TREATMENT STABILITY OF CLASS II PATIENTS TREATED WITH 2 OR 4 PREMOLAR EXTRACTION PROTOCOLS

A Thesis

by

CAMERON LEIGH KEYLOR

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Peter H. Buschang
Committee Members,	Gayle Glenn
	Emet Schneiderman
Head of Department,	Madhu K. Nair

May 2022

Major Subject: Oral Biology

ABSTRACT

Purpose

To determine the difference in treatment outcome and stability of Class II molar and Class I molar finished orthodontic cases using the American Board of Orthodontics Objective Grading System (OGS) in combination with irregularity index and buccal segment relationships.

Materials and Methods

The sample consisted of 48 patients who were retrospectively selected from four orthodontists. Group 1 included 24 patients treated with 2 maxillary first premolar extractions treated to a bilateral Class II molar relationship. Group 2 comprised 24 patients treated with 4 premolar extractions treated to a bilateral Class I molar relationship. Study models, panoramic radiographs, and lateral cephalograms were collected at three timepoints: pretreatment (T1), posttreatment (T2), and post-retention (T3). Molar and canine relationships, OGS scores, and mandibular incisor irregularity were measured at all three timepoints.

Results

There were not statistically significant between-group differences in canine relationships at any of the time points, but there were significant differences in molar relationship after treatment and post-retention. Group 1 experienced greater increases in malalignment posttreatment, but the difference was not statistically significant. The total OGS scores of both groups decreased significantly from T1 to T2, indicating improvements in treatment outcome, and decreased from T2 to T3, indicating posttreatment settling. Group 1 had a median posttreatment and post-retention score that was significantly greater than Group 2 suggesting a better treatment outcome for the group treated to a Class I molar finish. After treatment, Group 2 had significantly better alignment and occlusal relationships than Group 1. At post-retention,

significant differences remained for alignment and occlusal relationships; however, there were also significant differences in buccolingual inclination and marginal ridges.

Conclusions

Posttreatment and post-retention occlusion is worse in bilateral Class II molar finished patients than in Class I molar finished patients, with Class II molar finished patients scoring 17% and 44% higher total OGS scores. The anterior components of occlusion do not contribute significantly to the posttreatment between-group differences between 2 and 4 premolar extraction patterns. Posttreatment differences in overall OGS scores between 2 and 4 premolar extraction patterns are primarily due to the posterior components of occlusion, specifically alignment and occlusal relationships.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Peter Buschang, for his guidance and support throughout this study. His invaluable feedback and statistical analysis expertise made this research project possible. Thank you to my committee members, Drs. Gayle Glenn and Emet Schneiderman, for helping advise and overseeing the project to its completion.

I would also like to thank my co-residents and the Class of 2022. Your friendship has made me a better person and clinician. I am proud to call each of you colleagues. Thank you to my faculty members who have selflessly imparted a great deal of knowledge during my time in residency.

To my parents, Dave and Lisa, and my sister Kendall – thank you for your unconditional love and support with which you have provided me from day one. Lastly, thank you to my fiancé, Matthew, for your constant encouragement and positivity throughout my residency journey.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

The present work was supported by a thesis committee at Texas A&M College of Dentistry. A committee chair, Peter H. Buschang (Department of Orthodontics), and two committee members, Gayle Glenn (Department of Orthodontics) and Emet Schneiderman (Department of Biomedical Sciences) provided guidance and insight throughout the project and thesis preparation.

Thesis work was completed by the resident under the advisement of Peter H. Buschang (Department of Orthodontics).

Funding Sources

There were no funding sources for this thesis project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION & LITERATURE REVIEW

The primary objective of the proposed study was to compare the differences in the long-term stability of Class II molar and Class I molar finished cases. The changes that occur during treatment and over the post-retention period were investigated. To appreciate and understand the importance of this objective, it is pertinent to review the literature that is currently published.

Introduction to Occlusion:

It is well established within the orthodontic literature that normal occlusion is of foundational importance for function and stability; therefore, it is essential to elucidate a normal occlusal relationship. The most fundamental assessment of occlusion begins with Angle's Classification System. In 1899, Edward Angle published his "Classification of Malocclusion" in The Dental Cosmos based on the relationship of the buccal groove of the mandibular first molar and the mesiobuccal cusp of the maxillary first molar.[1] A normal occlusion includes a molar relationship in which the maxillary first and mandibular first molars exist in a Class I relationship. A Class II classification represents a relationship in which the maxillary first molar is positioned more anteriorly than the first mandibular molar buccal groove. Lastly, a Class III relationship signifies the opposite in which the maxillary first molar is positioned more posteriorly. Angle was not alone in defining a normal occlusion based on molar relationships. Andrews' "Six Keys to Normal Occlusion" describes the significant characteristics shared by all of the non-orthodontically treated patients whom he evaluated. Of note, molar relationship is the first of his six keys. He stated that the distal surface of the distobuccal cusp of the upper first molar should make contact and occlude with the mesial surface of the mesiobuccal cusp of the lower second molar. As a result, the canines and premolars have a cusp-embrasure relationship buccally and a cusp fossa relationship lingually.[2] The untreated models in Andrews' study consistently demonstrated a Class I relationship. Molar relationships contribute individually and collectively with the other five keys to the total scheme of occlusion and can be viewed as an essential component of successful orthodontic treatment.

Molar Relationships Effect on Function and Stability:

It is evident that the molar relationship is of utmost importance in defining and classifying occlusion, but it is also essential for function. The relationship between occlusion and posterior occlusal contacts has been extensively researched. Studies have shown that the biological basis for ideal Class I posterior occlusion includes maximum posterior occlusal contacts and near contacts.[3] Subjects with normal, Class I molar occlusion have significantly larger areas of contact and near contact (ACNC). Owens el al. revealed that patients with normal occlusion have significantly larger ACNC than those Class I, II or III malocclusions.[3] As a result, patients with normal Class I occlusion demonstrate better function as seen with enhanced masticatory performance,[4-7] normalized chewing cycles,[8] and greater bite forces.[9] The culmination of these various factors lends itself to the inherent stability that is apparent in Class I occlusal relationships.[10]

Increased bite forces are indirectly related to occlusal stability. A study by Bakke revealed that the number of occlusal tooth contacts is important in determining the maximum attainable bite force.[9] The occlusal surfaces of the molars play a key role in an individual's highest achievable bite force. In addition, malocclusions are often associated with reduced maximum bite forces. As a result, it is important to have well distributed contacts throughout the posterior dentition to enhance the amount of force delivery and provide stability during occlusal contact.

Class I has also been shown to be the most stable molar relationship.[2, 10] In a sample of untreated individuals (Bolton study), Harris and Behrents determined that the long-term stability of a Class I molar relationship is notably higher than Class II or Class III relationships. Class II and Class III malocclusions rarely self-correct and often become more severe with age. Furthermore, none of the patients' malocclusions spontaneously corrected to Class I. Arya et al. reported similar results in an untreated group of individuals from the Oregon dental school; molar malrelationships are rarely self-correcting.[11] Some would contend that there are other cusp-ridge-fossa-embrasure relationships that produce stability besides an ideal Class I relationship.[12, 13] For example, a full step mesiocclusion in which the distobuccal cusp of the maxillary first molar resides in the buccal groove of the mandibular first molar. Of the cases that presented with a *full step* Class II molar relationship, 86% changed with age while 14% remained the same. In contrast, none of the Class I relationships moved from the cusp-in-fossa occlusion.[10]

Stability vs. Relapse:

Long term stability should be a universal orthodontic treatment goal for all practitioners. When evaluating stability, it is imperative to define and differentiate between instability and relapse in orthodontic treatment. Relapse is a return towards the pretreatment condition, and the changes that occur are related to the treatment.[14-16] It describes changes that result from improper treatment, inadequate mechanics, or poor patient cooperation.[17] Relapse is often due to a disruption in the equilibrium; therefore, it happens quickly after orthodontic treatment. Instability is a result of changes in the craniofacial complex that would be anticipated whether or not orthodontic treatment took place.[17] Long term studies are required to properly evaluate and determine stability of orthodontic treatment and risk factors that contribute to instability.

Posttreatment irregularity is not only due to relapse but also due to instability.[17, 18] Normal age-related changes are seen in both untreated and treated individuals. Sinclair and Little evaluated dental casts from the Burlington Growth Center Study (Ontario, Canada) to determine the nature and extent of the developmental maturation process of the normal dentition.[19] Their results demonstrated there is a maturational decrease in arch length and intercanine width and minimal overall change in overjet, overbite, and intermolar width. Finally, malalignment, specifically incisor irregularity, increased from age 13-20 years old. The changes found in the untreated normal occlusion group were similar in nature but lesser in extent when compared to a sample of treated individuals.[19, 20] The most frequent overall findings are that malignment increases and arch dimensions decrease over time.

The current literature suggests that long-term stability of malocclusion corrections is an achievable goal. A connection between quality of the occlusion and the function and stability of the dentition has been suggested. The greater the quality of the orthodontic finished occlusion, the better the occlusal status at the post retention stage.[21] Begg and Kesling suggest that intercuspal relationships are imperative for the guided eruption of permanent teeth and coordinated facial growth.[22] A study by Ostyn el al. concluded that interdigitation plays a critical role in the control of vertical and anteroposterior facial growth and comprises an important factor in jaw relationships.[23] Experimental reduction of the posterior cusps in a group of growing monkeys resulted in critical discrepancies in jaw growth and arch development in anteroposterior, vertical, and transverse dimensions. In addition, it has been suggested that idealized intercuspation and good occlusal contacts could be the keys to a stable long term orthodontic result.[24] Kahl-Nieke et al findings support the supposition that intercuspation leads to stability. Their study indicated that a residual Class II or Class III molar relationship, which does not exhibit ideal occlusal contacts and intercuspation, are associated with post retention incisor irregularity and crowding.[25]

Class II Malocclusion Treatment Stability Treated to Class I Molar Relationships:

Correction of Class II malocclusions includes a variety of treatment modalities ranging from surgery, growth modification, functional appliances and elastics to extraction of premolars. There is evidence in the literature to suggest that Class II correction is stable when patients are treated to a Class I relationship. Patients presenting with Class II Division 1 malocclusion who are treated with fixed appliances have posttreatment molar relationships that are stable,[26, 27] while posttreatment overjet is slightly less stable.[27-29] Patients with similar malocclusions treated with functional appliances also present with dentoskeletal stability including sagittal occlusal relationships that are stable.[30, 31] In comparison to untreated controls, those treated with functional appliances exhibit slightly worse occlusal relationships but not clinically significant. The studies that evaluate functional appliances demonstrate greater posttreatment changes when no fixed treatment accompanies the appliance, the teeth are not finished in a Class I relationship, and hyperdivergent patients are treated.[31]

Likewise Class II Division 2 patients have stable posttreatment molar relationships when treated non-extraction with or without functional appliances[32]. An average of 7 years posttreatment showed a 0.5mm sagittal molar change toward a more Class II relationship.[33] Similarly, long-term 15 year posttreatment results showed a 0.1mm change in anteroposterior molar relationship with no change in canine relationship; however, posttreatment overbite was relatively unstable with an average of 40% return of initial overbite at 15 years.[6]

Extraction vs. Non-extraction Class II Malocclusion Treatment:

Class II malocclusion correction can also be accompanied by extractions. Often times the need for extractions is driven by facial esthetics and severe tooth size arch length discrepancies.[34] When comparing treatment results of a 4 premolar extraction vs. non-extraction protocol, individuals who underwent extractions tend to have more retrusive lips, flatter facial profiles, and more upright maxillary and mandibular incisors. Comparatively, the non-extraction groups had the opposite tendencies. In general, extractions do not significantly prevent or alter the *direction* of posttreatment changes, and the trends in posttreatment changes are similar in the maxillary and mandibular arches as well as in male and female patients.[27, 34, 35] Primary studies comparing extraction and non-extraction treatments have reported greater posttreatment malalignment in patients receiving extractions rather than non-extraction treatments. The differences range from 0.2-0.8 mm. In a systematic review conducted by Swidi et al., mandibular anterior teeth exhibited greater posttreatment irregularity in individuals whose premolars were extracted than in patients whose premolars were not extracted. The differences in long-term changes in mandibular incisor irregularity between extraction and non-extraction treatment was 0.34 mm, which is clinically insignificant.[36]

Some evidence suggests that regardless of extraction or non-extraction treatment, increases in intercanine width and arch length achieved during treatment of Class II patients tends to relapse.[33, 34] On the other hand, there is literature that supports that there is no difference in the percentage of relapse between groups with or without intercanine width increases.[27, 35]

When patients present with significant pretreatment mandibular anterior crowding, the intercanine width can be significantly decreased; therefore, it is nearly inevitable that the intercanine width will increase with treatment as the crowding unravels. This can help explain that disparity in results regarding stability of intercanine width. Posttreatment changes seen in both extraction and non-extraction groups include decreases in arch length and increases in tooth size arch length discrepancy accompanied by an increase in incisor irregularity.[27, 33, 34, 37, 38] The changes that are seen after orthodontic treatment in patients treated with or without extractions are related more to differential growth of the maxilla and mandible rather than the final position of the teeth at the end of treatment.[27]

Class I and Class II patients have similar percentages of relapse when comparing the two malocclusions in a non-extraction or extraction protocol. However, there is a tendency for Class II patients to have a greater relapse than Class I patients.[35, 39] Freitas et al. evaluated mandibular anterior crowding relapse and stability 5 years post-retention in Class I and Class II non-extraction patients. Class II patients had a higher percentage of incisor irregularity, but the difference was not statistically significant. This could be attributed to the anteroposterior relapse tendency in Class II patients. As the Class II relationship relapses and overjet increases, there could be an alteration in the equilibrium and balance of pressures from the tongue and lips. As a result, the relapse potential is greater in Class II individuals. Little et al. measured mandibular incisor irregularity in Class I and Class II mandibular first premolar extraction patients an average of 10 years post-retention. They determined that Class I Division 1 females had more post-retention irregularity than Class I females. However, when males and females were pooled, there was no significant difference.

Class II Malocclusions Treated to Class II Molar Relationships:

Since Class I occlusal relationships and stability have been well documented, this begs the question as to why orthodontists occasionally extract only maxillary premolars and finish treatment with a Class II molar relationship. The orthodontic literature from the time of Bolton contains articles with the molars finished in a Class II relationship.[40] Classic orthodontic textbooks by Bishara and Proffit document cases in which molars are treated to finish with Class II relationships.

In 1963, Kessel defended his thesis arguing that nongrowing patients with a particular type of Class II Division 1 malocclusion could justifiably receive single arch extractions as a treatment modality.[12] He contended that patients 12 years and older exhibiting a Class II skeletal and dentoalveolar relationship with minimal crowding in the mandibular arch would benefit considerably from maxillary premolar extractions. He even went as far as to speculate that retention is augmented considerably due to the favorable axial inclinations of upper and lower anterior teeth.[12] In agreement, Graber asserted that the concept of occlusion must be modified at times where skeletal dysplasias and unfavorable apical base relationships make normal occlusal interdigitation difficult to obtain.[41] There are, however, objections to this form of treatment. Loughlin summarized his concerns with the notion that when the arches are left in a Class II relationship, inclined planes, ridges, and grooves which were not meant to occlude are left to do just that.[42] He reasoned that with 2 maxillary premolar extractions there is a disharmony in tooth size in occlusion essentially creating a Bolton discrepancy in the posterior dentition. He went as far as to say that even with judicious interproximal reduction in the posterior segment, the results are not as sound as a Class I molar occlusion.

Treatment time and efficiency are also important considerations driving orthodontic treatment modalities. It comes as no surprise that thorough treatment and the correction of more complex malocclusions requires longer treatment times.[43, 44] Specifically focusing on the correction of complete Class II malocclusions, it has been reported that treatment times are shorter with a 2 maxillary premolar extraction protocol compared to 4 premolar extractions.[45] Janson et al. evaluated the PAR index score of pretreatment and posttreatment records to calculate the percentage of PAR score reduction. Treatment efficiency was then determined by the ratio between percentage of PAR reduction and treatment time. Treatment efficiency was greater in 2 maxillary premolar extraction procedures compared to a non-extraction approach.[46] Orthodontists are constantly looking for ways to minimize patient treatment time; therefore, it makes sense as to why finishing with a Class II molar relationship is a popular treatment modality for Class II patients.

Class II Molar Treatment Outcomes:

Although full step Class II molar orthodontic finish is a fairly common treatment objective, there are few studies that have investigated either the treatment effect or the long-term stability of these treated cases. In 2004, Janson et. al[47] retrospectively evaluated the treatment success of 2 and 4 premolar extraction protocols using Grainger's treatment priority index (TPI) in the university setting. A sample of 131 patients were identified into 2 groups, including 81 patients treated with 2 maxillary premolar extractions and 50 patients treated with 4 premolar extractions. The patients were treated by multiple orthodontic residents undergoing orthodontic training. Selection criteria were based solely on the anteroposterior dental relationship, regardless of any other dentoalveolar or skeletal relationship. The TPI score was calculated prior to the initiation of treatment and at the time of fixed appliance removal. The study did not include a long-term post-retention timepoint. TPI score categories included overjet, overbite, tooth displacement (alignment), and posterior cross bites. The TPI does not specifically evaluate anteroposterior relationships of the molars or canines, nor does it focus on posterior occlusion. In addition, the assessment of each category is less stringent than other indices such as the ABO Objective Grading System. For example, tooth displacement does not consider teeth misaligned if there is less than a 45-degree rotation or 2 mm of displacement. The results of the study showed that the TPI was significantly ($p = <0.001$) smaller and the change from initial to final records was significantly greater ($p=0.01$) in the 2 premolar extraction group. The final overjet ($p=<0.001$), overbite ($p=<0.001$), and anterior-posterior canine discrepancy ($p=<0.001$) were all significantly smaller in the 2 premolar extraction group. The findings indicate that the 2 premolar extraction cases provided better occlusal results than 4 premolar extraction cases based off of final TPI scores.

Cansunar and Uysal performed a study retrospectively that measured occlusion using the ABO Objective Grading System (OGS). They compared 2 premolar extraction versus 4 premolar extraction versus non-extraction treatment outcomes. However, they did not evaluate long-term stability results.[48] A sample of 1098 patients was divided into 3 groups: 269 patients treated with maxillary premolar extractions, 267 patients treated with 4 premolar extractions, and 562 patients treated non-extraction. The study was conducted in a university setting in Turkey with data compiled from nine different postgraduate orthodontic programs. The cross-sectional design evaluated treatment success immediately posttreatment. The OGS evaluates the following categories: alignment, marginal rides, buccolingual inclination, occlusal contacts, occlusal relationship, overjet, interproximal contacts, and root angulation. The selection criteria for the sample included patients with regularly attended appointments who started and completed treatment with the same orthodontist. The patients' records included posttreatment dental casts and panoramic radiographs. The inclusion criteria did not incorporate the patient's skeletal or dentoalveolar relationships at the beginning of treatment.

The study showed that there were no significant between group differences for alignment, marginal rides, buccolingual inclination or overjet at the end of treatment. The significant between group differences identified included occlusal contacts (p=0.02), occlusal relationship (p= <0.001) and root angulation (p=0.01), with the non-extraction group having the best posttreatment scores, followed by the 2 premolar extraction group, and the 4 premolar extraction group. The percentage of passing cases, which was determined by an OGS score of <20 points, was not significant between groups. However, the highest percentage of passing cases was found in the non-extraction group, followed by 2 premolar extraction group, and finally the 4 premolar extraction group.

In 2010, Janson et al[49] performed another study in the same university setting evaluating treatment stability of Class II malocclusion treated with 2 maxillary premolar extractions versus patients treated non-extraction. This retrospective study evaluated two groups, including 29 patients treated without extractions and 30 patients treated with 2 maxillary premolar extractions. The patients were treated by multiple orthodontic residents undergoing orthodontic training. Treatment success was evaluated and scored using the Peer Assessment Rating (PAR) index at 3 time points: pretreatment, posttreatment, and post-retention (minimum of ~2.4 years after treatment was completed). The initial age of the non-extraction group was significantly (p=0.005) younger than the 2 maxillary premolar extraction group. As a result, the non-extraction patients had more growth remaining than the other group which could contribute to long-term stability. The PAR index evaluates the following categories: molar and premolar anterior-posterior relationship overjet, overbite, midlines, crossbite, and crowding. Selection criteria was based solely on the anterior posterior dental relationship, regardless of any other dentoalveolar or skeletal relationship. It is important to note that the PAR index evaluates buccal occlusion based on half step deviations from ideal intercuspation, which is not a precise method to determine occlusal discrepancies. The results of the study showed no intergroup differences in the overall PAR score at T3, no difference in posttreatment occlusal changes, and no difference in percentage of posttreatment occlusal changes. There were no significant correlations between the treatment or the posttreatment occlusal changes that occurred. Finally, there was no difference in stability between patients in treatment less than 2 years and those in treatment more than 3 years. The study concluded that there is no difference in occlusal stability in treatment of patients with complete Class II malocclusion without extraction or with 2 premolar extractions; therefore, finishing treatment with the molars in a Class I or Class II relationship provides similar stability in patients who present with Class II malocclusion.

The results of the treatment outcomes of patients treated to complete Class I molar relationships versus those treated to bilateral Class II molar relationships are controversial and incomplete. According to the previous literature, patients treated with 2 maxillary premolar extractions present with better occlusal results at posttreatment compared to patients treated with 4 premolar extractions and non-extraction. When considering long-term stability, the literature seems to support the notion that there is not a significant difference between patients treated with 2 maxillary premolar extractions and those treated non-extraction; however, there is not any data that compares long-term stability of 2 maxillary premolar extractions and 4 maxillary premolar extractions. In addition, there are shortcomings and limitations to the current literature that evaluates the posttreatment and long-term stability results of Class I and Class II molar relationships that were previously mentioned.

Summary and Introduction to Current Study:

If there is a difference in long-term stability between Class I and Class II molar finishes, it is imperative that orthodontists are informed as to which treatment will provide a better result. In knowing which components of occlusion are more prone to instability, orthodontists will be better equipped to devise retention protocols and provide patients with the best care. Additionally, with the rise of DIY aligner companies who make claims that posterior occlusion is inconsequential, orthodontists need more evidence that critically evaluates occlusion and its effect on long-term stability.

To date, there is not published literature that compares *long-term* stability of treatment outcomes between 2 premolar and 4 premolar extraction groups. The studies that compare 2 versus 4 premolar extractions do not include observation after the removal of appliances. The results of the treatment outcomes are controversial. In addition, all of the previously mentioned studies are conducted using samples from university settings where multiple practitioners undergoing orthodontic training treated the patients. The treatment quality is parallel to the orthodontist's level of orthodontic experience, thus studies conducted in private practice are provide better external validity. The inclusion criteria for the previously mentioned studies were limited and the results published in the current literature is inconsistent. Lastly, the ABO OGS includes more rigorous standards and criteria than the PAR or TPI scoring indices.

The aim of the current study was to determine the difference in long-term stability of Class II molar and Class I molar finished cases using the ABO Objective Grading System, in combination with irregularity index and linear measurement of the buccal segment relationships to analyze treatment and stability outcomes. Dental casts and radiographs were gathered for each patient at three time points: pretreatment, appliance debond (posttreatment), and long-term post retention. Data were collected from the patient records at four private practices to determine ABO OGS score, irregularity index, and molar/canine sagittal relationships. Based on the foregoing literature, it was assumed that the Class II malocclusions corrected to a Class I molar relationship would be more stable long-term. The hypothesis was that there will be a significant difference in long term stability between the two groups. The null hypothesis was that there will be no significant differences in long term stability between the two groups.

MATERIALS AND METHODS

Patient Sample:

The study included 48 patients who were retrospectively selected from four private practice board certified orthodontists. All 48 patients initially had bilateral Class II Division 1 malocclusions and were treated with standard fixed appliances.

Sample selection was based on the following inclusion criteria:

- Full step Class II molar and canine malocclusion
- Extractions of either 2 maxillary premolars or 2 maxillary and 2 mandibular premolars
- No missing teeth (excluding 3rd molars) at the beginning of treatment
- Anteroposterior skeletal relationships (ANB > 3)
- FMA between 20-30*
- Complete sets of orthodontic records (dental casts, panoramic radiographs, and lateral cephalograms)

Patients were excluded from the study if they met any of the following exclusion criteria:

- Broken study models
- Treatment was not completed

The retrospective longitudinal study was reviewed and approved by Texas A&M College of Dentistry's Institutional Review Board (IRB2020-0876-CS-EXM).

To ensure sufficient sample sizes, an a priori power analysis was performed using the descriptive statistics reported for similar treatment outcomes (Janson et al.) Based on a 5% chance of making a Type I statistical error, a 20% chance of a Type II error (power of 80%), and an effect size of 0.73, 24 patients were required per group.

Group 1 included 24 patients treated with 2 maxillary first premolar extractions (3 males and 21 females) treated to a bilateral Class II molar relationship. Group 2 comprised 24 patients treated with 4 premolar extractions (4 males and 20 females) treated to a bilateral Class I molar relationship. There was no statistically significant sex by group difference (Table 1). Of the patients treated with 4 premolar extractions, 15 were treated with maxillary first premolar and mandibular second premolar extractions and 9 were treated with 4 first premolar extractions. The four private practices provided similar proportions of patients in Groups 1 and 2 (Table 2).

Orthodontic Treatment:

The median pretreatment age for the 2 premolar extraction group was 14.3 (IQR 12.0, 20.7) years with a median debond age of 16.6 (IQR 15.5, 23.0) years. The median pretreatment and debond ages for the 4 premolar extraction group were 12.7 (IQR 12.2, 17.0) and 15.5 (IQR 14.4, 19.7) years respectively. The median post retention ages were 33.0 (IQR 26.5, 40.2) and 34.8 (IQR 29.2, 40.5) years for the 2 premolar and 4 premolar extraction groups, respectively. There were no statistically significant between-group differences in age at T1, T2, or T3 (Table 3). Treatment duration was 2.35 and 2.45 years for the 2 premolar and 4 premolar extraction groups, with no statistically significant between group difference in duration.

Among the four private practices, orthodontic treatment was completed using either 018” Alexander bracket prescription, 022” 0-degree standard edgewise or 018” standard edgewise. Anchorage included cervical pull or J-hook headgear at the discretion of the treating orthodontist. Inter-arch elastics were prescribed and used as needed to detail each patient’s occlusion.

Patients from each practice were retained similarly. A stainless steel wire was adapted to the lingual aspect of the mandibular anterior teeth and was fixed in place via bands on the mandibular canines. Patients were instructed to remain in fixed mandibular retention for 2-3 years posttreatment. Impressions were used to fabricate maxillary Hawley retainers. Post retention records were acquired an average of 15.1 years after completed orthodontic treatment.

Data Collection:

Pretreatment initial records (T1), day of debond posttreatment records (T2), and long-term post retention (T3) records were collected, including study models, panoramic radiographs, and lateral cephalograms. Lateral cephalograms were scanned into Dolphin Imaging Software and digitized. The lateral cephalograms were traced to measure Frankfort-mandibular plane angle (FMA) and ANB at pretreatment initial records only. Molar and canine relationships were measured at all three timepoints parallel to the occlusal plane in half millimeter increments from the mesiobuccal cusp tip of the maxillary first molar to the buccal groove on the mandibular first molar and the maxillary canine cusp tip to the embrasure between the mandibular canine and first premolar, respectively.[14] For the patients who finished with Class II molar relationships (maxillary premolar extractions only), the measurement was taken from the mesiobuccal cusp tip of the maxillary first molar to the embrasure between the mandibular first molar and second premolar.

The ABO Objective Grading System scores and mandibular incisor irregularity were also measured at all three occasions.[50, 51] The ABO OGS included 8 criteria/measurements: alignment, marginal ridges, buccolingual inclination, occlusal relationships, occlusal contacts, overjet, interproximal contacts, and root angulation. The ABO measuring gauge was used to score the casts (Figure 1). A score of 0 indicated ideal alignment or occlusion, while scores of 1 and 2 showed deviations from ideal. The individual component scores were summed to provide a total ABO OGS score.

Since Group 2 subjects had a total of 4 premolars extracted during treatment, Group 1 subjects had 2 more mandibular premolars present for scoring at posttreatment and post-retention. To equalize the number of teeth scored in both groups at T2 and T3, mandibular first and second premolars were randomly selected for elimination in Group 1 patients. Among Group 2 subjects, there were 9 who had mandibular first premolars extracted and 15 who had mandibular second premolars extracted. As a result, 9 sets of mandibular first premolars and 15 sets of mandibular second premolars were not scored from Group 1 patients.

Mandibular incisor irregularity was calculated using Little's irregularity index.[50] The linear distance from anatomic contact point to adjacent anatomic contact point of the mandibular anterior teeth were measured from canine to canine and summed to represent the irregularity index. Interproximal spacing between the mandibular anterior teeth was not scored as part of the irregularity index.

The primary investigator (C.L.K) was blinded when performing each of the measurements. Duplicate measurements on 12 randomly selected casts and radiographs showed no statistically significant systematic errors. The method errors $\sqrt{(\Sigma \text{ deviations}^2/2N)}$ for OGS (0.61 points), molar relationship (0.05 mm), canine relationship (0.05 mm) and incisor irregularity (0.18 mm) were small. The intraclass correlations for OGS (0.999, $p<0.05$), molar and canine relationships (0.999, $p<0.05$), and incisor irregularity (0.994, $p<0.05$) were high.

Statistical Methods:

One investigator recorded each of the measurements on a Microsoft Excel spreadsheet. The data were transferred to SPSS Version 26.0 (IMB® Corp., Armonk NY) for statistical analyses. Normality of the data was assessed using skewness and kurtosis statistics. FMA and ANB were normally distributed. The rest of the data were not normally distributed requiring medians and interquartile ranges (IQRs) to be used for descriptive purposes. Mann-Whitney U-Tests were used to evaluate the inter-group differences.

Bonferroni correction was performed for the statistics carried out on Objective Grading System component scores to counteract the problem of multiple comparisons ($p = \alpha/n$). The corrected p-value was $p = 0.006$. P-values greater than $p = 0.006$ and less than $p = 0.05$ were considered marginally significant.

RESULTS

Sample Breakdown:

There were no statistically significant pretreatment between-group differences in the FMA or ANB angles (Figure 2). The median pretreatment ABO OGS total scores were 75.5 (IQR 66.3, 80.8) and 73.0 (IQR 62.0, 81.3) for Groups 1 and 2 respectively, with no statistically significant between-group pretreatment difference (Figure 3).

Molar and Canine Relationships and Incisor Irregularity:

There were statistically significant changes in anteroposterior canine relationships for both groups during treatment followed by non-statistically significant changes in canine relationships posttreatment. The between-group differences were not significant (Table 4). In addition, there were no statistically significant between-group differences in canine relationships at any of the 3 time points (Figure 4a, Table 5).

There were significant improvements in anteroposterior molar relationships from pretreatment to posttreatment (T1-T2), but no statistically significant changes from posttreatment to post-retention (T2-T3) for either group. There were also statistically significant between-group differences from pretreatment to posttreatment ($p = 0.01$), with the 4 premolar extraction group showing greater improvement (Table 4). While there were no statistically significant differences in molar relationships prior to treatment, there were significant differences in molar relationship after treatment ($p = <0.001$) and post retention ($p = <0.001$) (Figure 4b, Table 5).

Both groups showed significant reductions in malalignment of the lower incisors during treatment. The 2 premolar extraction group experienced slightly greater increases posttreatment, but the difference was not statistically significantly ($p = 0.21$) (Table 5).

Overall ABO Objective Grading System Score:

The total OGS scores of both groups decreased significantly from T1 to T2, indicating improvements in treatment outcome, and decreased from T2 to T3, indicating some posttreatment settling (Table 6). The 4 premolar extraction group experienced a statistically significant reduction in OGS score from T2 to T3, while the 2 premolar extraction group reduction was not significant. There was no statistically significant between-group difference ($p = 0.32$) in the changes that occurred during treatment (T1-T2). From posttreatment to post-retention (T2-T3), the 2 maxillary premolar extraction group showed smaller decreases in total OGS scores than the 4 premolar extraction group, with the difference approaching statistical significance ($p = 0.08$). There was a statistically significant between-group difference in the changes that occurred from pretreatment to post-retention ($p = 0.04$), with the 4 premolar extraction group showing greater decreases (Figure 5, Table 6).

The 2 maxillary premolar extraction group had a median posttreatment ($p = 0.02$) and post-retention ($p = <0.001$) score that was significantly greater than the 4 premolar extraction group suggesting a better treatment outcome for the Group treated to a Class I molar finish (Figure 6, Table 7).

ABO Objective Grading System Component Scores:

Except for buccolingual inclination ($p = 0.02$) which was higher in the 2 premolar extraction group, there were no significant between-group differences for any of the Objective Grading Score components prior to treatment (T1) (Table 8).

After orthodontic treatment (T2), the 4 premolar extraction group had significantly better alignment ($p = <0.001$), interproximal contacts ($p = 0.01$), and occlusal relationships (p =

<0.001) than the 2 premolar extraction group. In contrast, root angulation (p = 0.01) was significantly better in the 2 premolar extraction group. After Bonferroni corrections, the differences for interproximal contact and root angulation were marginally significant (Figure 7, Table 9).

Similar results were observed post-retention (T3) with a few noteworthy differences (Figure 8, Table 10). Alignment and occlusal relationships were the highest scoring components for both groups. Significant between-group differences remained for alignment (p = <0.001) and occlusal relationships (p = 0.003); however, there were also significant differences in buccolingual inclination (p = <0.001) and marginally significant differences in marginal ridges (p = 0.02).

Changes T1-T2, T2-T3, T1-T3:

Each of the component scores decreased significantly during treatment for the 2 premolar extraction group except for occlusal contacts. Each of the component scores decreased significantly during treatment for the 4 premolar extraction group except for buccolingual inclination and root angulation. There were no statistically significant between-group differences in the treatment changes that occurred except for a marginally significant difference in occlusal relationship (p = 0.05), with the 4 premolar extraction group exhibiting a greater improvement (Table 11).

Alignment was the only component score that increased significantly posttreatment (Table 12). Marginal ridges, occlusal contacts, and root angulation decreased significantly in both groups. Buccolingual inclination significantly decreased in the 4 premolar extraction group, while interproximal contacts significantly decreased in the 2 premolar extraction group. There were significant between-group differences in root angulation (p = <0.001), with the 2 premolar extraction group exhibiting a smaller reduction in root angulation score. The changes for the remaining components were not significant.

Lastly, the changes from initial records to post retention showed that all of the component scores improved significantly for both groups (Table 13). Alignment ($p = 0.004$) and overjet ($p = 0.05$) were the only two components that revealed marginally significant between-group differences.

DISCUSSION

Mandibular incisor irregularity is not affected by the different extraction patterns. The 2 maxillary premolar extraction and 4 premolar extraction treatments in the present study did not exhibit significant between-group differences in mandibular incisor irregularity posttreatment or post-retention. The posttreatment incisor irregularity index was 0 mm for both groups, followed by 2.3 mm and 1.5 mm post-retention increases for the 2 and 4 premolar extraction groups, respectively. There was no significant between-group difference in posttreatment changes. Importantly, the incisor irregularity index was measured to the nearest half millimeter rather than to the nearest tenth of a millimeter. This method likely resulted in an inability to detect a significant between-group difference. Studies comparing extraction and non-extraction treatments post-retention have found no significant between-group difference in incisor irregularity.[27, 36, 37] However, no other study has evaluated posttreatment irregularity changes of 2 maxillary premolar and 4 premolar extraction cases. A large proportion of untreated subjects with acceptable occlusion develop incisor irregularity by their early twenties.[19, 52] Therefore, changes in posttreatment incisor irregularity occur regardless of treatment protocols with or without extractions.

Posttreatment occlusion is worse in bilateral Class II molar finished patients than in Class I patients. The 4 premolar extraction group in the present study had overall OGS scores that were 17% lower at the end of treatment, indicating better occlusion, than the 2 maxillary premolar extraction group. Overall posttreatment OGS scores have been previously reported to be the same for 2 maxillary premolar and 4 premolar extractions.[48] However, the severity of the initial malocclusion was not specified, which could explain why they had a higher percentage of passing cases and the lack of a statistically significant between-group difference. Janson et al.

found significantly ($p<0.001$) better posttreatment TPI scores in patients treated with 2 maxillary premolar extractions than patients treated with 4 premolar extractions. The difference between this and the present study could be due to the components of the TPI, which primarily focus on anterior occlusion.[53] The higher ABO OGS scores of the 2 maxillary premolar extraction group in the present study are likely attributed to the posterior, rather than anterior components of occlusion. Neither of the previously mentioned studies controlled for treatment providers; more skilled clinicians could have treated more patients in the 2 maxillary premolar extraction group, resulting in better scores.

The anterior components of occlusion do not contribute significantly to posttreatment between-group differences between 2 and 4 premolar extraction patterns. In the present study, there were no significant between-group differences in mandibular incisor irregularity, canine relationships, or overjet at the end of treatment. Better overjet and anteroposterior canine relationships have been reported in patients treated with 2 maxillary premolar extractions than 4 premolar extractions.[47] This suggests that the 4 premolar extraction patients were not being treated to Class I canine relationships, which could affect final overjet. The increased overjet they reported could also be due to anterior Bolton discrepancies that were not addressed with interproximal reduction. In the present study, both groups were treated to bilateral Class I canine relationships with proper overjet, which explains the lack of between-group differences for the anterior occlusal components.

Posttreatment differences in overall OGS scores between 2 and 4 premolar extraction patterns are primarily due to the posterior components of occlusion. Posttreatment occlusal relationships and alignment were significantly better in patients treated to Class I molar relationships. Patients treated with 2 maxillary premolar extractions have been previously shown to have better occlusal relationships than patients treated with 4 premolar extractions.[48] This could have been due to the patients' anteroposterior or vertical skeletal discrepancies, which were not controlled. If more of the patients treated to Class I molar relationships were vertical growers and treated with extrusive mechanics, a worse posterior occlusal finish might be expected. To ensure for proper intercuspation of the posterior dentition, the distal surface of the distobuccal cusp of the maxillary first molar should make contact and occlude with the mesial surface of the mesiobuccal cusp of the mandibular second molar to allow.[2] When 2 maxillary premolars are extracted and the buccal segments are finished with Class II molar relationships, the maxillary first molar's distobuccal cusp occludes in the buccal groove of the mandibular first molar rather than resting in the embrasure space. This explains why higher occlusal relationship scores might be expected in the 2 maxillary premolar extraction cases.

Posttreatment posterior alignment is a major reason why patients treated with 4 premolar extractions had better overall posttreatment OGS scores than patients treated with 2 maxillary premolar extractions. The patients in the present study treated with 2 maxillary premolar extractions finished treatment with mesial in rotations of their maxillary first molars. No statistically significant difference in alignment scores has been previously reported for 2 maxillary premolar versus 4 premolar extractions.[48] However, it was not specified if the patients were treated with mesial in rotations, which could explain why they found no between-group differences. The ABO OGS scoring system uses the mesiodistal central groove to assess the alignment of the maxillary posterior dentition.[51] This could explain the increased posttreatment malalignment for the patients treated with 2 maxillary premolar extractions in the present study.

Bilateral Class II molar finished patients also have worse occlusion post-retention. In the present study, the 4 premolar extraction group had overall OGS scores that were 44% lower than the 2 maxillary premolar extraction group 15.1 years posttreatment. Both groups showed posttreatment reductions in overall ABO OGS scores. The decrease was significant for the 4 premolar extraction group but not for the 2 maxillary premolar extraction group. The between-group difference in posttreatment changes approached statistical significance meaning cases treated to Class I molar relationships allowed for more posttreatment settling and greater stability. Other studies evaluating ABO OGS posttreatment changes found that overall scores tend to significantly improve.[54, 55] No other study has evaluated the long-term post-retention occlusion of 2 versus 4 premolar extraction cases using ABO OGS scores. Superior post-retention scores in the 4 premolar extraction group were likely partially due to the significantly lower scores that were already present at the end of treatment. The between-group difference of posttreatment changes in alignment, buccolingual inclination, and root angulation scores were also important contributing factors in the significant difference in overall post-retention OGS scores.

After treatment, ABO component scores significantly improved for marginal ridges, occlusal contacts, and root angulations in both groups. Median scores reduced 70% and 67% (occlusal contacts), 60% and 66% (marginal ridges), and 25% and 66% (root angulations) for patients treated with 2 maxillary premolar and 4 premolar extractions, respectively. There were not significant posttreatment between-group differences for occlusal contacts, but the 4 premolar extraction group had significantly lower marginal ridge scores posttreatment. The greater improvements in posttreatment root angulation among the 4 premolar extraction patients was expected because they finished treatment with higher scores. The improvements confirm previous literature showing posttreatment occlusal settling. It has been previously shown that marginal ridges (decrease of 1.25 points) and occlusal contacts (decrease of 2.5 points) scores improve significantly 10 years posttreatment.[54] The improvements can be explained by Van der Linden's cone-funnel hypothesis, which suggests that the dentition settles to gain occlusal and masticatory coordination following treatment because the teeth are permitted to function as individual units.[23, 56]

Posttreatment alignment scores and incisor irregularity worsen regardless of extraction pattern. Posttreatment alignment scores increased 3.5 points and 1.0 point for patients treated with 2 maxillary premolar and 4 premolar extractions, respectively. Even though the between-group difference was not statistically significant, the lack of significant difference could be due to inadequate power to detect a between-group difference. Previous studies have confirmed these findings with significant ($p = < 0.001$) changes in ABO alignment scores posttreatment.[54, 57] Comparisons of posttreatment anterior and posterior alignment changes have shown the posterior alignment remains relatively stable, while anterior alignment worsens.[54, 58] However no studies have specifically compared Class I and Class II molar finished cases. Previous studies have also reported that alignment is variable and unpredictable in the long-term.[39] In the present study, the posttreatment changes in alignment scores for both groups can be attributed to the increase in anterior incisor irregularity but could also be attributed to maxillary anterior alignment as well as posterior alignment.

The most noteworthy clinical implication of the present study is, when possible, the orthodontist should make his or her treatment objective to achieve bilateral Class I molar relationships for his or her patients. Class I molar occlusion provides better treatment outcomes and long-term stability than Class II molar occlusion. In comparison to malocclusions, Class I occlusion leads to greater areas of contact which in turn provides greater bite forces and enhanced function. This ultimately leads to greater long-term stability.[3, 5, 9] When Class II patients are treated with functional appliances, they demonstrate greater posttreatment changes when the teeth are not finished in a Class I relationship.[31] In untreated individuals, the intrinsic stability of Class I molar relationships is superior to Class II or III relationships.[10] In the present study, patients treated to Class I molar relationships had overall OGS scores that were 44% lower than the Class II molar finished group, which supports Class I molar relationships provide for better long-term stability.

This study is not without limitations. One of the limitations was retention compliance. Patients were instructed to remain in fixed retention for 2-3 years after treatment; however, it was not documented in every patient chart when fixed retention was removed. As a result, there is no way to know or compare if one group remained in fixed retention for a longer period of time than the other group. If the patients treated with 4 premolar extractions were kept in retention significantly longer, this could affect the alignment between-group difference in alignment scores.

Although a priori power analysis was performed, the sample size could have been too small to detect significant between-group differences for some of the OGS score components. Lastly, the mandibular incisor irregularity index should have been measured to the nearest tenth of a millimeter rather than half millimeter increments. This could have allowed for better detection of between-group differences in alignment of the mandibular anterior teeth.

CONCLUSIONS

- Mandibular incisor irregularity is not affected by the two different extraction patterns.
- Posttreatment occlusion is worse in bilateral Class II molar finished patients than in Class I molar finished patients, with Class II molar finished patients scoring 17% higher total OGS scores.
- The anterior components of occlusion do not contribute significantly to the posttreatment between-group differences between 2 and 4 premolar extraction patterns.
- Posttreatment differences in overall OGS scores between 2 and 4 premolar extraction patterns are primarily due to the posterior components of occlusion, specifically alignment ($p=<0.001$) and occlusal relationships ($p=<0.001$).
- Bilateral Class II molar finished patients have worse occlusion post-retention. Patients treated with 4 premolar extractions and treated to Class I molar relationships scored 44% better long-term than patients treated with 2 maxillary premolar extractions.
- After treatment, OGS component scores significantly improve for marginal rides, occlusal contacts, and root angulations in both groups.
- Posttreatment alignment scores and incisor irregularity worsen regardless of extraction pattern.

REFERENCES

1. Angle, E., *Classification of malocclusion.* Dental Cosmos, 1899(41): p. 248-357.

2. Andrews, L.F., *The six keys to normal occlusion.* American Journal of Orthodontics, 1972. **62**(3): p. 296-309.

3. Owens, S., et al., *Masticatory performance and areas of occlusal contact and near contact in subjects with normal occlusion and malocclusion.* American Journal of Orthodontics and Dentofacial Orthopedics, 2002. **121**(6): p. 602-609.

4. English, J.D., P.H. Buschang, and G.S. Throckmorton, *Does malocclusion affect masticatory performance?* Angle Orthod, 2002. **72**(1): p. 21-7.

5. Kobayashi, T., et al., *Masticatory function in patients with mandibular prognathism before and after orthognathic surgery.* Journal of Oral and Maxillofacial Surgery, 1993. **51**(9): p. 997-1001.

6. Bae, J., et al., *Comparison of masticatory efficiency according to Angle's classification of malocclusion.* The Korean Journal of Orthodontics, 2017. **47**(3): p. 151.

7. Henrikson, T., E.C. Ekberg, and M. Nilner, *Masticatory efficiency and ability in relation to occlusion and mandibular dysfunction in girls.* Int J Prosthodont, 1998. **11**(2): p. 125-32.

8. Buschang, P.H., et al., *Chewing cycle kinematics of subjects with deepbite malocclusion.* American Journal of Orthodontics and Dentofacial Orthopedics, 2007. **131**(5): p. 627-634.

9. Bakke, M., *Bite Force and Occlusion.* Seminars in Orthodontics, 2006. **12**(2): p. 120-126.

10. Harris, E.F. and R.G. Behrents, *The intrinsic stability of Class I molar relationship: A longitudinal study of untreated cases.* American Journal of Orthodontics and Dentofacial Orthopedics, 1988. **94**(1): p. 63-67.

11. Arya, B.S., B.S. Savara, and D.R. Thomas, *Prediction of first molar occlusion.* American Journal of Orthodontics, 1973. **63**(6): p. 610-621.

12. Kessel, S.P., *The rationale of maxillary premolar extraction only in Class II therapy.* American Journal of Orthodontics and Dentofacial Orthopedics, 1963. **49**(4): p. 276-293.

13. Proffit, W., *Contemporary Orthodontics*. 6th ed. 2015.

14. Dyer, K.C., J.L. Vaden, and E.F. Harris, *Relapse revisited—again.* American Journal of Orthodontics and Dentofacial Orthopedics, 2012. **142**(2): p. 221-227.

15. Vaden, J.L., E.F. Harris, and R.L.Z. Gardner, *Relapse revisited.* American Journal of Orthodontics and Dentofacial Orthopedics, 1997. **111**(5): p. 543-553.

16. Thilandera, B., *Orthodontic relapse versus natural development.* American Journal of Orthodontics and Dentofacial Orthopedics, 2000. **117**(5): p. 562-563.

17. Horowitz, S.L. and E.H. Hixon, *Physiologic recovery following orthodontic treatment.* American Journal of Orthodontics, 1969. **55**(1): p. 1-4.

18. IL, N., *Growth considerations in stability of orthodontic treatment*. 1993, Philaadelphia: W.B. Saunders

19. Sinclair, P.M. and R.M. Little, *Maturation of untreated normal occlusions.* American Journal of Orthodontics, 1983. **83**(2): p. 114-123.

20. Little, R.M., *Stability and Relapse of Dental Arch Alignment.* British Journal of Orthodontics, 1990. **17**(3): p. 235-241.

21. de Freitas, K.M., et al., *Influence of the quality of the finished occlusion on postretention occlusal relapse.* Am J Orthod Dentofacial Orthop, 2007. **132**(4): p. 428.e9-14.

22. PR Begg, P.K., *Begg's orthodontic theory and technique*. 1971: WB Saunders.

23. Ostyn, J.M., et al., *The role of interdigitation in sagittal growth of the maxillomandibular complex in Macaca fascicularis.* American Journal of Orthodontics and Dentofacial Orthopedics, 1996. **109**(1): p. 71-78.

24. Parkinson, C.E., et al., *A new method of evaluating posterior occlusion and its relation to posttreatment occlusal changes.* Am J Orthod Dentofacial Orthop, 2001. **120**(5): p. 503-12.

25. Kahl-Nieke, B., H. Fischbach, and C.W. Schwarze, *Post-retention crowding and incisor irregularity: a long-term follow-up evaluation of stability and relapse.* Br J Orthod, 1995. **22**(3): p. 249-57.

26. Harris, E.F., et al., *Effects of patient age on postorthodontic stability in Class II, Division 1 malocclusions.* American Journal of Orthodontics and Dentofacial Orthopedics, 1994. **105**(1): p. 25-34.

27. Paquette, D.E., J.R. Beattie, and L.E. Johnston, *A long-term comparison of nonextraction and premolar extraction edgewise therapy in "borderline" Class II patients.* American Journal of Orthodontics and Dentofacial Orthopedics, 1992. **102**(1): p. 1-14.

28. Elms, T.N., P.H. Buschang, and R.G. Alexander, *Long-term stability of Class II, Division 1, nonextraction cervical face-bow therapy: I. Model analysis.* American Journal of Orthodontics and Dentofacial Orthopedics, 1996. **109**(3): p. 271-276.

29. Ciger, S., M. Aksu, and D. Germeç, *Evaluation of posttreatment changes in Class II Division 1 patients after nonextraction orthodontic treatment: Cephalometric and model analysis.* American Journal of Orthodontics and Dentofacial Orthopedics, 2005. **127**(2): p. 219-223.

30. Bock, N.C., et al., *Long-term (≥15 years) post-treatment changes and outcome quality after Class II:1 treatment in comparison to untreated Class I controls.* European Journal of Orthodontics, 2018. **40**(2): p. 206-213.

31. Bock, N.C., J. Von Bremen, and S. Ruf, *Stability of Class II fixed functional appliance therapy—a systematic review and meta-analysis.* The European Journal of Orthodontics, 2016. **38**(2): p. 129-139.

32. Millett, D.T., et al., *Treatment and stability of Class II Division 2 malocclusion in children and adolescents: A systematic review.* American Journal of Orthodontics and Dentofacial Orthopedics, 2012. **142**(2): p. 159-169.e9.

33. Canut, J., *A long-term evaluation of treated Class II division 2 malocclusions: a retrospective study model analysis.* The European Journal of Orthodontics, 1999. **21**(4): p. 377-386.

34. Bishara, S.E., D.M. Cummins, and A.R. Zaher, *Treatment and posttreatment changes in patients with Class II, Division 1 malocclusion after extraction and nonextraction treatment.* American Journal of Orthodontics and Dentofacial Orthopedics, 1997. **111**(1): p. 18-27.

35. Freitas, K.M.S., et al., *Postretention relapse of mandibular anterior crowding in patients treated without mandibular premolar extraction.* American Journal of Orthodontics and Dentofacial Orthopedics, 2004. **125**(4): p. 480-487.

36. Swidi, A.J., A.E. Griffin, and P.H. Buschang, *Mandibular alignment changes after full-fixed orthodontic treatment: a systematic review and meta-analysis.* Eur J Orthod, 2019. **41**(6): p. 609-621.

37. Artun, J., J.D. Garol, and R.M. Little, *Long-term stability of mandibular incisors following successful treatment of Class II, Division 1, malocclusions.* Angle Orthod, 1996. **66**(3): p. 229-38.

38. Rossouw, P.E., C.B. Preston, and C. Lombard, *A longitudinal evaluation of extraction versus nonextraction treatment with special reference to the posttreatment irregularity of the lower incisors.* Semin Orthod, 1999. **5**(3): p. 160-70.

39. Little, R.M., T.R. Wallen, and R.A. Riedel, *Stability and relapse of mandibular anterior alignment-first premolar extraction cases treated by traditional edgewise orthodontics.* Am J Orthod, 1981. **80**(4): p. 349-65.

40. Bolton, W.A., *The clinical application of a tooth-size analysis.* American Journal of Orthodontics, 1962. **48**(7): p. 504-529.

41. Graber, T.M., *Clinical Cephalometrics*. Vol. 38. 1952: Am. Journal of Orthodontics.

42. Loughlin, F.J., *Aims, aids, and auxiliaries.* American Journal of Orthodontics, 1952. **38**(5): p. 315-334.

43. Turbill, E.A., S. Richmond, and J.L. Wright, *The time-factor in orthodontics: what influences the duration of treatments in National Health Service practices?* Community Dent Oral Epidemiol, 2001. **29**(1): p. 62-72.

44. O'Brien, K.D., et al., *The effectiveness of Class II, division 1 treatment.* Am J Orthod Dentofacial Orthop, 1995. **107**(3): p. 329-34.

45. Janson, G., et al., *Orthodontic treatment time in 2- and 4-premolar-extraction protocols.* American Journal of Orthodontics and Dentofacial Orthopedics, 2006. **129**(5): p. 666-671.

46. Janson, G., et al., *Class II treatment efficiency in maxillary premolar extraction and nonextraction protocols.* American Journal of Orthodontics and Dentofacial Orthopedics, 2007. **132**(4): p. 490-498.

47. Janson, G., et al., *Class II treatment success rate in 2- and 4-premolar extraction protocols.* American Journal of Orthodontics and Dentofacial Orthopedics, 2004. **125**(4): p. 472-479.

48. Akinci Cansunar, H. and T. Uysal, *Comparison of orthodontic treatment outcomes in nonextraction, 2 maxillary premolar extraction, and 4 premolar extraction protocols with the American Board of Orthodontics objective grading system.* American Journal of Orthodontics and Dentofacial Orthopedics, 2014. **145**(5): p. 595-602.

49. Janson, G., et al., *Treatment stability in patients with Class II malocclusion treated with 2 maxillary premolar extractions or without extractions.* American Journal of Orthodontics and Dentofacial Orthopedics, 2010. **138**(1): p. 16-22.

50. Little, R.M., *The Irregularity Index: A quantitative score of mandibular anterior alignment.* American Journal of Orthodontics, 1975. **68**(5): p. 554-563.

51. Casko, J.S., et al., *Objective grading system for dental casts and panoramic radiographs.* American Journal of Orthodontics and Dentofacial Orthopedics, 1998. **114**(5): p. 589-599.

52. Peter H. Buschang, J.D.S., *Incisor crowding in untreated persons 15-50 years of age: United States, 1988-1994.* The Angle Orthodontist, 2003. **73**(5): p. 502-508.

53. Grainger, R.M., *Orthodontic treatment priority index.* Vital and Health Statistics, 1967.

54. Nett, B.C. and G.J. Huang, *Long-term posttreatment changes measured by the American Board of Orthodontics objective grading system.* American Journal of Orthodontics and Dentofacial Orthopedics, 2005. **127**(4): p. 444-450.

55. Kashner, J., *Does higher quality of treatment translate into a higher quality long-term result.* Diss. University of Washington, 1999.

56. Van Der Linden, F.P.G.M., *Facial growth and facial orthopedics*, ed. Chicago. 1986: Quitessence Publishing.

57. *Why case reports do not pass the ABO Phase III clinical examination.* American Journal of Orthodontics and Dentofacial Orthopedics, 1996. **110**(5): p. 559-560.

58. Fidler, B.C., et al., *Long-term stability of Angle Class II, Division 1 malocclusions with successful occlusal results at end of active treatment.* American Journal of Orthodontics and Dentofacial Orthopedics, 1995. **107**(3): p. 276-285.

APPENDIX A FIGURES

Figure 1: American Board of Orthodontics Objective Grading System Ruler

Figure 2: Mean ± 1 standard error Frankfort-mandibular plane angle (FMA) and ANB angle at pretreatment (T1)

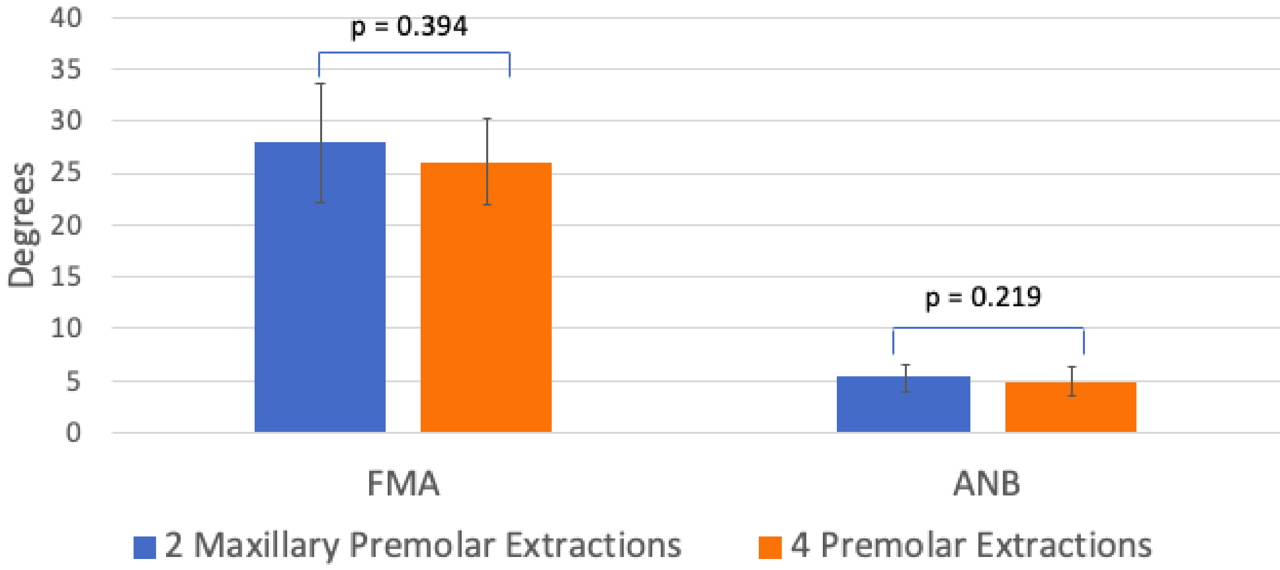


Figure 3: Pretreatment medians and interquartile ranges for the overall objective grading system (OGS) scores

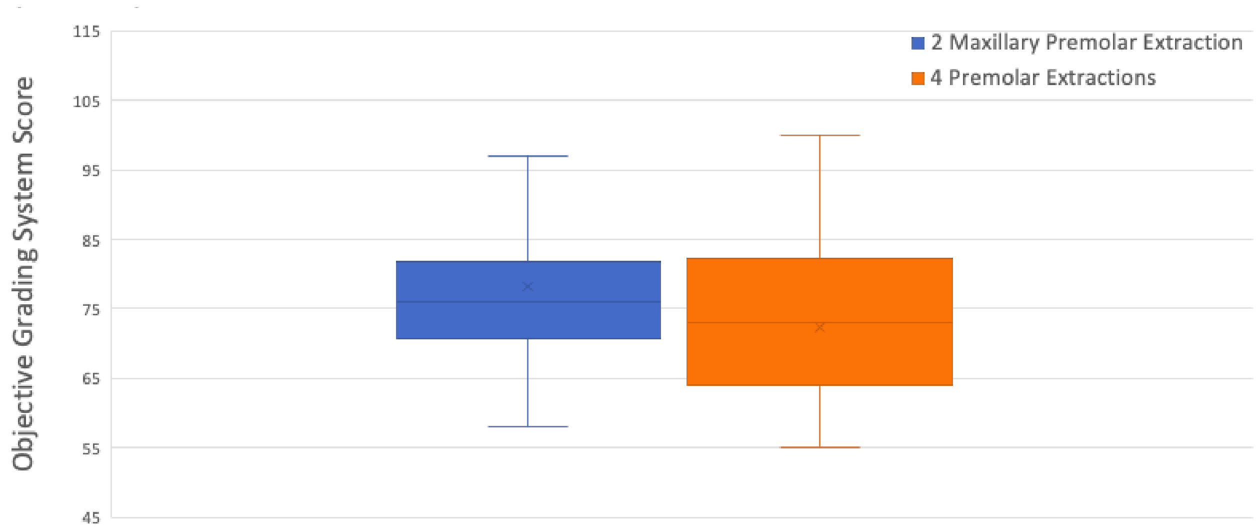


Figure 4a: Median canine relationships at pretreatment (T1), posttreatment (T2), and post-retention (T3) for the 2 maxillary premolar and 4 premolar extraction groups

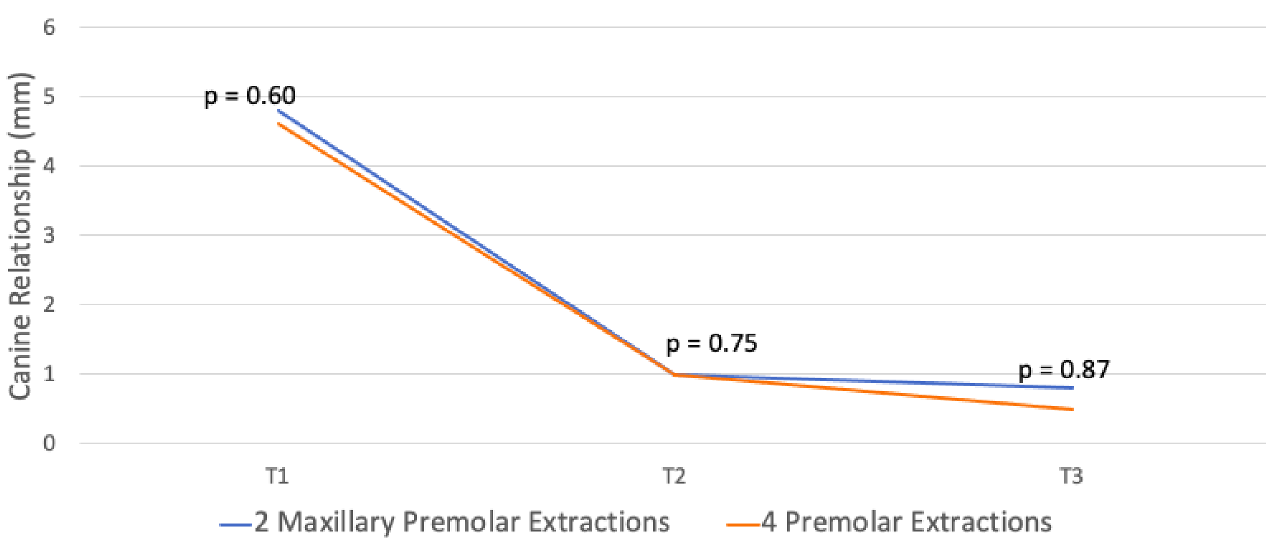


Figure 4b: Median molar relationships at pretreatment (T1), posttreatment (T2), and post-retention (T3) for the 2 maxillary premolar and 4 premolar extraction groups

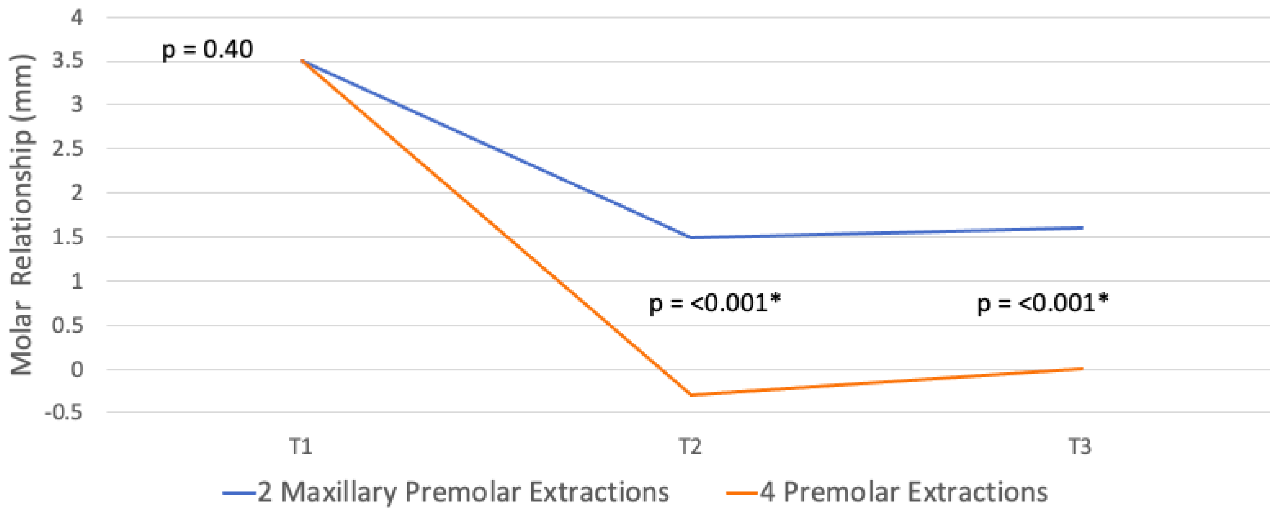


Figure 5. Median changes in overall objective grading system (OGS) score of the 2 maxillary premolar and 4 premolar extraction groups

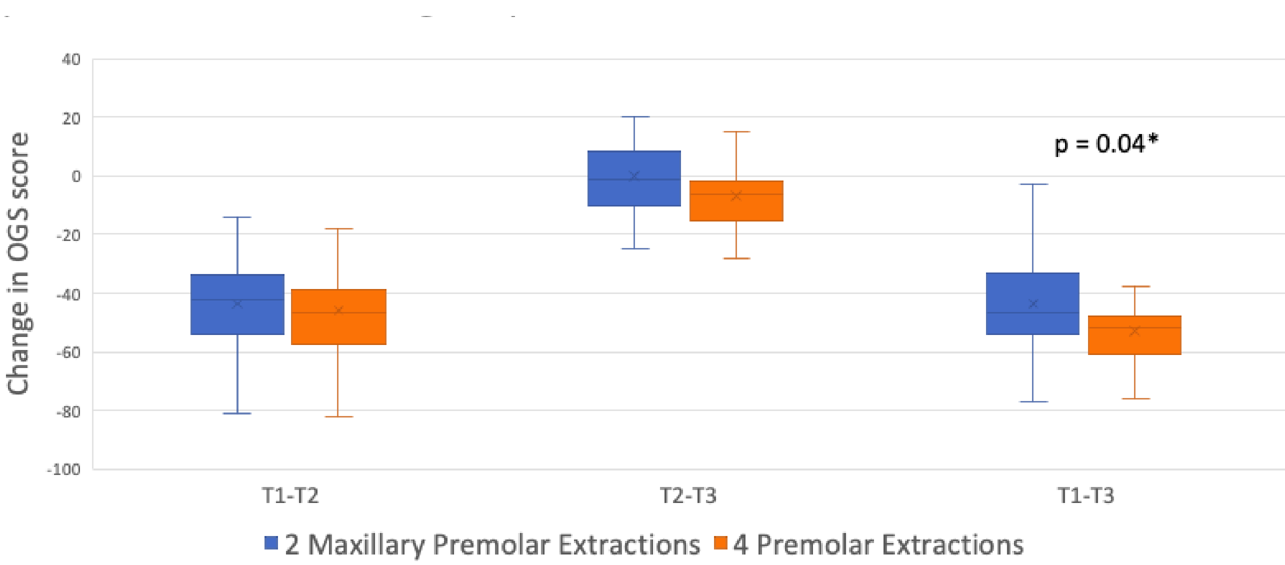


Figure 6. Median overall objective grading system (OGS) scores at pretreatment (T1), posttreatment (T2), and post-retention (T3)

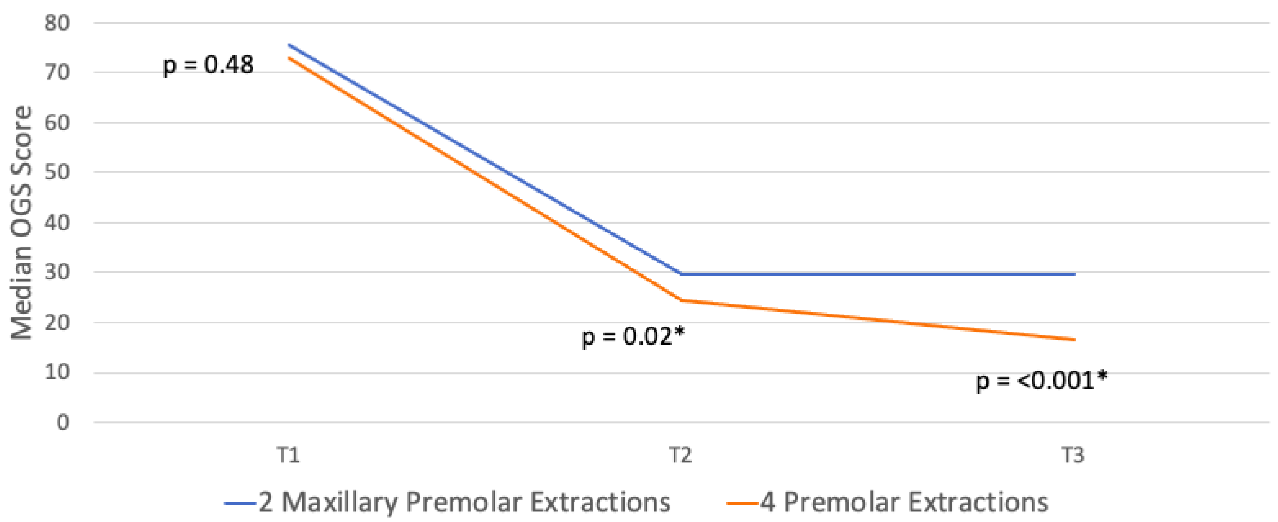


Figure 7. Median objective grading system component scores posttreatment (T2)

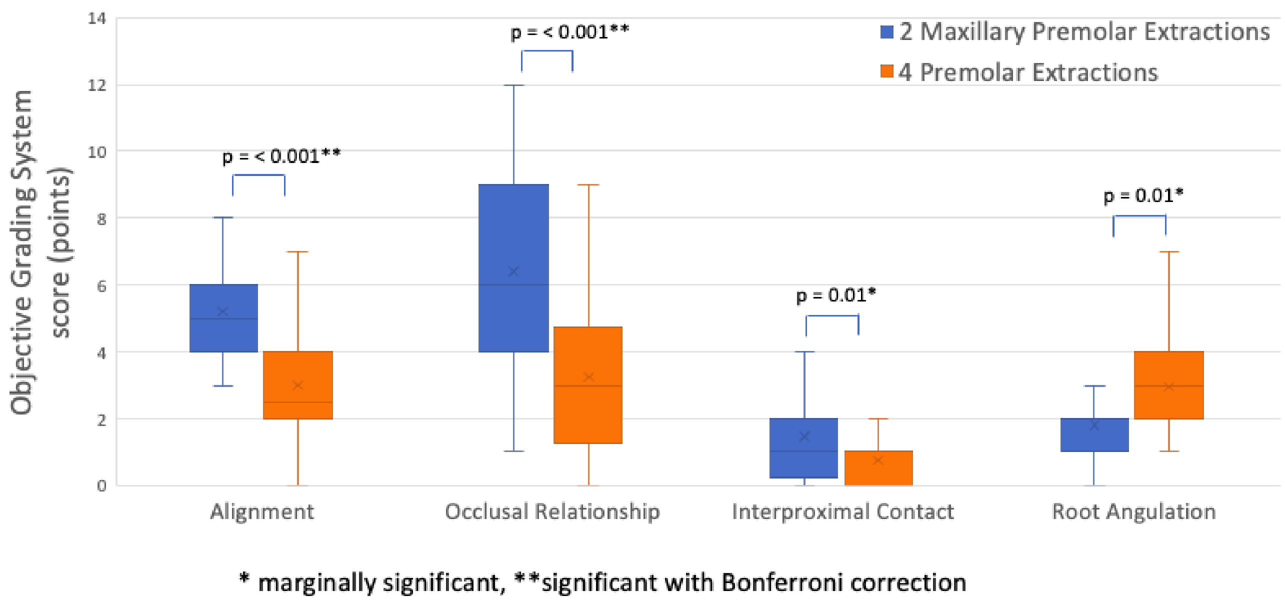
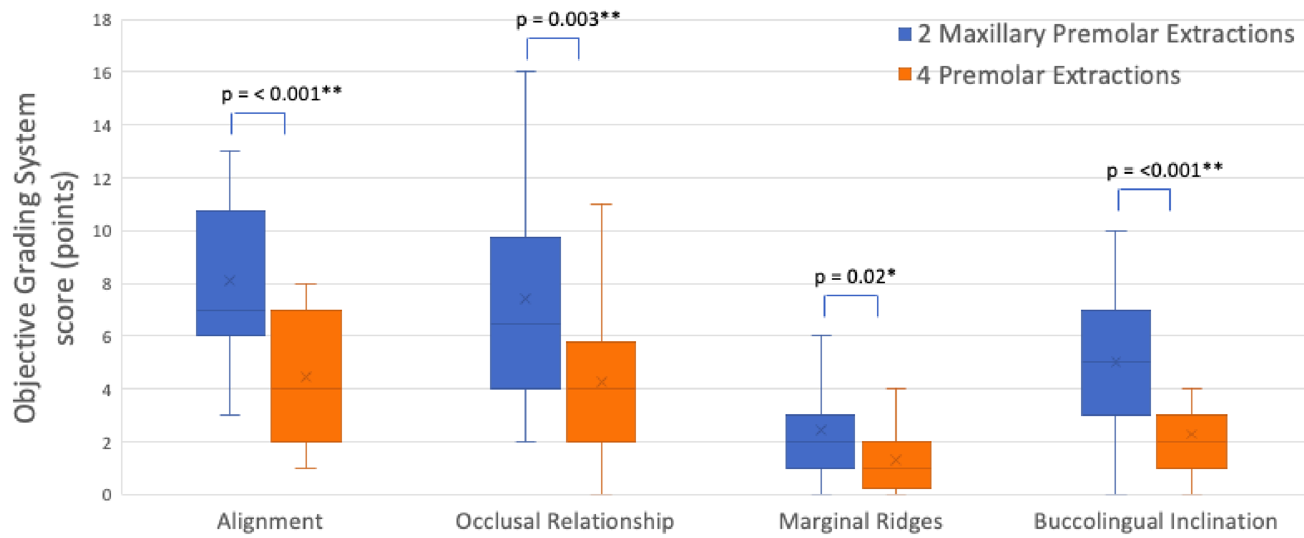


Figure 8. Median objective grading system component scores post-retention (T3)



* marginally significant, **significant with Bonferroni correction

APPENDIX B TABLES

Table 1. Sex distributions for the 2 maxillary premolar extraction and 4 premolar extraction groups

	Females	**Males**	**Total**	**Chi-Square**
2 Maxillary Premolar Extractions	21	3	24	$\chi^2 = 0.17$ p = 0.68
4 Premolar Extractions	20	4	24	
Total	41	7	48	

Table 2. Group practice distributions for the 2 maxillary premolar extraction and 4 premolar extraction groups

	Practice 1	**Practice 2**	**Practice 3**	**Practice 4**	**Chi-Square**
2 Maxillary Premolar Extractions	15	4	4	1	$\chi^2 = 1.15$ p = 0.77
4 Premolar Extractions	12	7	4	1	

Table 3. Median ages and interquartile ranges at each timepoint for the 2 maxillary premolar extraction and 4 premolar extraction groups

	2 Maxillary Premolar Extractions		**4 Premolar Extractions**		**Mann Whitney P-value**
	Median (50%)	***IQR (25%,75%)***	***Median (50%)***	***IQR (25%,75%)***	
Age @ T1	14.3	12.0, 20.7	12.7	12.2, 17.0	p= 0.35
Age @ T2	16.6	15.5, 23.0	15.5	14.5, 20.2	p = 0.13
Age @ T3	33.1	26.5, 40.2	34.8	29.2, 40.5	p = 0.68

Table 4. Median changes in molar and canine relationship during treatment (T1-T2) and posttreatment (T2-T3)

		2 Maxillary Premolar Extractions		4 Premolar Extractions		Mann-Whitney P-value
		Median (50%)	*IQR (25%,75%)*	*Median (50%)*	*IQR (25%,75%)*	
Change in Canine Relation	***T1-T2***	-3.8	-4.9, -3.3	-3.6	-4.9, -2.3	0.54
	T2-T3	0	-0.9, 0.5	0	-0.7, 0.8	0.65
	T1-T3	-4.0	-5.4, -3.5	-3.4	-4.8, -2.5	0.19
Change in Molar Relation	***T1-T2***	-2.5	-3.6, -1.6	-3.8	-4.6, -3.0	p = 0.01*
	T2-T3	0	-0.6, 0.7	0	-0.3, 0.9	p = 0.63
	T1-T3	-2.6	-3.5, -1.5	-3.5	-4.2, -2.6	p = 0.06

Table 5. Median molar relationships, canine relationships, and incisor irregularity at pretreatment (T1), posttreatment (T2), and post-retention (T3)

		2 Maxillary Premolar Extractions		4 Premolar Extractions		Mann-Whitney P-value
		Median (50%)	*IQR (25%,75%)*	*Median (50%)*	*IQR (25%,75%)*	
Canine relationship	***T1***	4.8	4.0, 5.8	4.6	3.8, 5.5	p = 0.60
	T2	1.0	0, 1.9	1.0	0, 1.4	p = 0.75
	T3	0.8	0, 1.5	0.5	0, 1.8	p = 0.87
Molar relationship	***T1***	3.5	2.8, 5.6	3.5	2.3, 4.4	p= 0.40
	T2	1.5	1.0, 1.9	-0.3	-1.0, 0	p = <0.001*
	T3	1.6	0.8, 2.0	0	-1.0, 0.5	p = <0.001*
Mandibular incisor irregularity	***T1***	3.5	2.5, 6.1	5.0	2.6, 6.4	p = 0.20
	T2	0	0, 0.5	0	0, 0.5	p = 0.90
	T3	2.3	0.5, 3.4	1.5	0.5, 2.4	p = 0.21

Table 6. Median overall changes in objective grading system (OGS) scores during treatment (T1-T2), posttreatment (T2-T3), and from pretreatment to post-retention (T1-T3)

		2 Maxillary Premolar Extractions		4 Premolar Extractions		Mann-Whitney P-value
		Median (50%)	***IQR (25%,75%)***	***Median (50%)***	***IQR (25%,75%)***	
Change in Total OGS Score	***T1-T2***	**-42.0**	-54.0, -33.8	**-46.5**	-57.5, -39.0	p = 0.32
	T2-T3	-1.0	-10.0, 8.5	**-6.5**	-15.0, -2.0	p = 0.08
	T1-T3	**-46.5**	-54.0, -33.0	**-52.0**	-60.8, -48.0	p = 0.04*

bold medians = p statistically significant (p<0.05) within group differences

Table 7. Median overall objective grading system (OGS) scores at pretreatment (T1), posttreatment (T2), and post-retention (T3)

		2 Maxillary Premolar Extractions		4 Premolar Extractions		Mann-Whitney P-value
		Median (50%)	***IQR (25%,75%)***	***Median (50%)***	***IQR (25%,75%)***	
Total OGS Score	***T1***	75.5	66.3, 80.8	73.0	62.0, 81.3	p = 0.48
	T2	29.5	26.5, 38.3	24.5	15.8, 35.3	p = 0.02*
	T3	29.5	24.3, 37.0	16.5	13.0, 21.8	p = <0.001*

Table 8. Median objective grading system component pretreatment (T1) scores

	2 Maxillary Premolar Extractions		4 Premolar Extractions		Mann-Whitney P-value
	Median (50%)	*IQR (25%,75%)*	*Median (50%)*	*IQR (25%,75%)*	
Alignment	15.5	13.0,18.8	16.5	14.0, 19.8	p= 0.69
Marginal Ridges	8.0	5.0, 10.0	8.0	4.0, 9.8	p = 0.36
Buccolingual Inclination	6.5	4.0, 8.0	4.0	3.0, 6.0	p = 0.02*
Occlusal Contacts	6.0	2.0, 11.0	6.5	5.0, 11.3	p = 0.85
Occlusal Relation	18.5	15.0, 20.0	18.0	16.0, 20.0	p = 0.89
Overjet	12.0	10.0, 14.0	12.5	12.0, 15.0	p = 0.33
Interproximal Contact	3.0	2.0, 8.8	0.5	0.0, 6.8	p = 0.05
Root Angulation	3.5	3.0, 4.8	3.5	3.0, 5.0	p = 0.87

Table 9. Median objective grading system component posttreatment (T2) scores

	2 Maxillary Premolar Extractions		4 Premolar Extractions		Mann-Whitney P-value
	Median (50%)	*IQR (25%,75%)*	*Median (50%)*	*IQR (25%,75%)*	
Alignment	5.0	4.0, 6.0	2.5	2.0, 4.0	p = <0.001*
Marginal Ridges	5.0	3.0, 6.0	3.0	2.0, 5.0	p = 0.07
Buccolingual Inclination	3.0	2.3, 5.0	3.0	2.0, 6.0	p = 0.63
Occlusal Contacts	5.0	2.0, 7.8	4.5	2.3, 7.8	p = 0.92
Occlusal Relation	6.0	4.0, 9.0	3.0	1.3, 4.8	p = <0.001*
Overjet	3.0	2.0, 4.0	3.0	1.3, 5.8	p = 0.93
Interproximal Contact	1.0	0.25, 2.0	0.0	0.0, 1.0	p = 0.01
Root Angulation	2.0	1.0, 2.0	3.0	2.0, 4.0	p = 0.01

Table 10. Median objective grading system component post-retention (T3) scores

	2 Maxillary Premolar Extractions		4 Premolar Extractions		Mann-Whitney P-value
	Median (50%)	***IQR (25%,75%)***	***Median (50%)***	***IQR (25%,75%)***	
Alignment	7.0	6.0, 10.8	4.0	2.0, 7.0	p = <0.001*
Marginal Ridges	2.0	1.0, 3.0	1.0	0.3, 2.0	p = 0.02
Buccolingual Inclination	5.0	3.0, 7.0	2.0	1.0, 3.0	p = <0.001*
Occlusal Contacts	1.5	0, 3.0	1.5	0.0, 3.0	p = 0.96
Occlusal Relation	6.5	4.0, 9.8	4.0	2.0, 5.8	p = 0.003*
Overjet	3.0	1.3, 7.5	1.0	1.0, 4.0	p = 0.07
Interproximal Contact	0	0, 1.0	0.0	0.0, 0.0	p = 0.24
Root Angulation	1.5	1.0, 2.0	1.0	1.0, 1.0	p = 0.18

Table 11. Median changes (T1-T2) in objective grading system component treatment scores

	2 Maxillary Premolar Extractions		4 Premolar Extractions		Mann-Whitney P-value
	Median (50%)	***IQR (25%,75%)***	***Median (50%)***	***IQR (25%,75%)***	
Alignment	**-9.5**	-14.0, -8.0	**-13.5**	-17.0, -9.3	p = 0.08
Marginal Ridges	**-3.0**	-5.0, -1.3	**-3.5**	-5.8, -0.3	p = 0.93
Buccolingual Inclination	**-2.5**	-5.0, -1.0	-1.0	-3.0, 2.0	p = 0.09
Occlusal Contacts	-1.0	-7.5, 1.8	**-3.0**	-7.0, 0.8	p = 0.77
Occlusal Relation	**-10.0**	-14.8, -7.3	**-14.0**	-17.0, -10.3	p = 0.05
Overjet	**-8.0**	-11.8, -5.3	**-8.5**	-12.0, -6.3	p = 0.51
Interproximal Contact	**-2.5**	-6.8, 0	**-0.5**	-6.8, 0.8	p = 0.31
Root Angulation	**-2.0**	-3.0, -1.0	0	-1.0, 0	p = 0.10

bold medians = p statistically significant (p<0.05) within group differences

Table 12. Median changes (T2-T3) in objective grading system component posttreatment scores

	2 Maxillary Premolar Extractions		4 Premolar Extractions		Mann-Whitney P-value
	Median (50%)	*IQR (25%,75%)*	*Median (50%)*	*IQR (25%,75%)*	
Alignment	**3.5**	1.0, 5.0	**1.0**	0, 3.0	p = 0.07
Marginal Ridges	**-2.5**	-4.0, 0	**-2.0**	-3.0, -1.0	p = 0.93
Buccolingual Inclination	0	-0.8, 2.0	**0**	-3.0, 1.0	p = 0.03
Occlusal Contacts	**-2.0**	-5.0, 0	**-3.0**	-5.0, -0.3	p = 0.48
Occlusal Relation	1.0	-1.8, 2.8	1.0	-0.8, 3.0	p = 0.85
Overjet	0	-2.8, 3.8	-2.0	-3.0, 0.8	p = 0.18
Interproximal Contact	**0**	-2.0, 0	0	-1.0, 0	p = 0.21
Root Angulation	**0**	-1.0, 0	**-1.0**	-2.0, -1.0	p = <0.001*

bold medians = p statistically significant (p<0.05) within group differences

Table 13. Median changes (T1-T3) in objective grading system component pretreatment to post-retention scores

	2 Maxillary Premolar Extractions		4 Premolar Extractions		Mann-Whitney P-value
	Median (50%)	*IQR (25%,75%)*	*Median (50%)*	*IQR (25%,75%)*	
Alignment	**-8.5**	-10.0, -5.3	**-13.0**	-16.0, -9.0	p = 0.004*
Marginal Ridges	**-5.0**	-7.0, -3.3	**-6.0**	-8.0, -3.0	p = 0.53
Buccolingual Inclination	**-2.0**	-3.0, 0	**-2.0**	-4.0, -0.3	p = 0.65
Occlusal Contacts	**-4.0**	-10.0, 0	**-5.5**	-8.0, -1.3	p = 0.69
Occlusal Relation	**-11.5**	-13.0, -6.3	**-13.5**	-16.0, -10.0	p = 0.09
Overjet	**-8.0**	-10.0, -4.0	**-9.5**	-12.0, -7.3	p = 0.05
Interproximal Contact	**-2.5**	-7.8, -1.3	**-0.5**	-6.8, 0	p = 0.10
Root Angulation	**-2.0**	-4.0, -1.0	**-2.0**	-4.0, -2.0	p = 0.39

bold medians = p statistically significant (p<0.05) within group differences