Spring 2017

Comparison of surgical and non-surgical orthodontic treatment approaches on occlusal and cephalometric outcomes in patients with severe Class II division I malocclusions

Sheila Meghnot Daniels
University of Iowa

COMPARISON OF SURGICAL AND NON-SURGICAL ORTHODONTIC TREATMENT APPROACHES ON OCCLUSAL AND CEPHALOMETRIC OUTCOMES IN PATIENTS WITH SEVERE CLASS II DIVISION I MALOCCLUSIONS

by

Sheila Meghnot Daniels

A thesis submitted in partial fulfillment
of the requirements for the Master of
Science degree in Orthodontics in
the Graduate College of
The University of Iowa

May 2017

Thesis Supervisor: Associate Professor Veerasathpurush Allareddy

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

____

MASTER'S THESIS

____

This is to certify that the Master's thesis of

Sheila Meghnot Daniels

has been approved by the Examining Committee for the thesis requirement for the Master of Science degree in Orthodontics at the May 2017 graduation.

Thesis Committee: ___________________________________
Veerasathpurush Allareddy, Thesis Supervisor

____
Kyungsup Shin

____
Satheesh Elangovan

ACKNOWLEDGMENTS

When I was three years old, my dream was to become an orthodontist. To say this degree came solely of my own accord would simply be untrue. I'd like to take a moment to thank all the wonderful people who have been a part of this journey:

For my mom, for her sacrifices and endless support. My siblings, Layla and Arya – my partners in crime and now my confidants. Brother, that statistics degree really came in handy. For my family in Michigan who've cheered me on as I've travelled the country on this mission- and at times, travelled for me.

For my wonderful friends - you know who you are. I can't forget the unfailing support of my good friend, Carrie, who gave me the best news of my life when she said she was moving to Richmond during my second year of dental school.
A special thanks to Kristen. Without our productivity reports, this would have taken much longer to complete.

Dr. Sath Allareddy, my thesis supervisor, whose speech two years ago is what made me want to be a Hawkeye. It has been a privilege and an honor. You're "the best."

Thanks to my thesis committee for their time and feedback. My mentors: Drs. Thomas Southard, Steven Marshall, Kyungsup Shin, Clayton Parks, Michael Callan, and Thomas Stark – your compassion and dedication to this specialty is exhilarating.

And for Tim. Since having met you, I have become the best version of myself.

ABSTRACT

Introduction: The Class II malocclusion is most oft treated by orthodontic practitioners in the United States. Multiple studies have aimed to identify a standardized way to treat these cases. However, due to a variety of factors (bias, inadequate sample size, patient burnout), this has proven quite challenging. This study follows subjects over the course of treatment to determine if any information can be obtained which will help future doctors treat these cases.

Purpose: This study aimed to examine end-of-treatment outcomes of severe Class II Division I malocclusion patients treated with surgical or non-surgical approaches. This study tests the hypotheses that occlusal outcomes (ABO-OGS) at the end of treatment will be similar while cephalometric outcomes will differ between these groups.

Study Design: 60 patients were identified which fit the inclusion criteria (40 non-surgical, 20 surgical). Initial and final casts were graded using parameters outlined by the American Board of Orthodontics. Initial information on each subject was gathered. End of treatment outcomes were compared using Mann-Whitney U tests and multivariable linear regression models.

Results: Adjustments were made for multiple confounders (age, gender, complexity of case, and skeletal patterns). The final deband was found to be similar in both groups (23.8 for surgical group versus 22.5 for non-surgical group). Those treated surgically had a significantly larger reduction in ANB angle, 3.4 degrees reduction versus 1.5 degrees reduction in the non-surgical group ($p=0.002$). The surgical group also showed increased maxillary incisor proclination ($p=0.001$) compared to candidates treated non-surgically.

Conclusion: Orthodontic cast outcomes were similar in both surgical and non-surgical treatment groups, where cephalometric outcomes differed. Data showed differences in treatment design impacted both final tooth position and angulation.

PUBLIC ABSTRACT

This study aimed to examine end-of-treatment outcomes of severe Class II Division I malocclusion patients treated with surgical or non-surgical approaches. This study tests the hypotheses that occlusal outcomes (ABO-OGS) at end of treatment will be similar while cephalometric outcomes will differ between these groups. A total of 60 patients were included: 20 of which underwent surgical correction and 40 of which did not. The end of treatment ABO-OGS and cephalometric outcomes were compared by Mann-Whitney U tests and multivariable linear regression models. Following adjustment for multiple confounders (age, gender, complexity of case, and skeletal patterns), the final deband score (ABO-OGS) was similar for both groups (23.8 for surgical group versus 22.5 for non-surgical group). Those treated surgically had a significantly larger reduction in ANB angle (p=0.002) and increased maxillary incisor proclination (p=0.001) compared to those treated non-surgically.

<u>TABLE OF CONTENTS</u>

LIST OF TABLES

Table

LIST OF FIGURES

Figure

INTRODUCTION

Three types of occlusion exist in the orthodontic population with each exhibiting distinguishable traits which differentiate them from each other (Class I, II, and III). This research focuses on the Class II division which clinically presents as a patient with a normally positioned maxilla and a relatively retrusive mandible. Dentally, Class II malocclusions are divided into two subcategories: Division I (Div I: characterized by increased overjet and a retrognathic mandible) and Division II (in which maxillary lateral incisors or canines are more proclined than the central incisors which are retroclined, typically). The more common of the Class II malocclusions is the Division I subset. Studies have found that an overjet of 5 mm or more is characteristic in over a quarter of the child/adolescent population and a little less than 10% of the adult population (Proffit et al., 2007). Left untreated, Class II malocclusions can pose a variety of complications both present and future. These include those in the functional, psychological, and sociological realms. A recent Cochrane review found that in early treatment groups, Class II patients with increased overjet are twice as much at risk of incisor trauma as those without increased overjet (Thiruvenkatachari et al., 2013; Thiruvenkatachari et al., 2015). QoL (quality of life) indicators agree that a child's perception of life might be negatively affected due to a Class II relationship (Martins-Junior et al., 2012). Along with this, sociological factors must be considered such as teasing or bullying. In fact, a study found that Class II Division I patients are more likely to deal with these issues when compared to children without this malocclusion (Seehara et al., 2011). Therefore, we can appreciate how orthodontic interference can help patients who present Class II in a number of ways.

Treatment options for Class II Division I malocclusions are three-pronged: orthopedics (headgear/Herbst appliance), masking/camouflage (where extractions or compensations may be utilized to provide a harmonious occlusal scheme without addressing an underlying skeletal discrepancy), and orthognathic surgery (either of the maxilla, mandible, or both. Literature and clinical practice has shown each option to be an effective means of treatment (Janson et al., 2013; Marsico et al., 2011; Perillo et al., 2010; Burden et al., 2007; Kinzinger et al., 2005; Lohrmann et al., 2006; Pancherz et al., 1994; Baccetti et al., 2009; Sloss et al., 2008; Cassidy et al., 1993; Cacciatore et al., 2014; Cacciatore et al., 2014; Wigal et al., 2011; Berger et al., 2005). The decision as to which path to take depends on two factors: time (as in, age of patient) and magnitude (amount of discrepancy: mild, moderate, or severe) (Janson et al., 2010). A significant skeletal component is expected and usually present in severe Class II Division I malocclusions. In these cases, the ideal method of treatment is orthognathic surgery – primarily a BSSO advancement – since this is the only treatment which addresses the skeletal base discrepancy. However, patients are not always accepting of this procedure. In these cases, one of the other modes of treatment (orthopedic attempt or camouflage/masking) may be attempted in lieu of the surgery. The orthopedic option, if the patient is young, can achieve better facial harmony. Yet, while a masking treatment can address occlusal discrepancies, it will not improve skeletal position and therefore profile esthetics (Brady, 2016, Wigal et al., 2011; Berger et al., 2005; Mihalik et al., 2003; Kinzinger et al., 2009; Tucker et al., 1995; Proffit et al., 1992; de Lir Ade et al., 2013; Ko et al., 2011; Millett et al., 2012; Chaiyongsirisern et al., 2009)

Especially in the United States, Class II malocclusions are possibly the most commonly encountered by practitioners in private practice and in residency. However, there is still no standard method of treatment or agreement on best practice modality. This could be because treatment of this malocclusion is multi-factorial, depending on age, timing of treatment, and patient concerns and desires. Current literature in this area is lacking in non-bias or low-bias articles with high levels of evidence Therefore, it is prudent for current researchers to work to provide well-founded additions to the pool (Brady 2016, Thiruvenkatachari et al., 2013; Chaiyongsirisern et al., 2009; Jambi et al., 2013; Millett et al., 2006; Koletsi et al., 2014).

This research will focus on cast-based treatment of Class II Div I malocclusions, both surgically and non-surgically. To date, quite a few articles have analyzed how each of the three treatment options (orthopedics, masking, and surgery) compare post-treatment. Initial discrepancy index measurements will be performed for each case as well as post-treatment ABO cast grading. Cephalometric outcomes will also be reviewed. Only one research project compared outcomes of surgical vs. non-surgical treatment of patients in the mid- to late adolescents (Lembesi et al., 2014). This is an important age to assess treatment outcomes, because it is one of the most common ages for initiation of orthodontic treatment. As such, proper diagnosis and treatment planning is imperative because, if an irreversible option is entertained and carried to completion, the practitioner might condemn the patient from being able to choose a more ideal option later in life.

REVIEW OF THE LITERATURE

Psychologically speaking, studies have looked into the effects of different malocclusions on a child's personal life. One such study specifically looked at bullying, which was defined as a more severe type of stress; one which causes harm and distress to a child. This was a cross-sectional study of patients in the adolescent population at three hospitals in the United Kingdom. 336 patients aged 10-14 were asked to participate and a questionnaire was used to measure the amount and severity of bullying they experienced, their self-esteem, and their quality of life index. Orthodontic treatment needs were also identified for each case. Results showed that about 13% of patients experienced bullying and that bullying was significantly related to Class II Division I incisor inclination, increased overbite, and a higher rating for the need of orthodontic intervention. It also found that participants who were bullied ranked with lower self-confidence, athletic competence, and social adaptability compared to a control group. Due to negative scores in a variety of lifestyle areas, it was concluded that they also had an overall negative quality of life score (Seehara et al., 2011).

An even younger group of children, aged 8-10 years, was studied to determine if differing malocclusions affected their quality of life. The Dental Aesthetic Index (DAI) was used to determine severity of the cases and the Child Perceptions Questionnaire (CPQ) was used to identify the quality of life. This study found that there was a positive correlation between CPQ and DAI scores with upper anterior irregularity greater or equal to 2mm, anterior open bite greater or equal to 2mm, and a diastema greater or equal to 2mm (Martins-Junior et al., 2012).

Statistical analysis plays a key role in many orthodontic pieces of literature. Many articles report findings taken from randomized controlled trials, (RCTs), and there exists a few publications which discuss data gathered from such reports. Koletsi, Pandis, and Fleming sought to assess the quality of reports of sample size calculations in trials published as RCTs in eight leading orthodontic journals during a twenty-year period up to September 2012. They wanted to identify factors associated with correct performance of sample size calculations in these journals, to determine the number of participants typically recruited to clinical trials in orthodontics, and also to assess the accuracy of calculations. A standard data collection form was completed for 10 selected papers. Details of the *a priori* sample size calculation were recorded – in particular, the conduct of a sample size calculation. If conducted, the target sample size, number of participants recruited, number of participants lost to follow-up, type of analysis (intention-to-treat [ITT] or per-protocol basis), and details of the power were recorded. Of 139 RCTs identified, complete sample size calculations were reported in 41 studies (29.5%). Parallel designs were typically used (113 studies; 81%), with 80% (111 studies) involving two arms and 16% (22 studies) having three arms. Data analysis was conducted on an ITT basis in 18 studies (13%). Satisfactory information to allow verification of the sample size calculation was provided in only 41 trials (29.5%) and based on the complete calculations presented in these 41 RCTs, a median of 46 participants were required to demonstrate sufficient power to highlight meaningful. The median number of participants recruited was 60, with a median of 4 participants being lost to follow-up. There was good agreement between projected numbers required and those verified (median discrepancy: 5.3%), although only a minority of trials (29.5%) could be assessed. Overall, most studies (70.6%) failed to present a complete calculation allowing the calculation to be verified, but for the small number of trials (29.5%) where sample recalculation was possible, there was good overall agreement between recruited and required samples. According to the complete sample size calculations acquired, the median number of participants needed in orthodontic research studies is 46, which is usually feasible. The article concluded that although sample size calculations are frequently reported in trials published as RCTs in orthodontic specialty journals, presentation needs to be improved (Millett et al., 2006).

Lembesi investigated the quality of reporting of randomized controlled trials (RCTs) in leading orthodontic journals and explored possible predictors of improved reporting. The contents of the 50 most recent issues of the American Journal of Orthodontics and Dentofacial Orthopedics (AJODO), the Angle Orthodontist (Angle), the European Journal of Orthodontics (EJO) and the Journal of Orthodontics (JO) were electronically searched. The modified CONSORT checklist was used for evaluation of the quality of reporting of orthodontic RCTs. This checklist consisted of 30 questions related to items derived from the CONSORT guidelines. These guidelines were developed in order to standardize reporting of clinical trials in the hopes of allowing better informed healthcare decisions. 128 RCTs were identified with the mean modified CONSORT percentage score across all studies being 68.97%. When the researchers analyzed completeness of reporting, JO ranked first (modified CONSORT score: 76.21%), followed by AJODO (modified CONSORT score: 73.05%). The univariable regression analysis indicated that AJODO, EJO, and JO showed significantly greater modified CONSORT percentage scores compared to Angle. Journal of publication, each additional year of publication, region of authorship, statistical significance for the primary outcome, and methodologist involvement were all significant predictors of improved modified CONSORT scores in the multivariable model. Based on this survey of the four leading orthodontic journals, the methodological and reporting quality of RCTs in orthodontics was found to be “suboptimal” in assorted CONSORT areas. Therefore, enhanced conduct and reporting is needed to aid in making informed orthodontic clinical decisions (Koletsi et al., 2014).

Severity of the initial anteroposterior Class II malocclusion is key in terms of final case result that can be obtained. A study used case selection criteria to confirm the importance given to the classification of Class II malocclusion in the AJODO orthodontic journal. A PubMed search was conducted to find any and all papers referencing Class II malocclusion since the first publication of the American Journal of Orthodontics and Dentofacial Orthopedics (AJODO). Of the 359 papers found, 72 (20.06%) papers reported on the initial severity of the Class II malocclusion sample. Of the remaining articles (the ones which did not report on occlusal severity), researchers determined that a severity recording would have been beneficial to the final results. Therefore, the authors stated that stricter precautions must be taken when determining the quality of results obtained from such cases which lack information on the quantification of occlusal discrepancy in Class II malocclusion cases (Janson et al., 2010).

Past literature has analyzed the effects of functional appliances only. A study done in collaboration between Colorado and Iowa compared soft-tissue profiles after treatment with headgear and Herbst appliances. Both of these appliances are common class II correctors. For the study, 48 pairs of lateral cephs were analyzed pre-treatment and post-treatment to generate silhouettes which were judged by both laypersons and orthodontic residents. The Likert 7 point scale was used for grading. Nonparametric procedures and intraclass correlation were used to compare initial, final, and profile esthetic scores upon change and also correlation between examiners. Participants in both the headgear and Herbst-treated groups showed improvement in profile. There was not a statistically significant difference between the two treatment options. Agreement between evaluations of the layperson group and the orthodontic resident group was strong. Therefore, this study concluded that treatment with either the headgear or Herbst functional appliance results in improved, equally appealing profile changes (Sloss et al., 2008).

Wigal et al. looked at the stability of Class II treatment with the Herbst appliance in early stages of dentition (Cacciatore et al., 2014). Cephalometric radiographs were taken before, after, and after a second phase of appliances. Changes in overjet and mandibular position were recorded, and the study found that overcorrection with a Herbst appliance resulted in less overjet later in life as well as to a more favorable molar position. This study did not, however, include a look at surgical cases.

Berger looked at both Frankel and Herbst appliances placed in young children (mean age: 10 years old). He found that in this group, after the appliance was removed there was continued growth in an advantageous direction. He then compared this group to a group of surgical patients (mean age 27 years old). Both groups had stable results with similar ceph measurements. Therefore, Berger concluded early correction might be a good way to avoid surgery later in life (Berger et al., 2005). This study did not, however, address magnitude nor time (i.e:, was the severity of the apical base discrepancy such that functional appliances alone would not be completely effective in providing correction?).

Tucker reviewed surgical versus camouflage treatment in a 1995 article. In this article, he discussed patient evaluation and motivations for treatment as strong influences on the final decision for treatment, be it surgical or non-surgical. Tucker mentioned that the efficacy of surgical treatment can be determined by observing a few key cephalometric outcomes: overbite, overjet, occlusal scheme, and long term facial improvement. He includes the idea that other factors such as confidence and self-perception might also be good indicators of treatment success (Tucker et al., 1995).

Millett et al. recognized that there exists within orthodontics a variety of treatment methods for these patients – namely a non-extraction treatment or one done in conjunction with extractions. His group was interested in determining for children with a Class II

Division 2 malocclusion whether orthodontic treatment done without the removal of permanent teeth produces a result that is different from no orthodontic treatment at all or orthodontic treatment including extraction of permanent teeth. Randomized controlled trials (RCTs) or controlled clinical trials (CCTs) of orthodontic treatments were identified and included in the study. Active interventions were orthodontic braces (removable, fixed, functional) or headgear with or without extraction of permanent teeth. The control was no treatment or delayed treatment (as in, there was a first phase early in life and a second phase planned for adolescence). However, no RCTs or CCTs meeting the criteria could be identified by the researchers and so, the authors concluded that it “was not possible to provide any evidence-based guidance to recommend or discourage any type of orthodontic treatment to correct Class II division 2 malocclusion in children” (Jambi et al., 2013).

Further studies have delved further to look at functional appliance therapy versus a bilateral sagittal split osteotomy (BSSO) advancement. One study looked explicitly at the Herbst appliance and how correction with this appliance might compare long-term to a surgical advancement. Cephalometric radiographs of 16 patients in each group were evaluated at three time points: start of treatment, deband, and three-year retention. Researchers found that both groups showed significant improvement in overbite, overjet, and mandibular advancement as measured by SNB, SNPg, and Pg/OLp. The surgical group expectedly showed a greater change in these parameters than the non-surgical, expectedly. This study concluded that for borderline Class II malocclusions, Herbst treatment might be beneficial if the surgical option cannot be entertained (Chaiyongsirisern et al., 2009).

Another piece of literature also wanted to assess the success of Herbst appliance usage versus orthognathic surgery. This sample population in this study was adults, making this one of the only pieces of literature that used a functional appliance on an individual much past their growth spurt. Sixty-nine adults with Class II Division I malocclusion were included in this study. Forty-six subjects were treated with an orthodontic-surgical plan (BSSO without genioplasty) and 23 were treated with the Herbst appliance. Changes were observed via cephalometric films taken before and after treatment. This study found that both methods of treatment corrected individual to a Class I occlusal scheme. Surgical candidates were mainly corrected in the sagittal dimension by skeletal changes whereas Herbst-treated patients had mostly dental means of correction. Facial convexity was decreased in both groups with the surgical group having more of a correction than the Herbst group. However, long-term effects were not observed for either group. So, while the authors concluded the Herbst appliances is a valid method of Class II Division I correction, there could be unaccounted for relapse which occurs to a greater extent than in patients who opt for a surgical plan. Another fallback of this study was that it did not provide the exact initial starting Class II relationship (i.e. the degree of Class II character was not quantified) – therefore understanding the severity of initial treatment was not possible (Ruf et al., 2004).

Taking severity into account, there exists a varying spectrum of Class II Division I cases. Therefore, some research has aimed to understand which treatment option, orthodontic or surgical, is best for these "borderline" patients. Twenty-seven orthodontic only and 26 surgical adult cases were chosen upon examining 108 cases for similar initial starting measurements. Researchers thought the fact that some were treated only orthodontically and other surgically was proof that some were treated only orthodontically and others surgically was thought by the researchers proof that treatment decisions for borderline patients was provider-dependent. Candidates in the orthodontically treated group were under a seven year recall while surgical candidates were recalled at about five years. Areas observed at follow-up were skeletal and dental stability, profile, and TMJ function. Orthodontic patients showed greater upper incisor and lip retroclination and decrease in upper arch length. Surgical cases showed an advancement of the mandible results in profile changes. The study found that although profile was markedly different and improved in surgical cases, both groups of subjects were pleased with their final treatment results. During follow-up, dental relapse was minor, but there was surgical relapse (this was due to condylar resorption in a few candidates). The authors state that in the borderline orthodontic/surgical patient, orthodontics alone can provide a satisfactory result unless the magnitude of the condition is so great that only surgical means can correct it (Cassidy et al., 1993).

The choice between each of the three treatment modes involves many factors – a concept investigated by a study headed by Camilla Tulloch. This article emphasized how treatment decisions can be difficult in younger children who are still growing – here, the degree of Class II nature needs to be weighed when planning treatment. Analysis of patient pre-treatment cephalometric radiographs was used to try to predict how age and severity factor into the decision to take one mode of treatment over another. Decrease in overjet was a metric used to determine treatment success. It was found that that while patient age plays a role in planning decisions, it was not related to treatment outcome. In the end, it was decided that neither age nor case severity alone could predict a reduction in overjet, but that multiple factors (psychological and social for instance) might also play a role (Tulloch et al., 1999).

A 1992 article published by Proffit compared outcomes of orthodontic only to surgical-orthodontic Class II malocclusions in adults. Thirty-three adults were treated with the masking approach and 57 were treated surgically. Cephalometric radiographs and casts were reviewed and scored to determine treatment outcomes. To identify effectiveness of treatment, he reviewed final outcomes to see if they fell within the range of normal. Major findings from this article were that both treatments improved dental occlusion and that the facial esthetics of surgical patients improved compared to their initial start point. Orthodontic only treatments had no significant change in profile (Proffit et al., 1992).

Proffit published another article in 1992 – this one looking at the effects and indications for surgical treatment versus orthodontic correction. Three groups of patients were identified: forty patients treated successfully with orthodontics alone, 21 patients treated unsuccessfully with orthodontics only, and 40 patients treated successfully with surgery. Patients started with at least 6mm of overjet, started treatment at twelve years of age or older and were debanded at less than 21 years. Satisfactory treatment was judged to have overjet less than or equal to 4mm, overbite of 0-4mm, and had appealing facial esthetics. In 40% of the patients successfully treated surgically, the surgery was not primarily in the mandible, but also in conjunction with some vertical movement of the maxilla. This study concluded that for Class II patients who are past their pubertal growth spurt, with if the overjet is greater than 10mm, the distance from pogonion to nasion perpendicular is 18mm or more, mandibular body length is less than 70mm, or facial height is more than 125mm, surgery is most likely the best mode of treatment (Lembesi et al., 2014).

Another study in 2010 assessed long-term stability of adolescent versus adult surgery for treatment of deficient mandibles. This article mentioned that in patients with mandibular deficiency, growth is not anticipated after the adolescent growth spurt so mandibular advancement is done after the age of thirteen. It might be judicious to mention that this varies considerably amongst institutions and private practitioners' preferences. Researchers wanted to determine how long-term stability for younger patients compared to that of adults. To do this, surgical patients were brought in for follow-up at one year and five years. Participants were divided into two groups: one group was comprised of 32 adolescents in their late teens and the other, 52 adult patients. One year out from surgery, group one (younger patients), showed more change in the location of B point (horizontal and vertical position), gonion (horizontal position only), and mandibular plane angle. About 20% in both groups had increased overjet of about 2-4mm. Changes in the younger group were expectedly greater than in the adult groups (Proffit et al., 2010).

Still other articles have looked at the long-term effects of surgical treatment of Class II patients. One study comparing the three methods of treatment (orthopedics, camouflage, and surgery) through cephalometric films concurred that overjet was diminished with each treatment type. In the surgical and functional appliance groups, there were changes to profile due to an increase in chin projection and increased body of ramus. Surgical candidates were seen to be the only of the three modalities to have an increase in the vertical dimension (Kinzinger et al., 2009; Millett et al., 2012). Another article aimed to study long-term stability of skeletal, dental, and soft tissue components at one and three years post-treatment. This article also found that of the BSSO advancement cases, relapse was due to increased mandibular plane growth – versus surgeries such as maxillary advancement and maxillary advancement with mandibular setback which were more stable (de Lir Ade et al., 2013).

Perhaps some of the most useful literature regarding Class II malocclusion has been put forth by the Cochrane Collaboration. This international collaboration aims to provide and compile information from various scientific fields and make this available to the research community. Four studies in particular were especially pertinent to this research (Thiruvenkatachari et al., 2015; Chaiyongsirisern et al., 2009; Jambi et al., 2013).

As mentioned, Class II cases can be treated in three ways: orthopedics, masking (involves extraction of teeth), and surgery (can involve extraction of teeth). Millett et al. researched treatment options for young patients with a deep bite and upright incisors (Jambi et al., 2013). This study aimed to determine how non-extraction orthodontic treatment compared to orthodontic treatment with extraction or no treatment in general. Studies included in this literature were randomized control and controlled clinical trials. Unfortunately, no studies fulfilled their criteria and so the authors concluded more research needs to be done in this area as it is clearly lacking.

Jambi et al. further delved into the orthopedic method of correction. With Class II cases, distalizing the upper first molars can improve the Class II molar relationship. Functional appliances and headgear are commonly used as Class II correctional appliances, either as anchorage or to create space for correction. This paper found 10 studies (354 participants) published between 2005 and 2011 in Europe and Brazil. All participants were below sixteen years of age and there was about an equal number of females as males. Researches were interested to see how molar movement differed between intraoral and extraoral appliances (i.e. headgear). They found that, although quality of evidence was not high, the amount of molar distalization with extraoral appliances was higher, but that there was less loss of anterior anchorage with the headgear (Chaiyongsirisern et al., 2009).

In 2015, Thiruvenkatachari et al. revisited their 2008 study regarding early treatment of Class II malocclusions. Their goal this time was to review any more current literature to see if early treatment was indeed superior to only one phase of treatment. A systematic review of databases yielded a number of randomized controlled trials, three of which were used for their research. 353 participants total were included and the samples were comprised of children and adolescents less than sixteen years of age, without cleft lip/palate or any syndrome. They classified the research into 3 groups of studies: Florida, North Carolina, and the United Kingdom. The Florida studies were found to have high bias associated with them: randomization and attrition due to patient drop-out. Of the North Carolina studies there was a high risk for blinding. Only the United Kingdom study had a low risk of bias. When

Thiruvenkatachari compiled data from these studies, they found that when an early treatment with functional appliances was compared to that of adolescent treatment only, there were no differences in final overjet, peer assessment, and self-perception. However, incisal trauma was significantly lower in the early treatment group. Furthermore, headgear was found to help one in six children lower the risk of trauma. Functional appliances decreased risk in one in ten children. So where treatment outcomes between a 2-Phase and a 1-Phase treatment did not differ, an early phase did help decrease chance of incisal trauma (Thiruvenkatachari et al., 2013; Thiruvenkatachari et al., 2015).

MATERIALS AND METHODS

This study received Institutional Review Board Approval: IRB protocol #201509787. Records for this study were obtained from past-treated cases in the University Of Iowa Department Of Orthodontics.

Study Design:

This was a retrospective study which, when designed, was meant to included fifty surgical and fifty non-surgical Class II Division I malocclusion cases treated consecutively in the Department of Orthodontics.

Inclusion criteria were as follows:

1. Class II Division 1 Malocclusion (Class II molar relationship with proclined upper incisors)
2. Class II molar relationship quantified as at least 3mm
3. Initial overjet of ≥6mm when measured on casts
4. Patient was debanded between ages 13 and < 20 years of age
5. Patients with craniofacial anomalies or syndromes were eliminated from the study
6. Treatment types (2): non-surgical orthodontic-only or a combination of orthodontic/surgical treatment
7. Availability of full records

Data was gathered on all subjects. Initial and final lateral cephalometric radiographs were scanned into Dolphin Imaging software. The following cephalometric landmarks were traced and used for recording measurements: Sella, Porion, Orbitale, Nasion, A Point, B Point, U1 Incisal Edge, U1 Root Tip, L1 Incisal Edge, L1 Root Tip, Menton, and Constructed Gonion. Figure 1 and Figure 2 provide a visual representation.

Definitions of key landmarks, measurements, and angles are below.

- Sella (S): Center of the pituitary fossa of the sphenoid bone.
- Porion: Most superior aspect of the external auditory meatus.
- Orbitale: Lowest point of the roof of the orbit.
- Nasion (N): Marks the intersection of the nasofrontal suture and internasal suture, midsaggittaly.
- A point: Traced slightly in front of the apex of the most facial upper incisor; deepest point in the anterior concavity of the maxilla.
- B point: Deepest point in the concavity of the anterior border of the symphysis.
- Menton (Me): Marks the most inferior point on the symphysis.
- Constructed Gonion: Intersection of the mandible and a line that bisects the angle formed by Me-Go and Articulare-Ramus point.
- SNA angle: Angle formed when a line from Sella (S) to Nasion (N) is made and connected to a line from Nasion to A-point. This angle is used to determine the anterior posterior position of the maxilla.
- SNB angle: Angle formed when a line from Sella (S) to Nasion (N) is made and connected to a line from Nasion to B-point. This angle is used to determine the anterior posterior position of the mandible.
- ANB angle: Angle made when SNB is subtracted from SNA. Normal range is 0-3 degrees. This is used to measure relative position of the maxilla and mandible in relation to each other.
- Frankfurt Horizontal: A line through the inferior border of the orbit and the superior margin of the external auditory meatus.

- FMIA angle (Frankfort Mandibular Plane Incisor Angle): Angle formed from the intersection between a line through Frankfurt horizontal and a line along the straight axis of the most facial lower incisor.
- IMPA angle (Incisor to Mandibular Plane Angle): Angle formed from the intersection between a line through the mandibular plane and a line along the straight axis of the most facial lower incisor.
- U1 to SN angle: Intersection of two lines: one connecting S-N and one along the long axis of the upper incisor.
- Overbite: Superior-inferior overlap of the upper incisors to the lower incisors.
- Overjet: Anterior-posterior distance between the most facial upper incisor and the lower incisors.

Outcomes gathered in this study were: deband lateral cephalometric outcomes (ANB, FMIA, IMPA, U1 to SN, Overbite, Overjet), cast grading outcomes (measured through the ABO COGS), and retention protocol. Independent variables in this study were: the type of treatment (surgical vs non-surgical), the initial discrepancy index (DI), initial cephalometric variables (ANB, FMIA, IMPA, U1 to SN, Overbite, Overjet), starting age of treatment, and gender (Brady, 2016).

Cast grading was performed on pre and post-treatment casts. Initial casts were graded using parameters determined by the American Board of Orthodontics Initial Discrepancy Index Form (Figure 3) which is used to quantify the difficulty of an untreated case. This form takes into account both intraoral and cephalometric characteristics. Final casts were graded using the Final Cast Grading Form, also provided by the American Board of Orthodontics (Figure 4), which helped to provide a numerical representation of the finish of cases – high numbers indicated more occlusal discrepancies in a finished case.

Retention protocol included acrylic Hawley retainers with a labial bow, a fixed retainer bonded to anterior teeth, a tooth positioner, or a combination of one or more of these methods. Some candidates did not return for retention visits.

Reliability analysis was performed. To calculate intra-examiner reliability, one researcher measured initial casts and final casts for 20 cases two times within a one-week interval to over 0.99 positive correlation. Inter-examiner reliability was performed between two examiners. Both examiners used the initial discrepancy index form provided by the American Board of Orthodontists and also the ABO Cast Grading form, which details proper instruction for cast grading at deband. In addition, both examiners took the same online tutorial for final cast grading, thereby having the same degree of training prior to measuring data. Both examiners were blinded as to which cases were treated surgically and which were treated non-surgically. Correlation for inter-examiner reliability was >0.99. Cephalometric tracing also was reported with over 0.99 correlation found for both intra and inter-examiner reliability: two examiners each traced the same ten radiographs two times over the course of two consecutive weeks (intra-examiner) and a second examiner traced the same ten later to compare results (inter-examiner).

Statistical analyses were performed using the Mann-Whitney test. This test is applicable when testing the null hypothesis that two samples come from the same population (have the same median) and to compare differences between two random, independent samples. The Mann-Whitney test is considered the nonparametric counterpart to the independent samples t-test in that the assumptions that the difference between the samples of concern is normally distributed and that the variances of the two populations are equal are not applied. Therefore, the Mann-Whitney test is more appropriate for analyses where the validity of the t-test is suspect. Multiple linear regression was also performed to examine the association between treatment and final lateral cephalometric numbers (adjusted for initial cephalometric numbers, age at start of treatment, initial DI, gender). This statistical method is an extension of simple linear regression and is used to find an equation model that describes the relationship between two or more explanatory variables (i.e. independent variables or predictors) and a response variable (i.e. a dependent variable) by fitting a linear equation to the observed data.

Figure 1: Cephalometric Landmarks

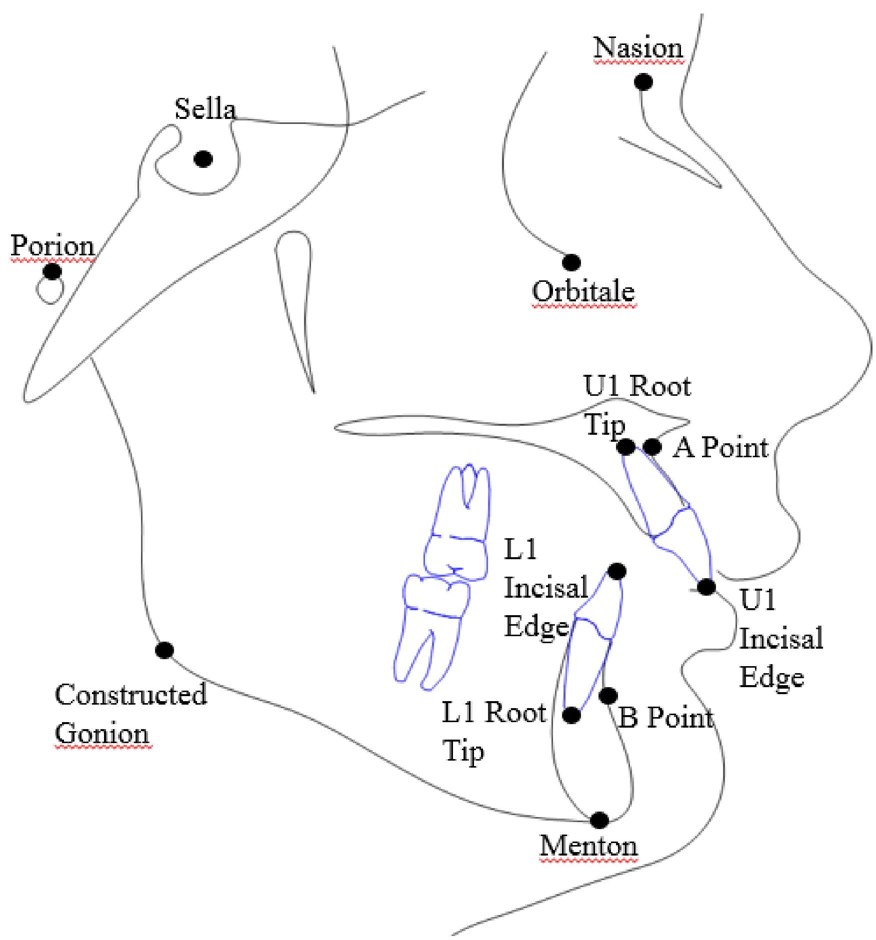


Figure 2: Angular and Linear Cephalometric Measurements

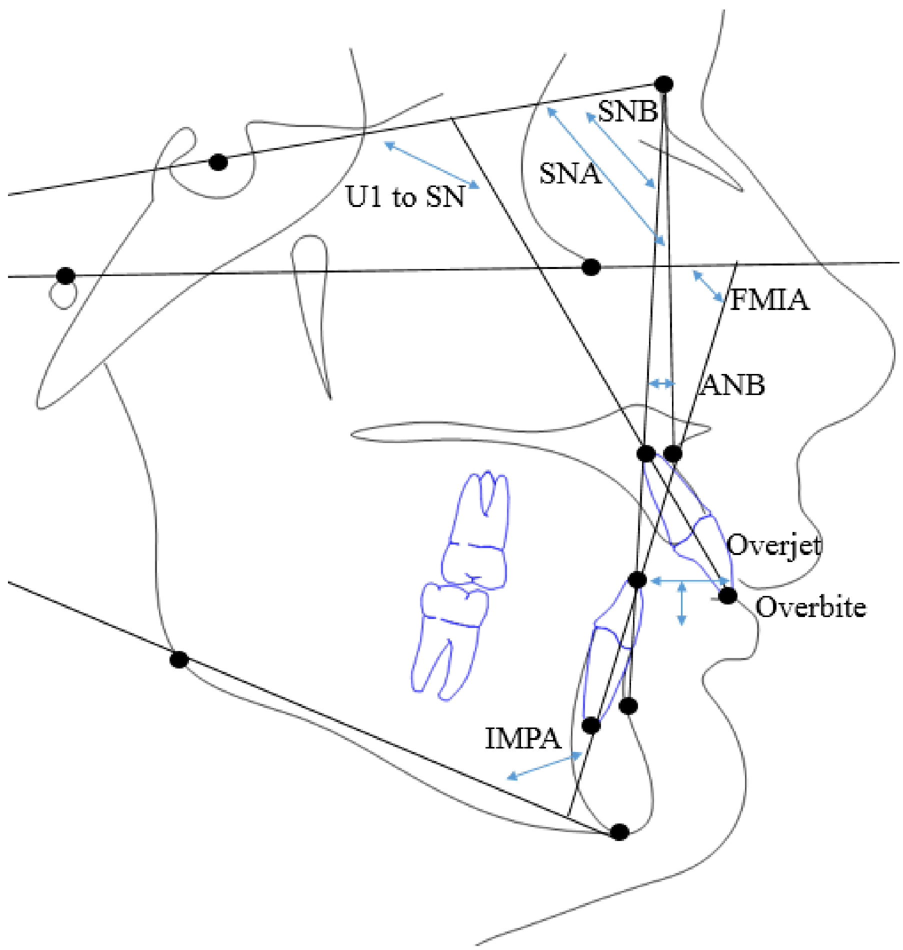


Figure 3: Initial Discrepancy Index Form

EXAM YEAR _______ **ABO DISCREPANCY INDEX**

ABO ID # _______ **CASE#** ______ **PATIENT** ______________________________

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	= ___
1̄ to MP ≥ 99°	@1pt	= ___
Each full degree > 99°	___x 1 pt	= ___
	Total	[ ]

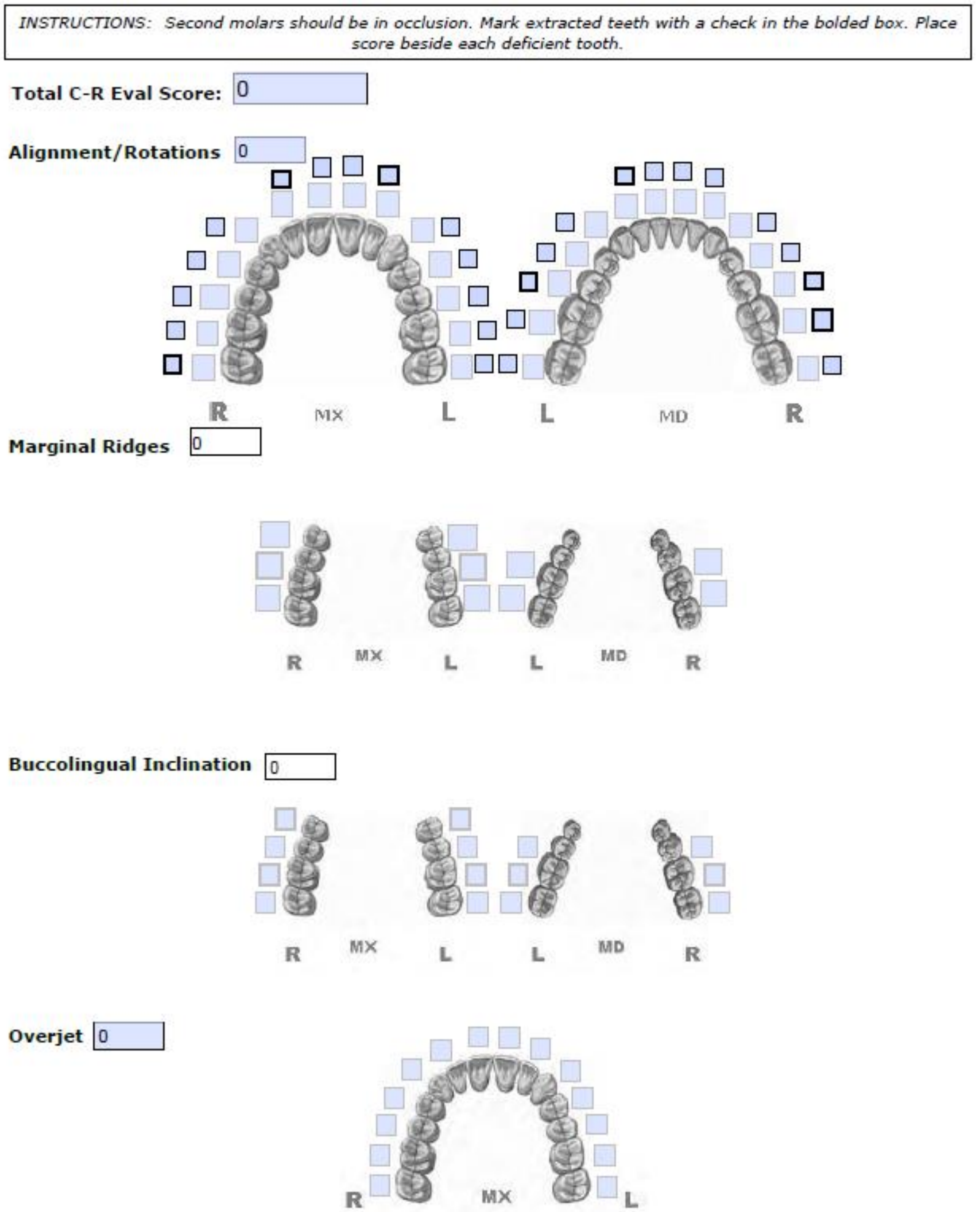
OTHER (See Instructions)

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 4: ABO Final Cast Grading Form


INSTRUCTIONS: Second molars should be in occlusion. Mark extracted teeth with a check in the bolded box. Place score beside each deficient tooth.

Total C-R Eval Score: 0

Alignment/Rotations 0

R MX L L MD R

Marginal Ridges 0

R MX L L MD R

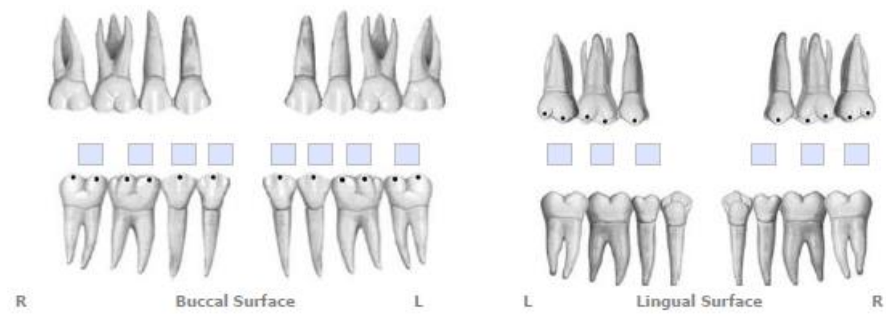
Buccolingual Inclination 0

R MX L L MD R

Overjet 0

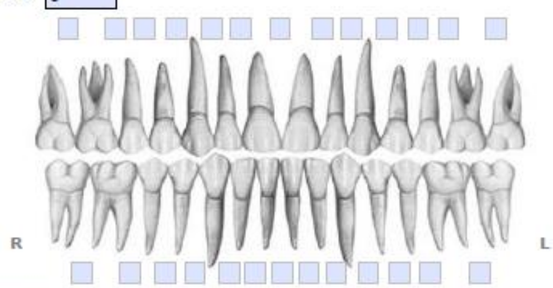
R MX L

Occlusal Contacts 0

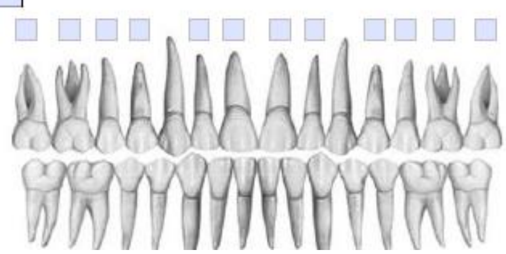


Occlusal Relationships 0

Interproximal Contacts 0



Root Angulation 0



RESULTS

After records were gathered, 60 patients were identified which fulfilled the inclusion criteria: forty non-surgical and twenty surgical cases (surgery was done in conjunction with orthodontic treatment) were included in the study. The study cohort was comprised of 28 female patients (21 in the non-surgical group and 7 in the surgical group) and 32 male patients (19 in the non-surgical group and 13 in the surgical group). Two patients identified as Hispanic and two as Caucasian-African American mixed. The remaining 56 patients were Caucasian.

The mean age of the surgical group at the start of treatment was 14.8 years (compared to 12.9 years in the non-surgical group) [$p<0.001$]. The mean age of the surgical group at the end of treatment was 17.4 years (compared to 15.4 years in the non-surgical group) [$p<0.001$]. The duration of treatment for the surgical group was 2.6 years compared with 2.5 years in the non-surgical group.

The mean initial discrepancy index score in the surgical group was 28.1 (compared to 20.0 in the non-surgical group) [$p=0.008$]. The final ABO deband score was 23.8 in the surgical group (compared to 22.5 in the non-surgical). Initial TSALD in the upper and lower arches for the surgical group was -1.0 and -3.4, respectively (in comparison to the non-surgical group: 0.1 upper and 0.1 lower). The initial SNA angle in the surgical group was 78.3 (compared to 78.6 in the non-surgical group). The initial SNB angle in the surgical group was 72.3 (compared to 74.8 in the non-surgical group) [$p=0.024$]. The initial ANB angle in the surgical group was 6.0 (compared to 3.9 in the non-surgical group) [$p=0.001$].The initial FMIA angle in the surgical group was 60.8 (compared to 60.4 in the non-surgical group). The initial IMPA angle in the surgical group was 91.7 (compared to 95.5 in the non-surgical group). The initial U1 to SN angle in the surgical group was 105.7

(compared to 107.6 in the non-surgical group). The initial ceph overbite in both the surgical and non-surgical group was 4.6. The initial ceph overjet in the surgical group was 10.1 (compared to 8.1 in the non-surgical group) [p=0.007]. The deband SNA angle in the surgical group was 77.7 (compared to 77.8 in the non-surgical group). The deband SNB angle in the surgical group was 75.1 (compared to 75.3 in the non-surgical group). The deband ANB angle in the surgical group was 2.6 (compared to 2.4 in the non-surgical group). The deband FMIA angle in the surgical group was 58.6 (compared to 56.0 in the non-surgical group). The deband IMPA angle in the surgical group was 92.3 (compared to 100.4 in the non-surgical group) [p<0.001]. The deband U1 to SN angle in the surgical group was 102.7 (compared to 101.6 in the non-surgical group). The deband ceph overbite was 1.5mm in the surgical (compared to 1.8 in the non-surgical group). The deband ceph overjet in the surgical group was 3.1mm (compared to 2.9 in the non-surgical group). The initial cast overjet in the surgical group was 10.1mm (compared to 8.3 in the non-surgical group) [p<0.002]. Finally, initial overbite in the surgical group was 3.9mm (compared to 4.6 in the non-surgical group) (Table 1).

Table 1: Comparison of Descriptives Between Treatment Groups

Characteristic	Non-Surgical			Surgical			
	Mean	Median	Std. Deviation	Mean	Median	Std. Deviation	p-value
Initial Discrepancy Index	20.0	18.5	6.8	28.1	25.0	13.8	0.008
Final ABO Deband Score	22.5	21.0	8.2	23.8	23.0	9.7	0.666
Initial TSALD Upper	0.1	0.0	3.1	-1.0	-0.7	6.0	0.415
Initial TSALD Lower	0.1	0.5	4.3	-3.4	-3.7	4.0	0.415
Starting Age (months)	154.6	151.5	20.9	177.1	179.0	16.1	p<0.0001
Starting Age (years)	12.9	12.6	1.7	14.8	14.9	1.3	p<0.0001
Deband Age (months)	184.8	180.0	18.4	208.7	207.0	15.2	p<0.0001
Deband Age (years)	15.4	15.0	1.5	17.4	17.3	1.3	p<0.0001
Treatment Duration (years)	2.5	2.3	0.8	2.6	2.6	0.8	0.227
Treatment Duration (months)	29.5	28.0	10.0	31.5	31.8	9.5	0.227
Initial SNA	78.6	78.4	3.6	78.3	78.1	2.6	0.820
Initial SNB	74.8	74.3	3.3	72.3	72.5	3.3	0.024
Initial ANB	3.9	4.1	1.8	6.0	6.0	2.1	0.001
Initial FMIA	60.4	60.0	7.2	60.8	59.4	9.3	0.969
Initial IMPA	95.5	94.4	6.6	91.7	90.5	8.4	0.068
Initial U1 to SN	107.6	107.7	6.4	105.7	107.1	9.2	0.666
Initial Ceph Overbite (mm)	4.6	5.0	1.8	4.6	5.1	3.7	0.931
Initial Ceph Overjet (mm)	8.1	8.3	2.0	10.1	9.4	2.6	0.007
Deband SNA	77.8	77.3	4.0	77.7	78.5	2.6	0.772
Deband SNB	75.3	74.9	4.2	75.1	75.6	3.6	0.944
Deband ANB	2.4	2.9	1.9	2.6	2.8	2.9	0.701
Deband FMIA	56.0	56.3	7.0	58.6	57.2	5.2	0.121
Deband IMPA	100.4	100.1	5.2	92.3	92.4	7.8	p<0.0001
Deband U1 to SN	101.6	100.1	7.6	102.7	101.9	10.2	0.772
Deband Ceph Overbite (mm)	1.8	2.0	0.7	1.5	1.7	1.0	0.146
Deband Ceph Overjet (mm)	2.9	2.8	1.2	3.1	2.8	0.9	0.354
Casts Initial Overjet (mm)	8.3	8.3	1.5	10.1	10.0	2.3	0.002
Cast Initial Overbite (mm)	4.6	5.0	1.8	3.9	4.5	2.8	0.080

All candidates in both groups were placed in orthodontic appliances.

59 candidates had no shift detected and there was 1 candidate in the non-surgical category with a small shift noted prior to treatment.

Final treatment plans in the non-surgical group was as follows: headgear only (2), headgear and elastic wear (14), headgear and upper first bicuspid extractions (3), one upper biscupid only (1), upper first bicuspid extractions only (3), four bicuspid extractions only (1), headgear in addition to upper premolar extractions and elastic wear (2), headgear and Forsus (3), Forsus correction only (1), Herbst and elastic wear (3), Herbst followed by headgear and elastic wear to hold correction (1), headgear as anchorage in conjunction with two bicuspid extractions (1), headgear as anchorage in conjunction with four bicuspid extractions (1), extraction of upper first premolars with TADs (1), one patient started with headgear then was offered surgery which was denied, and one patient chose the option of removing appliances once their occlusion was aligned, and another started with headgear then finished with elastics to TADs. Table 2 reviews these results.

Table 2: Final Treatment Plan in the Non-Surgical Treatment Group

Overall Treatment Type	Number of Candidates
Headgear only	2
Headgear and elastic wear	14
Headgear and upper first bicuspid extractions	3
One upper biscupid only	1
Upper first bicuspid extractions only	3
Four bicuspid extractions only	1
Headgear in addition to upper premolar extractions and elastic wear	2
Headgear and Forsus	3
Forsus correction only	1
Herbst and elastic wear	3
Herbst followed by headgear and elastic wear to hold correction	1
Headgear as anchorage in conjunction with two bicuspid extractions	1
Headgear as anchorage in conjunction with four bicuspid extractions	1
Extraction of upper first premolars with TADs	1
Started on HG and declines surgery	1
Deband once alignment was achieved	1
HG then elastics off TADs	1

Of the 20 surgical candidates, surgical breakdown was as follows: 1 piece maxillary impaction (2), BSSO advancement only (16), 2-jaw surgery (2).

Final treatment plans in the surgical group was as follows: non-extraction BSSO advancement (4), non-extraction with BSSO advancement and genioplasty (2), maxillary impaction and BSSO advancement (1), RME in conjunction with extraction of upper premolars and a BSSO/genioplasty (1), RME with 4 premolar extractions with BSSO/genioplasty (1), RME non-extraction with a BSSO/genioplasty (1), RME with four premolar extractions with maxillary impaction (1), extraction of all second premolars with a BSSO advancement (1), RME with extraction of lower first premolars then a BSSO (1), RME with extraction of lower first premolars then a BSSO with genioplasty (1), SARME with extraction of lower first premolars followed by a BSSO/genioplasty (1), extraction of four premolars with HG for anchorage followed by a surgery (1), extraction of lower first premolars and BSSO only (1), and 3 candidates had an unspecified surgery. These results are summarized in Table 3.

Table 3: Final Treatment Plan in the Surgical Treatment Group

Overall Treatment Type	Number of Candidates
Extraction of four premolars with HG for anchorage followed by a surgery	4
Non-extraction BSSO advancement and genioplasty	2
Maxillary impaction and BSSO advancement	1
RME in conjunction with extraction of upper premolars and a BSSO/genioplasty	1
RME with 4 premolar extractions with BSSO/genioplasty	1
RME non-extraction with a BSSO/genioplasty	1
RME with four premolar extractions with maxillary impaction	1
Extraction of all second premolars with a BSSO advancement	1
RME with extraction of lower first premolars then a BSSO	1
RME with extraction of lower first premolars then a BSSO with genioplasty	1
SARME with extraction of lower first premolars followed by a BSSO/genioplasty	1
Extraction of four premolars with HG for anchorage followed by a surgery	1
Extraction of lower first premolars and BSSO only	1
Unspecified surgery	3

Both groups utilized TADs or HG for anchorage purposes. The breakdown in each group was as follows: Non-surgical group: Headgear (29), headgear and TADs (1), TADs with no headgear (3), neither headgear nor TADs (7). Surgical group: headgear (3), TADs in the lower arch only (1), neither headgear nor TADs (16).

Final retention plan in each group was gathered. In the non-surgical group, options delivered were: Fixed U1-1 with Hawleys (2), Hawley retainers only (2), Hawley and bonded lower retainer (2), Hawley retainers only (33), tooth positioner then Hawley retainers (1). In the surgical group: Hawleys only (17), tooth positioner and Hawleys (2), and 1 candidate was given a tooth positioner and never returned for Hawleys.

Retainer Compliance was tracked to identify quality of treatment post-deband. Of the non-surgical candidates, 3 failed their recall appointments, 21 looked good at their first appointment recall, 6 looked “poor,” 1 had a poor post-deband outcome and was retreated, 3 had no retention checks post-deband, 1 looked good initially then had some relapse, 3 looked good at one appointment then no-showed their subsequent recalls, 1 looked good for one recall appointment then their retainer broke and they did not return, 1 looked good for one and a half years then lost their retainer and had a bonded lower 3-3 placed. Of the surgical candidates, 2 had fair outcomes at their first recall appointment, 1 had a fair recall and chose to have a bonded retainer placed in addition to Hawley retainers, 1 had a recall check but did not return for Hawley retainers, 9 looked good at their first recall, 1 look good at the second recall, 3 candidates had no recalls and 2 looked poor at their first recall.

Consent deband, indicating premature treatment completion, was tracked in each group. Of the non-surgical candidates, 32 did not have a consent deband. The remaining candidates opted for consent deband: 1 finished with a crossbite, 1 was due to patient burnout, 1 decided to stop treatment once correction could not be achieved and would consider surgery or extractions once older, and 5 consent debanded with no reason indicated. In the surgical group, 15 did not consent deband and the remaining 5 did, 1 with the reason being that they did not want to wear their elastics anymore.

Multivariable linear regression was performed to analyze results in the surgical population as they related to the following dependent variables: final ABO deband score, deband ANB angle, deband FMIA angle, deband IMPA angle, deband upper incisor to SN plane angle, deband cephalometric overbite, and deband cephalometric overjet. These results are summarized in Tables 4-9. The ABO deband score in the surgical treatment group was 0.854 points less than the non-surgical group. For every 1 year increase in starting age, the final ABO score in the surgical group decreased 1.656 points when all other predictors were held constant. For every 1 point increase in the initial discrepancy index score, the final ABO deband score increased 0.283 points when all other predictors were held constant (Table 4). The deband ANB angle in the surgical treatment group was 2.24 degrees less than the non-surgical group and this was statistically significant. For every 1 degree increase in initial ANB angle, the deband ANB angle increased 0.683 degrees when all other predictors were held constant. For every 1 degree increase in the initial FMIA angle, the deband ANB angle decreased 0.155 degrees when all other predictors were held constant. For every 1 degree increase in the initial IMPA angle, the deband ANB angle decreased 0.166 degrees when all other predictors were held constant (Table 5). The deband FMIA angle in the surgical treatment group was 0.649 degrees more than the non-surgical group. For every 1 year increase in start age, the deband FMIA angle increased 0.919 degrees when all other predictors were held constant. The deband FMIA angle for males was 3.01 degrees more than for females when all other predictors were held constant. For every 1 degree increase in Initial FMIA angle, the deband FMIA angle increased 0.793 degrees when all other predictors were held constant (Table 6). The deband IMPA angle in the surgical treatment group was 3.32 degrees less than the non-surgical group. For every 1 year increase in start age, the deband IMPA angle decreased by 0.972 degrees when all other predictors were held constant. For every 1 degree increase in Initial ANB angle, the deband IMPA angle decreased

0.911 degrees when all other predictors were held constant. For every 1 degree increase in Initial IMPA angle, the deband IMPA angle increased 0.532 degrees when all other predictors were held constant. For every 1mm increase in Initial ceph overjet, the deband IMPA angle increased 0.746 degrees when all other predictors were held constant (Table 7). The deband Upper Incisors to SN Plane angle in the surgical treatment group was 10.564 degrees more than the non-surgical group and this was statistically significant. For every 1 year increase in start age, deband upper incisor to SN plane angle in the surgical group decreased 1.33 degrees when all other predictors were held constant. For every 1 degree increase in Initial U1 to SN plane, deband upper incisor to SN plane angle increased 0.457 degrees when all other predictors were held constant (Table 8). The deband cepahlomateric overbite in the surgical treatment group was 0.606 mm less than the non-surgical group. For every 1 degree increase in Initial IMPA, deband cephalometric overbite increased 0.075 mm when all other predictors were held constant. For every 1 mm increase in initial ceph OJ, deband cephalometric overbite increased 0.156 mm when all other predictors were held constant (Table 9). Deband cephalometric overjet in the surgical treatment group was 0.188 mm more than the non-surgical group (Table 10).

Table 4: Multivariable Linear Regression - Final ABO Deband Score

Predictor Variable	Estimate (95% CI)	p-value
Final Surgical Treatment Plan	-0.854 (-7.75-6.04)	.804
Starting Age (years)	-1.656 (-2.99 - -0.31)	.017
Gender (Males)	-0.038 (-4.59 - 4.52)	.987
Initial Discrepancy Index	0.283 (-0.003-0.57)	.052
Initial ANB	0.677 (-0.68-2.04)	.321
Initial FMIA	0.091 (-0.41-0.59)	.716
Initial IMPA	-0.281 (-0.86-0.30)	.334
Initial U1 to SN	-0.092 (-0.54-0.36)	.685
Initial Ceph OB	0.589 (-0.57-1.75)	.312
Initial Ceph OJ	0.141 (-1.36-1.64)	.851

Table 5: Multivariable Linear Regression – Deband ANB Angle

Predictor Variable	Estimate (95% CI)	p-value
Final Surgical Treatment Plan	-2.24 (-3.62 - -0.86)	.002
Starting Age (years)	0.143 (-0.13-0.41)	.291
Gender (Males)	-0.813 (-1.73-0.10)	.080
Initial Discrepancy Index	0 (-0.06-0.06)	.999
Initial ANB	0.683 (0.41-0.96)	p<0.001
Initial FMIA	-0.155 (-0.26- -0.06)	.003
Initial IMPA	-0.166 (-0.28- -0.50)	.006
Initial U1 to SN	-0.003 (-0.09-0.09)	.949
Initial Ceph OB	0.190 (-0.04-0.42)	.106
Initial Ceph OJ	0.135 (-0.17-0.44)	.370

Table 6: Multivariable Linear Regression - Deband FMIA

Predictor Variable	Estimate (95% CI)	p-value
Final Surgical Treatment Plan	0.649 (-3.50-4.80)	.754
Starting Age (years)	0.919 (0.11-1.73)	.026
Gender (Males)	3.01 (0.27-5.75)	.032
Initial Discrepancy Index	0.031 (-0.14-0.20)	.715
Initial ANB	0.579 (-0.24-1.40)	.160
Initial FMIA	0.793 (0.5-1.09)	p<0.001
Initial IMPA	0.277 (-0.07-0.63)	.116
Initial U1 to SN	0.263 (-0.01-0.53)	.057
Initial Ceph OB	-0.567 (-1.26-0.13)	.108
Initial Ceph OJ	-0.355 (-1.26-0.55)	.433

Table 7: Multivariable Linear Regression - Deband IMPA

Predictor Variable	Estimate (95% CI)	p-value
Final Surgical Treatment Plan	-3.321 (-7.17-0.53)	.089
Starting Age (years)	-0.972 (-1.72 - -0.22)	.012
Gender (Males)	-2.02 (-4.56-0.52)	.116
Initial Discrepancy Index	-0.043 (-0.202-0.12)	.588
Initial ANB	-0.911 (-1.67- -0.15)	.019
Initial FMIA	-0.104 (-0.38-0.17)	.454
Initial IMPA	0.532 (0.21-0.86)	.002
Initial U1 to SN	-0.209 (-0.46-0.42)	.101
Initial Ceph OB	0.746 (0.10-1.39)	.024
Initial Ceph OJ	0.704 (-0.13-1.54)	.097

Table 8: Multivariable Linear Regression - Deband Upper Incisor to SN Plane

Predictor Variable	Estimate (95% CI)	p-value
Final Surgical Treatment Plan	10.564 (4.30-16.83)	.001
Starting Age (years)	-1.33 (-2.55 - -0.11)	.033
Gender (Males)	2.01 (-2.13-6.15)	.335
Initial Discrepancy Index	-0.161 (-0.42-0.10)	.217
Initial ANB	-0.925 (-2.16-0.31)	.138
Initial FMIA	0.362 (-0.09-0.81)	.115
Initial IMPA	0.306 (-0.22-0.83)	.248
Initial U1 to SN	0.457 (0.05-0.87)	.030
Initial Ceph OB	-0.236 (-1.3-0.82)	.654
Initial Ceph OJ	-1.14 (-2.5-0.23)	.100

Table 9: Multivariable Linear Regression - Deband Cephalomatric Overbite

Predictor Variable	Estimate (95% CI)	p-value
Final Surgical Treatment Plan	-0.606 (-1.27-0.05)	.071
Starting Age (years)	0.059 (-0.07-0.19)	.357
Gender (Males)	0.254 (-0.18-0.69)	.248
Initial Discrepancy Index	0.022 (-0.01-0.05)	.107
Initial ANB	-0.028 (-0.16-0.10)	.666
Initial FMIA	0.034 (-0.01-0.08)	.158
Initial IMPA	0.075 (0.02-0.13)	.009
Initial U1 to SN	-0.002 (-0.05-0.04)	.918
Initial Ceph OB	0.077 (-0.03-0.19)	.168
Initial Ceph OJ	0.156 (0.01-0.30)	.034

Table 10: Multivariable Linear Regression - Deband Cephalomatric Overjet

Predictor Variable	Estimate (95% CI)	p-value
Final Surgical Treatment Plan	0.188 (-0.82-1.20)	.709
Starting Age (years)	0.035 (-0.16-0.23)	.721
Gender (Males)	0.082 (-0.58-0.75)	.806
Initial Discrepancy Index	-0.007 (-0.05-0.04)	.750
Initial ANB	-0.079 (-0.28-0.12)	.429
Initial FMIA	0.011 (-0.06-0.08)	.757
Initial IMPA	0.016 (-0.07-0.10)	.699
Initial U1 to SN	-0.018 (-0.08-0.05)	.580
Initial Ceph OB	-0.032 (-0.20-0.14)	.702
Initial Ceph OJ	0.073 (-0.15-0.30)	.505

Table 11: Summary of Multiple Linear Regression Findings, Surgical Compared to Non-Surgical

Dependent Variable	**Final Surgical Treatment Plan**	**Estimate (95% CI)**	**p-value**
Final ABO deband score	-0.854	-7.75, 6.04	0.804
Deband ANB Angle	-2.24	-3.62, -0.86	0.002
Deband FMIA	0.649	-3.50, 4.80	0.754
Deband IMPA	-3.321	-7.17, 0.53	0.089
Deband Upper Incisor to SN Plane	10.564	4.30, 16.83	0.001
Deband Cephalometric Overbite	-0.606	-1.27, 0.05	0.071
Deband Cephalometric Overjet	0.188	-0.82, 1.20	0.709

DISCUSSION

Class II Division I malocclusions present in the European population more than any other malocclusion. Treatment approaches are three pronged: orthopedics (examples are seen in Table 12 below), masking (typically involving extractions of upper premolars or upper and lower premolars), or orthognathic surgery (one jaw, two jaws, or both). A combination of treatment approaches might also be utilized. Practioners must be wary of choosing irreversible treatment plans too early in life as they may eliminate other, better options later.

Final treatment in the non-surgical treatment group included headgear usage in 68% of patients. At the University of Iowa, patients are primarily treated with headgear if they present with a Class II Division I malocclusion and have not yet hit their pubertal growth spurt. The effects of a headgear are to prevent downward and forward movement of the maxilla, thereby holding A point in the same location in space. Maxillary molars will expectedly intrude and distalize to some extent. A Herbst appliance has similar effects to a headgear, however, it has the added effect of proclining the lower incisors a net total of three degrees.[37] Due to this finding, the ideal Herbst patient initiates treatment with upright lower incisors and a robust biotype in the lower anterior mandibular region. This appliance is useful in that its usage does not depend on patient compliance.

Table 12: Orthopedic Options for Class II Division I Treatment

Orthopedic Appliances for Early Correction of Class II Malocclusion
Headgear
Herbst
Twin Block
Frankel Appliance
Bionator

The option of extracting to camouflage the discrepancy is often chosen if the surgical option cannot be entertained and the patient is beyond their growth spurt. In these cases, practitioners must be mindful of soft tissue changes when teeth are extracted. In the maxilla, the upper lip comes back 1mm for every 2mm of tooth retraction and the lower lip comes back 1mm for every 1mm of dental retraction. While soft tissue changes were not documented in this study, we must be mindful of the implications of such a treatment on soft tissue profile and educate the patient accordingly.

The final option, a surgical option was analyzed by our data in great detail. When choosing this treatment mode, not only can we achieve skeletal base discrepancy improvement, but also soft tissue profile change. Functional improvement may also be an added positive finding, as a BSSO advancement might help a patient who initially presents with airway obstruction or sleep apnea.[38,39] Masticatory performance might also improve. An article by Jeryl English looked at Class I, II, and III occlusal relationships and the ability of each type of occlusion on chewing function. This article showed that Class III malocclusion had the worst ability to break down particles, followed by Class II, and finally Class I.[40]

The goal of our study was to compare initial treatment casts to final deband casts of severe Class II Division I malocclusions treated surgically and non-surgically. Parameters of cast grading were the AAO initial discrepancy index and the final cast grading instruction. Past research conducted at the University of Iowa in 2016 analyzed cephalometric outcomes in our same population, and this information was integrated into this study. Our aim was to compare the end of treatment outcomes in orthognathic surgery versus non-surgical treatment. We hypothesized that occlusal outcomes would be similar at deband (when braces were removed) and that cephalometric values would differ. Our findings showed that cast-grading outcomes were similar between the surgical and the non-surgical treatment groups and that cephalometric values differed.

For this study, the Mann-Whitney U test was utilized. This is the nonparametric counterpart to the student t-test which is used to check if there is a differences between the mean of two populations assuming the populations and random and independent – meaning the outcome of one does not affect the outcome on the other. When the t-test is used, there is also the assumption that there is a normal distribution. The Mann Whitney test is a more robust version of this analysis because it can be utilized for a larger number of scenarios. Here, one does not have to assume a normal distribution but the difference between the means of two populations is still under scrutiny. Multiple linear regression was used to fit a model to describe the relationship between two or more independent variables (here, we were look at a number of cephalometric and cast-specific items) and a dependent variable. Our dependent variables for this study were: final ABO deband score, deband ANB angle, deband FMIA angle, deband IMPA angle, deband upper incisor to SN plane angle, deband cephalometric overbite, and deband cephalometric overjet. Reliability data was gathered between two examiners for cast grading measurements. These values can be found in Tables 13 and 14. For cephalometric data, both inter and intra-examiner data was gathered (Table 15) Correlation in both groups was above 97%.

Table 13: Inter-examiner Reliability Analysis – Initial Discrepancy Index

Intraclass Correlation Coefficient		
	Intraclass Correlation	**95% Confidence Interval**
Single Measures	0.941	(0.858, 0.976)
Average Measures	0.970	(0.924, 0.988)

Table 14: Inter-examiner Reliability Analysis – ABO Final Cast Grading

Intraclass Correlation Coefficient		
	Intraclass Correlation	**95% Confidence Interval**
Single Measures	0.987	(0.968, 0.995)
Average Measures	0.994	(0.984, 0.997)

Table 15: Inter/Intra-Examiner Reliability Analysis – Cephalometric Data

Cephalometric Measurement	Inter-Examiner Reliability	Intra-Examiner Reliability
SNA	0.824	0.952
SNB	0.977	0.985
ANB	0.901	0.971
FMIA	0.953	0.980
IMPA	0.890	0.942
U1 to SN	0.921	0.966
Overbite	0.964	0.977
Overjet	0.995	0.991

Our results showed that the final deband score in the surgical group was 0.854 points less than the non-surgical group. The deband ANB angle was 2.24 degrees less than the non-surgical groups (p=0.002). This indicates that the maxilla/mandible relationship improved in the surgical group to a greater extent. This can be expected since skeletal positions are changing with this treatment plan.

Both FMIA and IMPA angle give information about lower incisor position. The deband FMIA angle in the surgical group was found to be 0.649 degrees more than the non-surgical group – meaning that when lower incisors position was measured from the Frankfort plane to the mandibular plane, the lower incisors proclined slightly. One of the drawbacks of the FMIA angle measurement is that it does not realize that when the mandibular plane angle changes, the incisor position remains the same. The IMPA angle is arguably more useful at providing a more accurate value of true incisor proclination or retroclination. The deband IMPA angle in the surgical treatment group was 3.321 degrees less than the non-surgical treatment group meaning the incisors were more upright at the end of treatment. We would expect these values to change according to which treatment plan, surgical or non-surgical, was chosen. For instance, if lower premolars were extracted before a BSSO advancement, we would expect some uprighting of the lower incisors when this space is closed. Also, initial crowding would have an effect on incisor position. If no extractions were done, the way to gain arch length to resolve lower anterior cording is to proline the lower incisors.

The deband upper incisors to SN plane angle, a measurement of upper incisor position, was shown to be 10.562 degrees greater in the surgical treatment group (p=0.001). We might expect this finding because, in non-surgical treatment plan for severe Class II Division I malocclusion, if a surgical option cannot be entertained, a masking approach by extracting upper premolars is most likely considered instead. During space closure of the extraction spaces, there will be uprighting of the upper incisors, thereby leading to a decreased upper incisor to SN plane angle. This finding could also be explained in that for a surgical treatment option where teeth might not need to be extracted, if there is an existing upper anterior crowding, incisors will be proclined to gain space for alignment.

Deband cephalometric overbite was found to be 0.606mm less in the surgical group while deband cephalometric overjet was shown to be 0.188mm greater than the non-surgical treatment group. These findings can be explained when we consider consequences of treatment plans in the vertical dimension in both treatment groups. In a non-surgical approach, if extractions are performed, not only will upper incisors upright, as mentioned before, but the bite will also deepen as extraction spaces are closed. If no extractions are performed, as in a surgical treatment plan, overjet will increase as upper incisors are proclined to allow for alignment of these teeth. One way to minimize the amount of proclination might be to consider interproximal reduction, or slenderizing of the mesial and distal aspect of these teeth.

One of the first pieces of literature analyzing need for orthognathic surgery based on severity was put forth by Proffit in 1992. When reviewing an adolescent population treated non-surgically (through camouflage treatment) or surgically, Proffit identified certain parameters which might be useful when deciding treatment. He evaluated cephalometric and cast measurement before and after treatment to determine efficacy – which he considered any correction to what he deemed "normal" standards. Our study found end-of-treatment occlusion to be similar in both groups. This was supported by Proffit's work as well. Our study showed that overjet was slightly higher in the surgical group, which was not seen in Proffit's results. This could be attributed to differing practitioners' approach to treatment or a variability in the success of the surgical treatment in either surgical population (Proffit, 1992).

Mihalik performed a long-term follow-up of Class II adults treated with camouflage treatment or surgical treatment and analyzed results post-deband. Patients in this population were recalled 12 years out of treatment. This group found that both groups showed acceptable correction of the malocclusion. This was echoed by our study which found ABO cast grading outcomes in the non-surgical vs. surgical population at the end of treatment were very similar (22.5 vs 23.8 respectively). At recall, Mahalik reported that in both populations, overbite increased to a small extent and overjet increased in the surgical group by 10-20%. Our study found that deband cepahlomatric overbite was less in the surgical group (versus non-surgical) and overjet was increased compared to the non-surgical group, although neither value was significant. This might be expected because with camouflage treatment, as the upper incisors are retracted, overbite increases (Mihalik, 2003).

In an adolescent population (less than 20 years of age), Tulloch discussed the difficulty in treatment planning as these patients might still be undergoing growth. This study emphasized that in severe cases, surgical treatment is most likely the best option. They looked at 500 patients in a study with similar inclusion criteria as ours: craniofacial anomalies or syndromes patients were excluded and patients has to present Class II with an overjet of 6mm or more. Patients were treated non-surgically or surgically and end-of-treatment outcomes were reviewed based on division into three categories: orthodontic success, orthodontic failure, and surgical success. This study assessed success of treatment through reduction in overjet to less than 4mm. Cephalometric radiographs were reviewed and patients were placed into two sub-groups based on gender. Initial ANB in this study was about 6 degrees, similar to that in our study but initial overjet measurement is both groups were significantly more (7.8 and 8.6mm) when compared to our population (2.9 vs. 3.1mm). This study found that 98% of patients in their entire population did not meet their criteria for correction of overjet. Since our study evaluated ANB change as a measure of AP correction, we were able to focus exclusively on skeletal position instead of tooth position. Tulloch concluded that neither gender nor age were associated with success of correction of overjet and concluded that more factors go into a “successful” or “unsuccessful” case than practitioners might think. Since these factors were held constant in our linear regression models, we were able to analyze differing variables without the risk of bias (Tulloch, 1999).

In Kinzinger’s study of Class II Division I populations, 60 young adults were evaluated after a surgical or non-surgical treatment. Their results showed changes in all skeletal categories, as can be presumed because with this method of treatment, the skeletal base is being influences directly. Each group in this study achieved a reduction in overjet. The surgical group was found to have significant protrusion of upper incisor position, as did our research. This might be attributed to the biomechanical differences in treating a surgical case versus a non-surgical case. One might imagine that not only will a camouflage treatment increase overbite as incisors are retracted, but they will also upright. If a surgical patient has minimal crowding, it might not be outside the realm of possibility that the practitioner might choose to procline the upper incisors to allow for alignment of teeth before the patient is sent for surgery (Kinzinger, 2009).

LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

The quality of existing literature in this area has primarily been plagued with varying degrees of bias due to the retrospective nature of the studies. Many articles have had higher dropout rates than anticipated, most likely due to the length of the study, which aimed to identify immediate and long-term final treatment outcomes. Our study, while subject to selection bias due to its retrospective nature, was able to include retention records for many candidates. By blinding examiners as to the type of treatment performed before cast grading was done, we were able to eliminate some bias which could have been associated with cast grading. Past literature has also trended towards having limited comprehensive information available to researchers that fit the study criteria and small sample sizes. Here, while we were able to provide a comprehensive documentation of each patient's overall treatment plan, we were limited by the number of cases which fit our inclusion criteria.

Studies such as this, done at a single institution are beneficial because researchers are able to draw patients from a homogeneous population. Therefore, since the same treatment approach was offered to all patients, variables associated with different treatment methods were eliminated. This will be consistent for all patients. Once this study achieves the capacity to extend to reach a broader number of institutions, we will be able to draw treatment data from a larger population who undoubtedly will have utilized different types of treatment for correction. Next steps might be capturing retention records past the point of deband and one year past deband. This could provide unseen differences between non-surgical and surgical treatment plans which were unidentifiable short-term.

CONCLUSION

We can conclude that amongst Class II Division I cases identified in this study, final deband cast outcomes were similar between the non-surgical and surgical treatment groups. Cephalometrically, there were some differences in deband outcomes between non-surgical and surgical populations. Those treated surgically had a significantly larger reduction in ANB angle (p=0.002) and increased maxillary incisor proclination (p=0.001) compared to those treated non-surgically. Further information should be gathered at other institutions to compile a more diverse picture of successful treatment options in the Class II Division I population.

REFERENCES

Baccetti, T., Franchi, L., Stahl, F. Comparison of 2 comprehensive Class II treatment protocols including the bonded Herbst and headgear appliances: a double-blind study of consecutively treated patients at puberty. Am J Orthod Dentofacial Orthop. 2009; 135(6):698.e1-10; discussion 698-699.

Berger, J.L., Pangrazio-Kulbersh, V., George, C., Kaczynski, R. Long-term comparison of treatment outcome and stability of Class II patients treated with functional appliances versus bilateral sagittal split ramus osteotomy. Am J Orthod Dentofacial Orthop 2005; 95: 250-258.

Bollhalder, J., et al. Dentofacial and upper airway characteristics of mild and severe Class II division 1 subjects. European Journal of Orthodontics. 2013; 35:447-453.

Brady, Patrick, et al. Cephalometric analysis of adolescents with severe Class II Division I malocclusions treated surgically and non-surgically. University of Iowa. 2016. 1-36.

Burden, D., Johnston, C., Kennedy, D., et al. A cephalometric study of Class II malocclusion treated with mandibular surgery. Am J Orthod Dentofacial Orthop 2007; 1-8.

Cacciatore, G., Alvetro, L., Defraia, E., Ghislanzoni, L.T.H., Franchi, L. Active-treatment effects of the forsus fatigue resistant device during comprehensive Class II correction in growing patients. Korean J Orthod 2014; 44(3):136-142.

Cacciatore, G., Ghislanzoni, L.T.H., Alvetro, L., Giuntini, V., Franchi, L. Treatment and posttreatment effects induced by the Forsus appliance: A controlled clinical study. Angle Orthodontist 2014; (84) 1010-1017.

Cassidy, D.W., Herbosa, E.G., Rotskoff, K.S., et al. A comparison of surgery and orthodontics in “borderline” adults with Class II, division 1 malocclusions. Am J Orthod Dentofacial Orthop 1993; 104:455-470.

Chaiyongsirisern, A., Rabie, A.B., Wong, R.W. Stepwise advancement Herbst appliance versus mandibular sagittal split osteotomy. Treatment effects and long-term stability of adult Class II patients. Angle Orthod. 2009; 79 (6):1084-1094.

de Lir Ade, L., de Moura, W.L., Oliveira Ruellas, A.C., Gomes Souza, M.M., Nojima, L.I. Long-term skeletal and profile stability after surgical-orthodontic treatment of Class II and Class III malocclusion. J Craniomaxillofac Surg. 2013; 41(4):296-302.

English, J., Buschang, P., Throckmorton, G. Does Malocclusion Affect Masticatory Performance? Angle Orthodontist. 2002; 72: 21-27.

Hansen, K., Koutsonas, T., Pancherz, H. Long-term effects of Herbst treatment on the mandibular incisor segment: A cephalometric and biometric investigation. Am J Orthod Dentofac Orthop. 1997; 112: 92-103.

Jambi, S., Thiruvenkatachari, B., O'Brien, K.D., Walsh, T. Orthodontic treatment for distalising upper first molars in children and adolescents. Cochrane Database Syst Rev. 2013; 11: 1-78.

Janson, G., Sathler, R., Fernandes, T.M., Branco, N.C., Freitas, M.R. Correction of Class II malocclusion with Class II elastics: a systematic review. Am J Orthod Dentofacial Orthop. 2013;143(3):383-392.

Janson, G., Sathler, R., Fernandes, T.M., Zanda, M., Pinzan, A. Class II malocclusion occlusal severity description. J Appl Oral Sci, 18 (2010), pp. 397–402.

Kinzinger, G., Diedrich, P. Skeletal effects in Class II treatment with the Functional Mandibular Advancer (FMA)? J Orofac Orthop 2005; 66:46-58.

Kinzinger, G., Frye, L., Diedrich, P. Class II treatment in adults: comparing camouflage orthodontics, dentofacial orthopedics and orthognathic surgery--a cephalometric study to evaluate various therapeutic effects. J Orofac Orthop. 2009; 70(1):63-91.

Ko, E.W., Hsu, S.S., Hsieh, H.Y., Wang, Y.C., Huang, C.S., Chen, Y.R. Comparison of progressive cephalometric changes and postsurgical stability of skeletal Class III correction with and without presurgical orthodontic treatment. J Oral Maxillofac Surg. 2011; 69(5):1469-1477.

Koletsi, D., Pandis, N., Fleming, P.S. Sample size in orthodontic randomized controlled trials: are numbers justified? Eur J Orthod. 2014; 36(1):67-73.

Lempesi, E., Koletsi, D., Fleming, P., Pandis, N. The reporting quality of Randomised Controlled Trials in Orthodontics. J Evid Base Dent Pract 2014; 14:46-52.

Lohrmann, B., Schwestka-Polly, R., Nagerl, H,. et al. The influence of orthodontics and mandibular sagittal split osteotomy on dental and skeletal variables- a comparative cephalometric study. Eur J Orthod 2006; 28:553-560.

Marsico, E., Gatto, E., Burrascano, M., Matarese, G., Cordasco, G. Effectiveness of orthodontic treatment with functional appliances on mandibular growth in the short term. Am J Orthod Dentofacial Orthop. 2011; 139(1):24-36.

Martins-Junior, P.A., Marques, L.S., Ramos-Jorge, M.L., Malocclusion: social, functional and emotional influence on children. J Clin Pediatr Dent. 2011; 37(1):103-108.

Mihalik, C.A., Proffit, W.R., 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.

Millett, D.T., Cunningham, S.J., O'Brien, K.D., Benson, P., Williams, A., de Oliveira, C.M. Orthodontic treatment for deep bite and retroclined upper front teeth in children. Cochrane Database Syst Rev. 2006; 10:1-63.

Millett, D.T., Cunningham, S.J., O'Brien, K.D., Benson, P.E., de Oliveira, C.M. Treatment and stability of class II division 2 malocclusion in children and adolescents: a systematic review. Am J Orthod Dentofacial Orthop. 2012; 142(2):159-169.

Pancherz, H., Anehus-Pancherz, M. Facial profile changes during and after Herbst appliance treatment. Eur J Orthod 1994; 16:275-286.

Perillo, L., Cannavale, R., Ferro, F., Franchi, L., Masucci, C., Chiodini, P., Baccetti, T. Meta-analysis of skeletal mandibular changes during Frankel appliance treatment. Eur J Orthod. 2011; 33(1):84-92.

Proffit, W.R., Fields, H.W., Sarver, D.M. Contemporary Orthodontics (4th edn), Mosby Elsevier. 2007; Chapter 1.

Proffit, W.R., Philips, C., Turvery, T.A. Long-term stability of adolescent versus adult surgery for treatment of mandibular deficiency. Int. J. Oral Maxillofac. Surg. 2010; 39: 327-332.

Proffit, W.R., Phillips, C., Douvartzidis, N. A comparison of outcomes of orthodontic and surgical-orthodontic treatment of Class II malocclusion in adults. Am J Orthod Dentofacial Orthop. 1992; 101(6):556-565.

Proffit, W.R., Phillips, C., Tulloch, J.F., Medland, P.H. Surgical versus orthodontic correction of skeletal class II malocclusion in adolescents: effects and indications. Int J Adult Orthod Orthognath Surg. 1992 (7): 209-220.

Ruf, S., Pancherz, H. Orthognathic surgery and dentofacial orthopedics in adults Class II Division 1 treatment: Mandibular sagittal split osteotomy versus Herbst appliance. Am J Orthod Dentofacial Orthop. 2004; 126: 140-156.

Schultz, T. et al. Class II correction improves nocturnal breathing in adolescents. Angle Orthodontist. 2011; 81:222-228.

Seehara, J., Fleming, P.S., Newton, T., DiBiase, A.T. Bullying in orthodontic patients and its relationship to malocclusion, self-esteem and oral health-related quality of life. J Orthod. 38: 247-56, 2011

Sloss, E.A., Southard, K.A., Qian, F., Stock, S.E., Mann, K.R., Meyer, D.L., Southard, T.E. Comparison of soft-tissue profiles after treatment with headgear or Herbst appliance. Am J Orthod Dentofacial Orthop. 2008; 133(4):509-514.

Thiruvenkatachari, B., Harrison, J.E., Worthington, H.V., O'Brien, K.D. Early orthodontic treatment for Class II malocclusion reduces the chance of incisal trauma: Results of a Cochrane systematic review. Am J Orthod Dentofacial Orthop 2015; 148:47-59.

Thiruvenkatachari, B., Harrison, J.E., Worthington, H.V., O'Brien, K.D. Orthodontic treatment for prominent upper front teeth (Class II malocclusion) in children. Cochrane Database Syst Rev. 2013; 7:209-220.

Tucker, M.R. Orthognathic surgery versus orthodontic camouflage in the treatment of mandibular deficiency. J Oral Maxillofac Surg. 1995; 53(5):572-578.

Tulloch, C., Lenz, B., Phillips, C. Surgical versus orthodontic correction for Class II patients: age and severity in treatment planning and treatment outcome. Semin Orthod. 1999; 5:231-240.

Wigal, T.G., Dischinger, T., Martin, C., Razmus, T., Gunel, E., Ngan, P. Stability of Class II treatment with an edgewise crowned Herbst appliance in the early mixed dentition: Skeletal and dental changes. Am J Orthod Dentofacial Orthop. 2011; 140(2):210-223.