index takes only a minute[31]. In a study of PAR, ICON, DAI and ABO-CRE, the ICON index was the quickest to use, with three minutes being required to measure each set of casts, while scoring with the PAR index took the longest time[99]. However, a strict recording of the time spent to score cases was not performed in that study[99]. Because no previous study reported the exact time required to score cases, we calculated the time spent to use the indices using an electronic timer. Our results showed significant difference between the indices. The ICON index took significantly less time to score both the pre-treatment (1.5 mins) and post-treatment (0.7 mins) statuses. The PAR and ICON indices took more time to score the pre-treatment status than the post-treatment outcome. This makes sense given the amount of initial malocclusion. An electronic data collection software (REDCap) was used in this study and all calculations were done electronically (e.g. multiplying scores by weights for the PAR or ICON indices), which may have contributed to a reduction in scoring time. For the ABO-DI, lateral cephalometric films were traced and measured before starting the scoring process. If the tracing was performed during the scoring process, it would have taken significantly more time. The ABO-CRE took five minutes to score, but more time might have been needed for an examiner who had just started using the index. The evaluator in this study was trained to use of ABO-CRE and applied the index to 40 cases before starting the data collection.

According to the recommendations of the recent meta-analysis and systematic review on orthodontic outcomes, researchers should evaluate baseline malocclusion when a comparison is to be made between different treatment groups[101]. The recommendations of the systematic review also state that both treatment duration and outcome should be considered when assessing treatment efficiency[101]. These three variables (baseline malocclusion, treatment duration and outcome) were taken into account and considered as the cornerstone of the present study. We aimed to take all possible variables into consideration in order to draw robust conclusions.

However, the present findings cannot be generalized to other group practices or providers of orthodontic treatments. Concomitantly, the methodology used in our study could be used to assess pre-treatment status and outcomes when treatments are performed by different providers (i.e. general dentists and pedodontists). Further, the outcomes of treatments with different appliances such as clear removable aligner appliances could also be assessed with a methodology similar to the one used in the current study.

Chapter 5: Conclusions

- Different outcome results were obtained using the PAR, ICON and ABO-CRE scoring systems.
- Based on the PAR and ICON results, there was no significant difference in treatment outcomes between orthodontic treatments performed by sole orthodontists versus two collaborating orthodontists.
- The final PAR scores were significantly higher when using the US weighting than when using the UK weighting.
- Sole orthodontists treated to the same ABO-CRE standards but with different treatment times.
- Based on the ABO-CRE, treatment outcomes were significantly better for cases treated by sole orthodontists than for cases treated by a collaboration of two orthodontists.
- Two collaborating orthodontists had fewer passed and more failed cases than solo orthodontists.
- The length of treatment time was not significantly different between cases treated by orthodontist B versus cases treated by both orthodontists.
- Cases treated by two collaborating orthodontists required two to four more visits than cases treated by one orthodontist.
- Treatment efficiency, as determined by the PAR scores, was not significantly different among the treatment groups.

Tables and Figures

Table I. Components of the PAR index [44] [63]

	Components	Unweighted	UK Weighting	US Weighting
1.	Upper right segment	×1	0	0
2.	Upper anterior segment	×1	×1	×1
3.	Upper left segment	×1	0	0
4.	Lower right segment	×1	0	0
5.	Lower anterior segment	×1	×1	0
6.	Lower left segment	×1	0	0
7.	Right buccal occlusion	×1	×1	×2
8.	Overjet	×1	×6	×4.5
9.	Overbite	×1	×2	×3
10.	Centerline	×1	×4	×3.5
11.	Left buccal occlusion	×1	×1	×2

Table II. Components of ICON [31]

Occlusal traits	ICON weighting
IOTN Aesthetic Component	7
Left + right buccal antero-posterior	3
Upper arch crowding	5
Overbite	4
Crossbite	5

Table III. ICON Complexity cut-off values [31]

Complexity grades	Score range
Easy	< 29
Mild	29–50
Moderate	51–63
Difficult	64–77
Very difficult	> 77

Table IV. ICON improvement grades [31]

Improvement grades	Score range
Greatly improved	> –1
Substantially improved	–25 to –1
Moderately improved	–53 to –26
Minimally improved	–85 to –54
Not improved or worse	< –85

Table V. Protocol for ICON scoring [31]

	Score	0	1	2	3	4	5
Aesthetic	1–10 As judged using IOTN AC						
Upper arch crowding	Score only the highest trait either spacing or crowding	Less than 2 mm	2·1 to 5 mm	5·1 to 9 mm	9·1 to 13 mm	13·1 to 17 mm	> 17 mm or impacted teeth
Upper spacing		Up to 2 mm	2·1–5 mm	5·1–9 mm	>9 mm		
Crossbite	Transverse relationship of cusp to cusp or worse	No crossbite	Crossbite present				
Incisor open bite	Score only the highest trait either open bite or overbite	Complete bite	Less than 1 mm	1·1–2 mm	2·1–4 mm	>4 mm	
Incisor overbite	Lower incisor coverage	Up to ⅓ tooth	⅓–⅔ coverage	⅔ up to full covered	Fully covered		
Buccal segment anteroposterior	Left and right added together	Cusp to embrasure relationship only, Class I, II or III	Any cusp relation up to but not including cusp to cusp	Cusp to cusp relationship			

Table VI: Intraclass correlations and ranges of the examiner scores with a calibrated ABO examiner on calibration cases (n = 40)

	ABO-CRE parameters							
	Alignment / rotations	Marginal ridges	Buccolingual inclination	Overjet	Occlusal contacts	Occlusal relationship	Interproximal contacts	Total
Mean scores of Examiner 1 scores	2.3 ± 2.0	2.6 ± 1.6	4.3 ± 2.9	2.3 ± 2.0	2.4 ± 2.2	6.1 ± 5.0	0.1 ± 0.5	20.6 ± 6.0
Mean scores of Examiner 2 scores	2.7 ± 2.0	2.1 ± 1.7	2.6 ± 2.4	2.7 ± 2.2	2.4 ± 2.2	5.5 ± 4.2	0.2 ± 0.6	19.0 ± 5.8
	P = 0.107	P = 0.965	P = 0.999	P = 0.129	P = 1	P = 0.814	P = 0.149	P = 0.955
Mean scores of the calibration cases	2.5 ± 2.0	2.4 ± 1.7	3.5 ± 2.7	2.5 ± 2.1	2.4 ± 2.2	5.8 ± 4.7	0.2 ± 0.6	19.8 ± 5.9
Shortest range	0	0	0	0	0	0	0	0
Largest range	4.0	4.0	5.0	4.0	2.0	4.0	1.0	9.0
Average range across the calibration cases	-0.4	0.5	1.7	-0.4	0	0.6	-0.1	1.6
Intraclass Correlation	0.90	0.82	0.81	0.91	0.98	0.97	0.84	0.88

For comparison b/w Ex.1 and Ex.2 scores , independent sample t-test was used

Table VII: Intra-examiner reliability on 25 randomly selected cases

	Used occlusal indices (n=25)							
	Pre-PAR UK	Pre-PAR US	Post-PAR UK	Post-PAR US	ABO-DI	ABO-CRE	Pre-ICON	Post-ICON
Intraclass Correlation	0.97	0.98	0.98	0.99	0.99	0.86	0.94	0.99

Table VIII: Time spent on scoring the sample using PAR, ICON and ABO indices

	Pre-treatment Indices			*Post-treatment Indices*		
	Pre-PAR	**Pre-ICON**	**ABO-DI**	**Post-PAR**	**Post-ICON**	**ABO-CRE**
Duration (mins)	2.0 ± 0.4	1.5 ± 0.3	2.0 ± 0.4	0.9 ± 0.2	0.7 ± 0.2	5.0 ± 1.0
	P <0.001			*P* <0.001		
	Pre-ICON vs Pre-PAR (*P*<0.001) Pre-ICON vs ABO-DI (*P*<0.001) Pre-PAR vs ABO-DI (*ns*)			Post-ICON vs Post-PAR (*P*<0.001) Post-ICON vs ABO-CRE (*P*<0.001) Post-PAR vs ABO-CRE (*P*=0.012)		

*** For comparison of means, ANOVA with Post Hoc Bonferroni adjustment was used**

Table IX: Sample description and characteristics

	Treatment groups				
	All *(n = 150)*	**Group A** *(n = 50)*	**Group B** *(n = 50)*	**Group C** *(n = 50)*	
Variable	*Mean ± SD*	*Mean ± SD*	*Mean ± SD*	*Mean ± SD*	*P value*
Age (years)	13.0 ± 1.8	13.4 ± 2.1	12.7 ± 1.8	12.9 ± 1.4	0.111†
Gender	*n (%)*	*n (%)*	*n (%)*	*n (%)*	0.707‡
Males	74 (49.3)	24 (48.0)	23 (46.0)	27 (54.0)	
Females	76 (50.7)	26 (52.0)	27 (54.0)	23 (46.0)	
Angle Classification					0.801‡
Class I	58 (38.7)	20 (40.0)	22 (44.0)	16 (32.0)	
Class II Div 1	59 (39.3)	20 (40.0)	17 (34.0)	22 (44.0)	
Class II Div 2	31 (20.7)	9 (18.0)	10 (20.0)	12 (24.0)	
Class III	2 (1.3)	1 (2.0)	1 (2.0)	0 (0.0)	
Extraction treatment					0.791‡
Yes	41 (27.3)	14 (28.0)	12 (24.0)	15 (30.0)	
No	109 (72.7)	36 (72.0)	38 (76.0)	35 (70.0)	
Surgical treatment					0.132⁋
Yes	2 (1.3)	2 (4.0)	0 (0.0)	0 (0.0)	
No	148 (98.7)	48 (96.0)	50 (100.0)	50 (100.0)	
Year of bonding					0.563⁋
2006-2010	63 (42.0)	19 (38.0)	20 (40.0)	24 (48.0)	
2010-2015	87 (58.0)	31 (62.0)	30 (60.0)	26 (52.0)	

† ANOVA with Post Hoc Bonferroni adjustment
‡ Chi square test
⁋ Fisher's exact test

Table X: Comparison between pre- and post- treatment scores and improvement grades

	Treatment groups				
	All	**Group A**	**Group B**	**Group C**	
	(n = 150)	*(n = 50)*	*(n = 50)*	*(n = 50)*	
Variable	*Mean ± SD*	*Mean ± SD*	*Mean ± SD*	*Mean ± SD*	*P value*
PAR Index (UK weighting)					
Pre-treatment	23.2 ± 9.6	22.1 ± 10.1	24.2 ± 9.2	23.3 ± 9.7	0.559†
Post-treatment	1.5 ± 1.4	1.4 ± 1.1	1.6 ± 1.4	1.4 ± 1.4	0.805†
Improvement rate	92.6 ± 7.4	91.9 ± 8.0	93.0 ± 6.7	93.7 ± 7.6	0.367†
Improved, n (%)	150 (100)	50 (100)	50 (100)	50 (100)	1.000‡
Greatly improved, n (%)	65 (43.3)	23 (46.0)	23 (46.0)	19 (38.0)	0.653‡
Greatly improved *by Gu et al 2017*, n (%)	87 (58.0)	29 (58.0)	30 (60.0)	28 (56.0)	0.923‡
PAR Index (US weighting)					
Pre-treatment	23.7 ± 8.6	23.3 ± 9.1	24.2 ± 7.9	23.7 ± 8.7	0.871†
Post-treatment	2.7 ± 2.7	2.7 ± 1.9	2.8 ± 2.2	2.7 ± 2.5	0.941†
Improvement rate	87.6 ± 11.2	86.2 ± 11.9	88.1 ± 9.9	88.3 ± 11.7	0.562†
Improved, n (%)	150 (100)	50 (100)	50 (100)	50 (100)	1.000‡
Greatly improved, n (%)	61 (40.7)	22 (44.0)	20 (40.0)	19 (38.0)	0.827‡
Greatly improved *by Gu et al 2017*, n (%)	85 (56.7)	29 (58.0)	29 (58.0)	27 (54.0)	0.899‡
ICON					
Pre-treatment	66.3 ± 17.0	65.2 ± 18.6	67.3 ± 16.5	66.5 ± 16.3	0.826†
Post-treatment	11.1 ± 3.4	11.2 ± 2.9	10.9 ± 3.3	11.3 ± 4.0	0.825†
Improvement grade	21.7 ± 20.8	20.4 ± 21.4	23.7 ± 20.3	21.2 ± 21.2	0.764†
Greatly improved, n (%)	129 (86.0)	43 (86.0)	44 (88.0)	42 (84.0)	0.847‡
Substantially improved, n (%)	21 (14.0)	7 (14.0)	6 (12.0)	8 (16.0)	0.847‡
ABO Indices					
ABO-DI	17.0 ± 9.6	16.3 ± 9.6	16.3 ± 8.6	18.5 ± 10.4	0.424†
ABO-CRE	22.9 ± 7.2	21.5 ± 6.5	22.0 ± 6.5	25.3 ± 7.9	0.014†*

Group A vs Group C (P = 0.019) | Group B vs Group C (P = 0.037) | Group A vs Group B (*ns*)

† ANOVA with Post Hoc Bonferroni adjustment
‡ Chi square test
* Significant difference between groups ($P < 0.05$)

Table XI: ICON Complexity grades among the sample

	Treatment groups				
	All *(n = 150)*	**Group A** *(n = 50)*	**Group B** *(n = 50)*	**Group C** *(n = 50)*	
Complexity grade	*n (%)*	*n (%)*	*n (%)*	*n (%)*	*P value*
Very Difficult	39 (26)	10 (20.0)	15 (30.0)	14 (28.0)	0.643‡
Difficult	43 (28.7)	16 (32.0)	14 (28.0)	13 (26.0)	
Moderate	43 (28.7)	14 (28.0)	12 (24.0)	17 (34.0)	
Mild	24 (16)	10 (20.0)	9 (18.0)	5 (10.0)	
Easy	1 (0.7)	0 (0)	0 (0)	1 (2.0)	

‡ Chi square test (for statistical analysis purposes , mild and easy grades were combined)

Table XII: ABO-DI Complexity grades among the sample

	Treatment groups				
	All *(n = 150)*	**Group A** *(n = 50)*	**Group B** *(n = 50)*	**Group C** *(n = 50)*	
Discrepancy Index	*n (%)*	*n (%)*	*n (%)*	*n (%)*	*P value*
Low (< 10)	31 (20.7)	14 (28.0)	10 (20.0)	7 (14.0)	0.430‡
Medium (≥ 10 to < 20)	70 (46.7)	19 (38.0)	24 (48.0)	27 (54.0)	
High (≥ 20)	49 (32.7)	17 (34.0)	16 (32.0)	16 (32.0)	

‡ Chi square test

Table XIII: ABO-CRE categories among the sample

ABO-CRE categories	All (*n = 150*) *n (%)*	Group A (*n = 50*) *n (%)*	Group B (*n = 50*) *n (%)*	Group C (*n = 50*) *n (%)*	*P value*
	Treatment groups				
Passed (≤ 27)	113 (75.3)	42 (84.0)	41 (82.0)	30 (60.0)	
Borderline (> 27 to ≤ 30)	13 (8.7)	2 (4.0)	2 (4.0)	9 (18.0)	0.020*‡
Failed (> 30)	24 (16.0)	6 (12.0)	7 (14.0)	11 (22.0)	

‡ Chi square test

Table XIV : ABO-CRE component scores between treatment groups

Variable	All (*n = 150*) *Mean ± SD*	Group A (*n = 50*) *Mean ± SD*	Group B (*n = 50*) *Mean ± SD*	Group C (*n = 50*) *Mean ± SD*	*P value*†
	Treatment groups				
Alignment / Rotations	2.9 ± 1.9	3.0 ± 1.8	2.9 ± 1.9	2.8 ± 1.9	0.906
Marginal ridges	2.7 ± 1.9	2.0 ± 1.7	2.3 ± 1.6	3.6 ± 2.2	<0.001*
		Group A vs Group C (*P*<0.001) | Group B vs Group C (*P* = 0.001) | Group A vs Group B (*ns*)			
Buccolingual Inclination	3.9 ± 2.7	3.5 ± 2.0	3.6 ± 2.9	4.8 ± 2.9	0.023*
		Group A vs Group C (*P* = 0.033) | Group B vs Group C (*ns*) | Group A vs Group B (*ns*)			
Overjet	3.3 ± 2.5	3.2 ± 2.5	3.1 ± 2.51	3.6 ± 2.6	0.520
Occlusal contacts	2.7 ± 2.6	2.6 ± 2.8	2.2 ± 2.2	3.3 ± 2.6	0.082
Occlusal relationship	6.7 ± 4.7	6.6 ± 4.2	7.3 ± 5.3	6.3 ± 4.6	0.518
Interproximal contacts	0.1 ± 0.4	0.1 ± 0.5	0.1 ± 0.3	0.1 ± 0.5	0.790
Root parallelism	0.6 ± 0.8	0.5 ± 0.7	0.5 ± 0.7	0.7 ± 0.8	0.233
Total ABO-CRE score	22.9 ± 7.2	21.5 ± 6.5	22.0 ± 6.5	25.3 ± 7.9	0.014*
		Group A vs Group C (*P* = 0.019) | Group B vs Group C (*P* = 0.037) | Group A vs Group B (*ns*)			

† ANOVA with Post Hoc Bonferroni adjustment
* Significant difference between groups ($P < 0.05$)

Table XV: Treatment duration and number of visits, missed, cancelled and emergency appointments

	Treatment groups				
	All *(n = 150)*	**Group A** *(n = 50)*	**Group B** *(n = 50)*	**Group C** *(n = 50)*	
Variable	*Mean ± SD*	*Mean ± SD*	*Mean ± SD*	*Mean ± SD*	*P value*†
Treatment duration (months)	25.1 ± 6.3	22.9 ± 4.6	26.2 ± 6.5	26.3 ± 7.0	0.011*
	Group A vs Group C (*P* = 0.023) | Group B vs Group C (*ns*) | Group A vs Group B (*P* = 0.027)				
No. of appointments	25.1 ± 5.8	23.4 ± 4.5	24.9 ± 5.8	27.0 ± 6.4	0.014*
	Group A vs Group C (*P* = 0.010) | Group B vs Group C (*ns*) | Group A vs Group B (*ns*)				
No. of missed appointments	1.8 ± 2.1	1.5 ± 2.2	1.7 ± 2.2	2.0 ± 1.9	0.077
No. of cancelled appointments	1.8 ± 1.6	1.4 ± 1.6	1.9 ± 1.7	2.0 ± 1.6	0.441
No. of emergency visits	1.7 ± 1.9	1.4 ± 1.9	1.7 ± 1.7	1.9 ± 2.0	0.425

† ANOVA with Post Hoc Bonferroni adjustment
* Significant difference between groups ($P < 0.05$)

Table XVI: Treatment duration, number of visits and ABO-CRE categories for cases with ABO-DI ≥ 20

	Treatment groups				
	All *(n = 49)*	**Group A** *(n = 17)*	**Group B** *(n = 16)*	**Group C** *(n = 16)*	
Variable	*Mean ± SD*	*Mean ± SD*	*Mean ± SD*	*Mean ± SD*	*P value*†
Treatment duration (months)	27.9 ± 5.7	26.3 ± 5.1	25.5 ± 5.7	32.0 ± 7.5	0.008*
		Group A vs Group C (*P* = 0.032) | Group B vs Group C (*P* = 0.014) | Group A vs Group B (*ns*)			
No. of appointments	27.7 ± 6.4	26.2 ± 5.2	24.8 ± 5.5	32.0 ± 6.2	0.002*
		Group A vs Group C (*P* = 0.016) | Group B vs Group C (*P* = 0.002) | Group A vs Group B (*ns*)			
ABO-CRE scores	22.9 ± 7.4	22.5 ± 8.4	22.6 ± 5.5	23.7 ± 8.2	0.876
ABO-CRE categories	*n (%)*	*n (%)*	*n (%)*	*n (%)*	
Passed (≤ 27)	38 (77.6)	12 (70.6)	14 (87.5)	12 (75.0)	0.584⁋
Borderline (> 27 to ≤ 30)	3 (6.1)	2 (5.9)	0 (0)	2 (12.5)	
Failed (> 30)	8 (16.3)	6 (23.5)	2 (12.5)	2 (12.5)	

† ANOVA with Post Hoc Bonferroni adjustment
* Significant difference between groups (*P* < 0.05)
⁋ Fisher's exact test

Table XVII: Treatment Efficiency Index (TEI) using UK and US PAR scores

	Treatment groups								
	All *(n = 150)*		**Group A** *(n = 50)*		**Group B** *(n = 50)*		**Group C** *(n = 50)*		
Variable	*Mean ± SD*		*Mean ± SD*		*Mean ± SD*		*Mean ± SD*		*P value*†
TEI (PAR-UK)	3.9 ± 1.0	*P*=0.039*	4.1 ± 0.8	*P*=0.095	3.8 ± 0.9	*P*=0.225	3.8 ± 1.1	*P*=0.379	0.113
TEI (PAR-US)	3.7 ± 1.0		3.9 ± 0.8		3.6 ± 0.9		3.6 ± 1.2		0.217

For comparison b/w UK and US weighting , independent sample t-test was used
†For comparison of means, ANOVA with Post Hoc Bonferroni adjustment was used

List of Figures:

TOTAL D.I. SCORE

For mm measures, round up to the next full mm.
Examiners will verify measurements in each category.

OVERJET

≥ 0 to < 1 mm (edge-to-edge)	= 1 pt
≥ 1 to ≤ 3 mm	= 0 pts
> 3 to ≤ 5 mm	= 2 pts
> 5 to ≤ 7 mm	= 3 pts
> 7 to ≤ 9 mm	= 4 pts
> 9 mm	= 5 pts
Negative Overjet (x-bite): 1 pt per mm per tooth	= ___pts
Total	

OVERBITE

> 1 to ≤ 3 mm	= 0 pts
> 3 to ≤ 5 mm	= 2 pts
> 5 to ≤ 7 mm	= 3 pts
Impinging (100%)	= 5 pts
Total	

ANTERIOR OPEN BITE

0 mm (edge-to-edge), 1 pt per tooth	=_____pts
then 1 pt per mm per tooth	=_____pts
Total	

LATERAL OPEN BITE

≥ 0.5 mm, 2 pts per mm per tooth

Total

CROWDING (only one arch)

≥ 0 to ≤1 mm	= 0 pts
> 1 to ≤ 3 mm	= 1 pts
> 3 to ≤ 5 mm	= 2 pts
> 5 to ≤ 7 mm	= 4 pts
> 7 mm	= 7 pts
Total	

OCCLUSAL RELATIONSHIP

Class I to End On	= 0 pts	
End-to-End Class II or III	= 2 pts per side	_____pts
Full Class II or III	= 4 pts per side	_____pts
Beyond Class II or III	= 1 pt per mm additional	_____pts
Total		

30130315

LINGUAL POSTERIOR X-BITE

> 0 mm, 1 pt per tooth Total

BUCCAL POSTERIOR X-BITE

> 0 mm, 2 pts per tooth Total

CEPHALOMETRICS (See Instructions)

ANB ≥ 6° or ≤ -2°	@4pts	= ___
Each full degree > 6°	___x 1 pt	= ___
Each full degree < -2°	___x 1 pt	= ___
SN-MP		
≥ 38°	@2pts	= ___
Each full degree > 38°	___x 2 pts	= ___
≤ 26°	@1pt	= ___
Each full degree < 26°	___x 1 pt	= ___
$\bar{1}$ to MP ≥ 99°	@1pt	= ___
Each full degree > 99°	___x 1 pt	= ___
	Total	

OTHER (See Instructions)

Supernumerary teeth	___x 1 pt	= ___
Ankylosis of permanent teeth	___x 2 pts	= ___
Anomalous morphology	___x 2 pts	= ___
Impaction (except 3rd molars)	___x 2 pts	= ___
Midline discrepancy (≥3 mm)	@ 2 pts	= ___
Missing teeth (except 3rd molars)	___x 1 pt	= ___
Missing teeth, congenital	___x 2 pts	= ___
Spacing (4 or more, per arch)	___x 2 pts	= ___
Spacing (mx cent diastema ≥ 2 mm)	@ 2 pts	= ___
Tooth transposition	___x 2 pts	= ___
Skeletal asymmetry(nonsurgical tx)	@ 3 pts	= ___
Addl. treatment complexities	___x 2 pts	= ___

Identify:

Total Other

Figure I ABO-Discrepancy Index Worksheet [127]

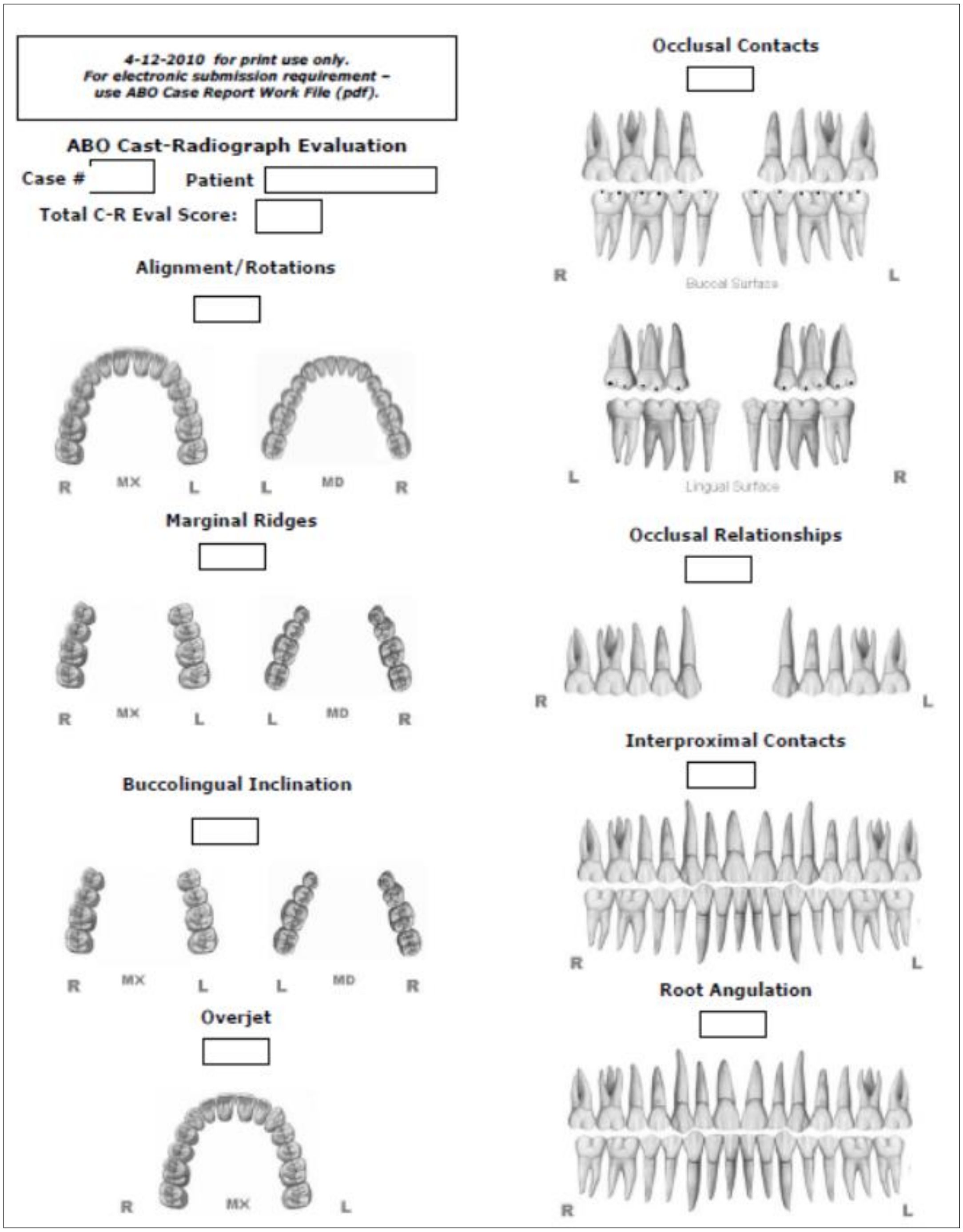

Figure II ABO-Cast and Radiographic Evaluation Worksheet [55]

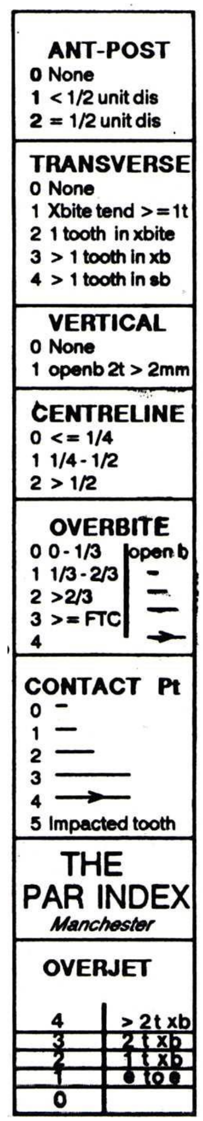

Figure III PAR Index ruler [44]

Bibliography

1. Hsieh T-J, Pinskaya Y, Roberts WE. Assessment of orthodontic treatment outcomes: early treatment versus late treatment. *Angle Orthod.* 2005;75(2):162-170.
2. Yang EY, Kiyak HA. Orthodontic treatment timing: a survey of orthodontists. *Am J Orthod Dentofacial Orthop.* 1998;113(1):96-103.
3. Tulloch JF, Phillips C, Proffit WR. Benefit of early Class II treatment: progress report of a two-phase randomized clinical trial. *Am J Orthod Dentofacial Orthop.* 1998;113(1):62-72, quiz 73-64.
4. Baccetti T, Franchi L, Giuntini V, Masucci C, Vangelisti A, Defraia E. Early vs late orthodontic treatment of deepbite: a prospective clinical trial in growing subjects. *Am J Orthod Dentofacial Orthop.* 2012;142(1):75-82.
5. Haruki T, Little RM. Early versus late treatment of crowded first premolar extraction cases: postretention evaluation of stability and relapse. *Angle Orthod.* 1998;68(1):61-68.
6. Von Bremen J, Pancherz H. Efficiency of early and late Class II Division 1 treatment. *Am J Orthod Dentofacial Orthop.* 2002;121(1):31-37.
7. Pandis N, Polychronopoulou A, Eliades T. Self-ligating vs conventional brackets in the treatment of mandibular crowding: a prospective clinical trial of treatment duration and dental effects. *Am J Orthod Dentofacial Orthop.* 2007;132(2):208-215.
8. Nabhan AF, Abbas NH, Fleming PS, Johal A, Sadek MM. Self-ligating brackets versus conventional pre-adjusted edgewise brackets for treating malocclusion. *The Cochrane Library.* 2016.
9. Miles PG. Self-ligating vs conventional twin brackets during en-masse space closure with sliding mechanics. *Am J Orthod Dentofacial Orthop.* 2007;132(2):223-225.
10. Read-Ward GE, Jones SP, Davies EH. A comparison of self-ligating and conventional orthodontic bracket systems. *Br J Orthod.* 1997;24(4):309-317.
11. Fleming PS, DiBiase AT, Lee RT. Randomized clinical trial of orthodontic treatment efficiency with self-ligating and conventional fixed orthodontic appliances. *Am J Orthod Dentofacial Orthop.* 2010;137(6):738-742.
12. Luppanapornlarp S, Johnston Jr LE. The effects of premolar-extraction: a long-term comparison of outcomes in “clear-cut” extraction and nonextraction Class II patients. *Angle Orthod.* 1993;63(4):257-272.
13. Paquette DE, Beattie JR, Johnston LE, Jr. A long-term comparison of nonextraction and premolar extraction edgewise therapy in "borderline" Class II patients. *Am J Orthod Dentofacial Orthop.* 1992;102(1):1-14.
14. Chua AL, Lim JY, Lubit EC. The effects of extraction versus nonextraction orthodontic treatment on the growth of the lower anterior face height. *Am J Orthod Dentofacial Orthop.* 1993;104(4):361-368.
15. Klapper L, Navarro SF, Bowman D, Pawlowski B. The influence of extraction and nonextraction orthodontic treatment on brachyfacial and dolichofacial growth patterns. *Am J Orthod Dentofacial Orthop.* 1992;101(5):425-430.

16. Verma SL, Sharma V, Tandon P, Singh GP, Sachan K. Comparison of esthetic outcome after extraction or non-extraction orthodontic treatment in class II division 1 malocclusion patients. *Contemp Clin Dent.* 2013;4(2):206.
17. Anthopoulou C, Konstantonis D, Makou M. Treatment outcomes after extraction and nonextraction treatment evaluated with the American Board of Orthodontics objective grading system. *Am J Orthod Dentofacial Orthop.* 2014;146(6):717-723.
18. Mascarenhas AK, Vig K. Comparison of orthodontic treatment outcomes in educational and private practice settings. *Journal of Dental Education.* 2002;66(1):94-99.
19. Cook DR, Harris EF, Vaden JL. Comparison of university and private-practice orthodontic treatment outcomes with the American Board of Orthodontics objective grading system. *Am J Orthod Dentofacial Orthop.* 2005;127(6):707-712.
20. Mislik B, Konstantonis D, Katsadouris A, Eliades T. University clinic and private practice treatment outcomes in Class I extraction and nonextraction patients: A comparative study with the American Board of Orthodontics Objective Grading System. *Am J Orthod Dentofacial Orthop.* 2016;149(2):253-258.
21. Proffit WR, Phillips C, Tulloch JF, Medland PH. Surgical versus orthodontic correction of skeletal Class II malocclusion in adolescents: effects and indications. *The International journal of adult orthodontics and orthognathic surgery.* 1992;7(4):209-220.
22. Tucker MR. Orthognathic surgery versus orthodontic camouflage in the treatment of mandibular deficiency. *J Oral Maxillofac Surg.* 1995;53(5):572-578.
23. Tulloch JF, Lenz BE, Phillips C. Surgical versus orthodontic correction for Class II patients: age and severity in treatment planning and treatment outcome. *Semin Orthod.* 1999;5(4):231-240.
24. Mihalik CA, Proffit WR, Phillips C. Long-term follow-up of Class II adults treated with orthodontic camouflage: a comparison with orthognathic surgery outcomes. *Am J Orthod Dentofacial Orthop.* 2003;123(3):266-278.
25. Jain M, Varghese J, Mascarenhas R, Mogra S, Shetty S, Dhakar N. Assessment of clinical outcomes of Roth and MBT bracket prescription using the American Board of Orthodontics Objective Grading System. *Contemp Clin Dent.* 2013;4(3):307-312.
26. Detterline DA, Isikbay SC, Brizendine EJ, Kula KS. Clinical outcomes of 0.018-inch and 0.022-inch bracket slot using the ABO objective grading system. *Angle Orthod.* 2010;80(3):528-532.
27. Kuncio D, Maganzini A, Shelton C, Freeman K. Invisalign and traditional orthodontic treatment postretention outcomes compared using the American Board of Orthodontics objective grading system. *Angle Orthod.* 2007;77(5):864-869.
28. McGuinness NJ, McDonald JP. The influence of operator changes on orthodontic treatment times and results in a postgraduate teaching environment. *Eur J Orthod.* 1998;20(2):159-167.
29. Peppers EC, Leggitt VL, Caruso JM, Neufeld R, Green J. The effect of changes in primary attending doctor coverage frequency on orthodontic treatment time and results. *Angle Orthod.* 2015;85(6):1051-1056.
30. Brook PH, Shaw WC. The development of an index of orthodontic treatment priority. *Eur J Orthod.* 1989;11(3):309-320.
31. Daniels C, Richmond S. The development of the index of complexity, outcome and need (ICON). *J Orthod.* 2000;27(2):149.

32. Draker HL. Handicapping labio-lingual deviations: a proposed index for public health purposes. *American Journal of Orthodontics.* 1960;46(4):295-305.
33. Grainger RM. Orthodontic treatment priority index. *Vital Health Stat 2.* 1967(25):1-49.
34. Salzmann J. Handicapping malocclusion assessment to establish treatment priority. *American Journal of Orthodontics.* 1968;54(10):749-765.
35. Summers CJ. The occlusal index: a system for identifying and scoring occlusal disorders. *American journal of orthodontics.* 1971;59(6):552-567.
36. Ingervall B, Ronnerman A. Index for need of orthodontic treatment. *Odontol Revy.* 1975;26(1):59-82.
37. Parker WS. The HLD (CalMod) index and the index question. *Am J Orthod Dentofacial Orthop.* 1998;114(2):134-141.
38. *AAO Bulletin: American Association of Orthodontists, St Louis* (Fall 1990).
39. Han H, Davidson WM. A useful insight into 2 occlusal indexes: HLD (Md) and HLD (CalMod). *Am J Orthod Dentofacial Orthop.* 2001;120(3):247-253.
40. Cons NC, Jenny J, Kohout FJ. *Dental Aesthetic Index.* College of Dentistry, University of Iowa; 1986.
41. World Health O. *Oral data collection instrument and examination criteria : International Collaborative Study of Oral Health Outcomes (ICS II).* Geneva: World Health Organization; 1989.
42. Espeland LV, Ivarsson K, Stenvik A. A new Norwegian index of orthodontic treatment need related to orthodontic concern among 11-year-olds and their parents. *Community dentistry and oral epidemiology.* 1992;20(5):274-279.
43. Ireland AJ, Cunningham SJ, Petrie A, et al. An index of orthognathic functional treatment need (IOFTN). *J Orthod.* 2014;41(2):77-83.
44. Richmond S, Shaw W, O'brien K, et al. The development of the PAR Index (Peer Assessment Rating): reliability and validity. *Eur J Orthod.* 1992;14(2):125-139.
45. Llewellyn SK, Hamdan AM, Rock WP. An index of orthodontic treatment complexity. *Eur J Orthod.* 2007;29(2):186-192.
46. Borzabadi-Farahani A. *An overview of selected orthodontic treatment need indices.* In Tech; 2011.
47. Holmes A, Willmot DR. The Consultant Orthodontists Group 1994 survey of the use of the Index of Orthodontic Treatment Need (IOTN). *Br J Orthod.* 1996;23(1):57-59.
48. Richmond S, Shaw W, Roberts C, Andrews M. The PAR Index (Peer Assessment Rating): methods to determine outcome of orthodontic treatment in terms of improvement and standards. *Eur J Orthod.* 1992;14(3):180-187.
49. Firestone AR, Beck FM, Beglin FM, Vig KW. Validity of the Index of Complexity, Outcome, and Need (ICON) in determining orthodontic treatment need. *Angle Orthod.* 2002;72(1):15-20.
50. Liepa A, Urtane I, Richmond S, Dunstan F. Orthodontic treatment need in Latvia. *Eur J Orthod.* 2003;25(3):279-284.
51. Richmond S, Ikonomou C, Williams B, Ramel S, Rolfe B, Kurol J. Orthodontic treatment standards in a public group practice in Sweden. *Swed Dent J.* 2001;25(4):137-144.
52. Richmond S, Ikonomou C, Williams B, Rolfe B. Orthodontic treatment standards in Greece. *Hellenic Orthodontic Review.* 2001;4(1).

53. Cangialosi TJ, Riolo ML, Owens SE, et al. The ABO discrepancy index: a measure of case complexity. *Am J Orthod Dentofacial Orthop.* 2004;125(3):270-278.
54. Casko JS, Vaden JL, Kokich VG, et al. Objective grading system for dental casts and panoramic radiographs. American Board of Orthodontics. *Am J Orthod Dentofacial Orthop.* 1998;114(5):589-599.
55. The American Board of Orthodontics : Grading System for Dental Casts and Panoramic Radiographs. 2012; https://www.americanboardortho.com/media/1191/grading-system-casts-radiographs.pdf. Accessed Nov 11, 2016.
56. Parrish LD, Roberts WE, Maupome G, Stewart KT, Bandy RW, Kula KS. The relationship between the ABO discrepancy index and treatment duration in a graduate orthodontic clinic. *Angle Orthod.* 2011;81(2):192-197.
57. Chalabi O, Preston CB, Al-Jewair TS, Tabbaa S. A comparison of orthodontic treatment outcomes using the Objective Grading System (OGS) and the Peer Assessment Rating (PAR) index. *Aust Orthod J.* 2015;31(2):157-164.
58. Mislik B, Konstantonis D, Katsadouris A, Eliades T. University clinic and private practice treatment outcomes in Class I extraction and nonextraction patients: A comparative study with the American Board of Orthodontics Objective Grading System. *Am J Orthod Dentofacial Orthop.* 2016;149(2):253-258.
59. Vu CQ, Roberts WE, Hartsfield JK, Jr., Ofner S. Treatment complexity index for assessing the relationship of treatment duration and outcomes in a graduate orthodontics clinic. *Am J Orthod Dentofacial Orthop.* 2008;133(1):9 e1-13.
60. Bellot-Arcís C, Almerich-Silla J, Paredes-Gallardo V, Gandía-Franco J. The use of occlusal indices in high-impact literature. *Community dental health.* 2012;29(1):45-48.
61. Wang X, Cai B, Cao Y, et al. Objective method for evaluating orthodontic treatment from the lay perspective: An eye-tracking study. *Am J Orthod Dentofacial Orthop.* 2016;150(4):601-610.
62. Richards MR, Fields HW, Jr., Beck FM, et al. Contribution of malocclusion and female facial attractiveness to smile esthetics evaluated by eye tracking. *Am J Orthod Dentofacial Orthop.* 2015;147(4):472-482.
63. DeGuzman L, Bahiraei D, Vig KW, Vig PS, Weyant RJ, O'Brien K. The validation of the Peer Assessment Rating index for malocclusion severity and treatment difficulty. *Am J Orthod Dentofacial Orthop.* 1995;107(2):172-176.
64. Bergtröm K, Halling A. Comparison of three indices in evaluation of orthodontic treatment outcome. *Acta Odontologica Scandinavica.* 1997;55(1):36-43.
65. Firestone AR, Beck FM, Beglin FM, Vig KW. Evaluation of the peer assessment rating (PAR) index as an index of orthodontic treatment need. *Am J Orthod Dentofacial Orthop.* 2002;122(5):463-469.
66. Andrews CK-s. *Validity and Reliability of Peer Assessment Rating Index Scores of Digital and Plaster Models*, The Ohio State University; 2008.
67. Mayers M, Firestone AR, Rashid R, Vig KW. Comparison of peer assessment rating (PAR) index scores of plaster and computer-based digital models. *Am J Orthod Dentofacial Orthop.* 2005;128(4):431-434.
68. Turbill EA, Richmond S, Wright JL. A critical assessment of orthodontic standards in England and Wales (1990-1991) in relation to changes in prior approval. *Br J Orthod.* 1996;23(3):221-228.

69. Hamdan AM, Rock WP. An appraisal of the Peer Assessment Rating (PAR) Index and a suggested new weighting system. *Eur J Orthod.* 1999;21(2):181-192.
70. McKnight MM, Daniels CP, Johnston LE, Jr. A retrospective study of two-stage treatment outcomes assessed with two modified PAR indices. *Angle Orthod.* 1998;68(6):521-524; discussion 525-526.
71. Birkeland K, Furevik J, Bøe OE, Wisth PJ. Evaluation of treatment and post-treatment changes by the PAR Index. *Eur J Orthod.* 1997;19(3):279-288.
72. Pangrazio-Kulbersh V, Kaczynski R, Shunock M. Early treatment outcome assessed by the Peer Assessment Rating index. *Am J Orthod Dentofacial Orthop.* 1999;115(5):544-550.
73. McGorray SP, Wheeler TT, Keeling SD, Yurkiewicz L, Taylor MG, King GJ. Evaluation of orthodontists' perception of treatment need and the peer assessment rating (PAR) index. *Angle Orthod.* 1999;69(4):325-333.
74. Abdelkader HM, Mohammed RE, Ali MM. The reliability and validity of the Index of Complexity, Outcome and Need (ICON) among sample of egyptian orthodontic patients. *International Journal of Clinical Dentistry.* 2015;8(2).
75. Costa RN, Abreu MH, Magalhaes CS, Moreira AN. Validity of two occlusal indices for determining orthodontic treatment needs of patients treated in a public university in Belo Horizonte, Minas Gerais State, Brazil. *Cad Saude Publica.* 2011;27(3):581-590.
76. Onyeaso CO. Relationship between Index of Complexity, Outcome and Need and Dental Aesthetic Index in the assessment of orthodontic treatment complexity and need of Nigerian adolescents. *Brazilian Research in Pediatric Dentistry and Integrated Clinic.* 2008;8(2):141-145.
77. Liao ZY, Jian F, Long H, et al. Validity assessment and determination of the cutoff value for the Index of Complexity, Outcome and Need among 12-13 year-olds in Southern Chinese. *Int J Oral Sci.* 2012;4(2):88-93.
78. Louwerse TJ, Aartman IH, Kramer GJ, Prahl-Andersen B. The reliability and validity of the Index of Complexity, Outcome and Need for determining treatment need in Dutch orthodontic practice. *Eur J Orthod.* 2006;28(1):58-64.
79. Torkan S, Pakshir HR, Fattahi HR, et al. An Analytical Study on an Orthodontic Index: Index of Complexity, Outcome and Need (ICON). *J Dent (Shiraz).* 2015;16(3):149-155.
80. Koochek A, Shue-Te Yeh M, Rolfe B, Richmond S. General Practice: The relationship between Index of Complexity, Outcome and Need, and patients' perceptions of malocclusion: a study in general dental practice. *British Dental Journal.* 2001;191(6):325-329.
81. Borzabadi-Farahani A, Borzabadi-Farahani A. Agreement between the index of complexity, outcome, and need and the dental and aesthetic components of the index of orthodontic treatment need. *Am J Orthod Dentofacial Orthop.* 2011;140(2):233-238.
82. Savastano NJ, Jr., Firestone AR, Beck FM, Vig KW. Validation of the complexity and treatment outcome components of the index of complexity, outcome, and need (ICON). *Am J Orthod Dentofacial Orthop.* 2003;124(3):244-248.
83. Onyeaso CO, BeGole EA. Orthodontic treatment standard in an accredited graduate orthodontic clinic in North America assessed using the Index of Complexity, Outcome and Need (ICON). *Hellenic Orthodontic Review.* 2006;9(1).

84. Templeton KM, Powell R, Moore MB, Williams AC, Sandy JR. Are the Peer Assessment Rating Index and the Index of Treatment Complexity, Outcome, and Need suitable measures for orthognathic outcomes? *Eur J Orthod.* 2006;28(5):462-466.
85. Deguchi T, Honjo T, Fukunaga T, Miyawaki S, Roberts WE, Takano-Yamamoto T. Clinical assessment of orthodontic outcomes with the peer assessment rating, discrepancy index, objective grading system, and comprehensive clinical assessment. *Am J Orthod Dentofacial Orthop.* 2005;127(4):434-443.
86. Pulfer RM, Drake CT, Maupome G, Eckert GJ, Roberts WE. The association of malocclusion complexity and orthodontic treatment outcomes. *Angle Orthod.* 2009;79(3):468-472.
87. Simister BG. *A Study of Treatment Duration Using the Discrepancy Index*, Saint Louis University; 2007.
88. Cameron DL. *The relationship between American Board of Orthodontics pretreatment dental cast Discrepancy Index scores and posttreatment dental cast Objective Grading System scores*, Saint Louis University; 2010.
89. Alexander TL, Hitchcock HP. Cephalometric standards for American Negro children. *American Journal of Orthodontics.* 1978;74(3):298-304.
90. Chung C, Kau M, Walker G. Racial variation of cephalometric measurements in Hawaii. *Journal of Craniofacial Genetics and Developmental Biology.* 1981;2(2):99-106.
91. Lau J, Hägg U. Cephalometric morphology of Chinese with Class II division 1 malocclusion. *British Dental Journal.* 1999;186(4):188-190.
92. Evanko AM, Freeman K, Cisneros GJ. Mesh diagram analysis: developing a norm for Puerto Rican Americans. *Angle Orthod.* 1997;67(5):381-388.
93. Bailey KL, Taylor RW. Mesh diagram cephalometric norms for Americans of African descent. *Am J Orthod Dentofacial Orthop.* 1998;114(2):218-223.
94. Birkeland K, Boe OE, Wisth PJ. Subjective assessment of dental and psychosocial effects of orthodontic treatment. *J Orofac Orthop.* 1997;58(1):44-61.
95. Hinman C. The Dental Practice Board Orthodontics—The current status. *British Journal of Orthodontics.* 1995;22(3):287-290.
96. Gu J, Tang JS, Skulski B, et al. Evaluation of Invisalign treatment effectiveness and efficiency compared with conventional fixed appliances using the Peer Assessment Rating index. *Am J Orthod Dentofacial Orthop.* 2017;151(2):259-266.
97. Dyken RA, Sadowsky PL, Hurst D. Orthodontic outcomes assessment using the peer assessment rating index. *Angle Orthod.* 2001;71(3):164-169.
98. Sohrabi K. *Does initial malocclusion predict the outcome of orthodontic treatment? ,* University of Washington; 2016.
99. Onyeaso CO, Begole EA. Relationship between index of complexity, outcome and need, dental aesthetic index, peer assessment rating index, and American Board of Orthodontics objective grading system. *Am J Orthod Dentofacial Orthop.* 2007;131(2):248-252.
100. Campbell CL, Roberts WE, Hartsfield JK, Jr., Qi R. Treatment outcomes in a graduate orthodontic clinic for cases defined by the American Board of Orthodontics malocclusion categories. *Am J Orthod Dentofacial Orthop.* 2007;132(6):822-829.

101. Papageorgiou SN, Hochli D, Eliades T. Outcomes of comprehensive fixed appliance orthodontic treatment: A systematic review with meta-analysis and methodological overview. *Korean J Orthod.* 2017;47(6):401-413.
102. 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(3):292-298; discussion 298.
103. Viwattanatipa N, Buapuean W, Komoltri C. Relationship between discrepancy index and the objective grading system in thai board of orthodontics patients. *Orthodontic Waves.* 2016;75(3):54-63.
104. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42(2):377-381.
105. Janson G, Baldo T, Garib D, Barros SE, Silva Poletto R, Bittencourt Dos Santos P. Efficiency of Class II subdivision malocclusion treatment with 3 and 4 premolar extractions. *Am J Orthod Dentofacial Orthop.* 2016;150(3):499-503.
106. Association CD. 2017; https://www.cda-adc.ca/en/services/internationallytrained/economic/. Accessed Dec 20th 2017, 2017.
107. Center AS. *2010 American Dental Association Dental Workforce Model 2008-2030.* American Dental Association; 2010.
108. Ag IS. How will dentistry look in 2020? 2012; http://www.straumann.com/content/dam/internet/straumann_com/Resources/investor-relations/publications-and-reports/capital-markets-day-2012/How%20will%20dentistry%20in%202020%20look_Straumann%20CMD2012_Achermann.pdf. Accessed Dec 20th 2017, 2017.
109. Orthodontists AAo. *You Have Options: Orthodontic Career Considerations for Residents and Graduates.* 2016.
110. Weiss J, Eiser HM. Psychological timing of orthodontic treatment. *American Journal of Orthodontics.* 1977;72(2):198-204.
111. Robb SI, Sadowsky C, Schneider BJ, BeGole EA. Effectiveness and duration of orthodontic treatment in adults and adolescents. *Am J Orthod Dentofacial Orthop.* 1998;114(4):383-386.
112. Skidmore KJ, Brook KJ, Thomson WM, Harding WJ. Factors influencing treatment time in orthodontic patients. *Am J Orthod Dentofacial Orthop.* 2006;129(2):230-238.
113. Vig PS, Weintraub JA, Brown C, Kowalski CJ. The duration of orthodontic treatment with and without extractions: a pilot study of five selected practices. *Am J Orthod Dentofacial Orthop.* 1990;97(1):45-51.
114. Fink DF, Smith RJ. The duration of orthodontic treatment. *Am J Orthod Dentofacial Orthop.* 1992;102(1):45-51.
115. Koochek AR, Yeh MS, Rolfe B, Richmond S. The relationship between Index of Complexity, Outcome and Need, and patients' perceptions of malocclusion: a study in general dental practice. *Br Dent J.* 2001;191(6):325-329.
116. Greco PM, English JD, Briss BS, et al. Posttreatment tooth movement: for better or for worse. *Am J Orthod Dentofacial Orthop.* 2010;138(5):552-558.

117. Lindauer SJ, Powell JA, Leypoldt BC, Tufekci E, Shroff B. Influence of patient financial account status on orthodontic appointment attendance. *Angle Orthod.* 2009;79(4):755-758.
118. Horsley BP, Lindauer SJ, Shroff B, et al. Appointment keeping behavior of Medicaid vs non-Medicaid orthodontic patients. *Am J Orthod Dentofacial Orthop.* 2007;132(1):49-53.
119. Bukhari OM, Sohrabi K, Tavares M. Factors affecting patients' adherence to orthodontic appointments. *Am J Orthod Dentofacial Orthop.* 2016;149(3):319-324.
120. Proffit WR, Fields HW, Sarver DM. *Contemporary Orthodontics-E-Book.* Elsevier Health Sciences; 2014.
121. Firestone AR, Hasler RU, Ingervall B. Treatment results in dental school orthodontic patients in 1983 and 1993. *Angle Orthod.* 1999;69(1):19-26.
122. Fox NA, Daniels C, Gilgrass T. 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). *Br Dent J.* 2002;193(4):225-230.
123. Yang-Powers LC, Sadowsky C, Rosenstein S, BeGole EA. Treatment outcome in a graduate orthodontic clinic using the American Board of Orthodontics grading system. *Am J Orthod Dentofacial Orthop.* 2002;122(5):451-455.
124. Vig KW, Weyant R, Vayda D, O'Brien K, Bennett E. Orthodontic process and outcome: efficacy studies--strategies for developing process and outcome measures: a new era in orthodontics. *Clin Orthod Res.* 1998;1(2):147-155.
125. Popowich K, Nebbe B, Heo G, Glover KE, Major PW. Predictors for Class II treatment duration. *Am J Orthod Dentofacial Orthop.* 2005;127(3):293-300.
126. Tsichlaki A, Chin SY, Pandis N, Fleming PS. How long does treatment with fixed orthodontic appliances last? A systematic review. *Am J Orthod Dentofacial Orthop.* 2016;149(3):308-318.
127. American Board of Orthodontics Discrepancy Index. https://www.americanboardortho.com/media/1186/discrepancy-index-worksheet-for-print.pdf. Accessed Nov 10, 2016.