SERIAL EXTRACTIONS VERSUS LATE PREMOLAR EXTRACTIONS

Kevin O'Shaughnessy

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Science in the School of Dentistry (Orthodontics).

Chapel Hill
2007

Approved by:

Lorne Koroluk

Garland Hershey

Andrea K. Biddle

ABSTRACT

KEVIN O'SHAUGHNESSY: Serial Extractions versus Late Premolar Extractions
(Under the direction of Lorne Koroluk)

The purpose of this study was to compare treatment times for serial extractions (SE) and late premolar extractions (LPE) for a cost-effectiveness analysis. A retrospective chart review identified 51 SE and 49 LPE patients. PAR scores were obtained at the initiation (PAR T1) and completion (PAR T2) of active orthodontic treatment. PAR T1 scores for the SE patients were significantly less than that for the LPE patients ($p<0.001$) whereas PAR T2 scores were not significantly different. Active-treatment time was significantly less for the SE group as compared to the LPE group ($p<0.001$). Total time (mos. pre-active + active-treatment) and total number of visits was significantly greater for the SE group as compared to the LPE group ($p<0.001$). Total chair time (min.) was not significantly different between groups. Serial extractions may reduce active-treatment time for severe crowding but a significant time and effort precedes active-treatment.

ACKNOWLEDGMENTS

I would like to thank the following people: Dr. Lorne Koroluk for his consistent guidance, support, and insight throughout the development and completion of this project; Dr. Andrea K. Biddle for her tremendous help with the economic modeling and statistics; Dr. Garland Hershey for his support throughout my orthodontic education and his help with writing the thesis; Dr. David Kennedy who was a gracious host in Vancouver, CA and allowed access to his patient record database which was used in this study; Dr. Phillips and Ms. Se Hee Kim for their help with the statistical analysis and interpretation; and Debbie Price for her help with the figures.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF ABBREVIATIONS

LPE	Late Premolar Extraction
mm	millimeters
PAR index	Peer Assessment Rating index
PAR T0	The Peer Assessment Rating index score at time point zero
PAR T1	The Peer Assessment Rating index score at time point one
PAR T2	The Peer Assessment Rating index score at time point two
SE	Serial Extraction
SE-Interceptive	A hypothetical fee structure whereby a patient is charged a fixed fee by the orthodontist for the serial extraction portion of the treatment
SE-Observation	A hypothetical fee structure whereby a patient is charged a per-visit observation fee by the orthodontist for the serial extraction portion of the treatment
SE-Discounted	A hypothetical fee structure whereby a patient is charged no fee by the orthodontist for the serial extraction portion of the treatment and is given a discount for the active-treatment portion
T0	Time point zero, representing the time when pre-serial extraction diagnostic records were taken for the SE group
T1	Time point one, representing the time when pre-active-treatment diagnostic records were taken for both SE and LPE groups
T2	Time point two, representing the time when final, post active-treatment records were taken for both SE and LPE groups
UNC	University of North Carolina at Chapel Hill, School of Dentistry, Department of Orthodontics

SECTION I

INTRODUCTION

The purpose of this study was to compare and contrast treatment outcomes and treatment timing factors of two treatment protocols, serial extraction (SE) and late premolar extraction (LPE). SE is a treatment modality used to correct severe tooth-size arch-length discrepancies with the early extraction of a series of primary and permanent teeth at specific dental developmental stages. Although this treatment was first described by Robert Bunon[4] in 1743, it was not widely implemented until the late 1940s after being described by Kjellgren[5] and by Hotz.[6] The alternative approach used to correct severe tooth-size arch-length discrepancies is to wait for the full eruption of the permanent dentition allowing crowding to develop before performing extractions (LPE). These late premolar extractions then are followed by comprehensive orthodontic treatment. Almost all SE patients require subsequent comprehensive orthodontic treatment as tipping of teeth, residual spacing and alignment problems commonly occur as the remaining permanent teeth erupt.

Only four published clinical studies were found that compare and contrast the effects of SE and LPE[1,7-9] and one publication[10] that reviewed several previously unpublished master's theses with varying thoroughness. Previous to these works, the only direct comparisons of SE to LPE could only be found in case reports and expert opinion articles. Many benefits were (and sometimes still are) claimed for SE over LPE, but these were supported mostly by clinical experience rather than data-based conclusions. Several of these claimed benefits may not be valid whereas others have since been supported by data as discussed in the literature review.

The goal of SE is to create space for the eruption of teeth over basal bone. It was theorized that this would lead to increased long term stability of the dentition, especially of the lower incisors. In contrast, if teeth were allowed to erupt in a severely crowded manner, the teeth would not be positioned over basal bone. LPE and orthodontic treatment to move the teeth over basal bone would somehow result in less long term stability than SE and orthodontic treatment. This hypothesis remains in question as noted by Little et al[7,11] who found no significant difference between the post-retention (minimum of ten years) Irregularity Index measures for SE and LPE groups.

Another purported benefit of SE over LPE is a shortened fixed appliance treatment time which was suggested by authors[12-15] based on their clinical experience and later supported by data-based studies.[1,7,10] Fixed orthodontic treatment time in months for SE versus LPE respectively was reported as follows: 16.8 versus 27.6 by Wagner and Berg,[1] 12 versus approximately 24 by Little et al[7] and 12.7 versus 19 by Ringenberg.[10]

Another purported benefit of SE over LPE is economic savings for the patient. Dale[12,16] suggested this twice. A definition of "costly" was not included nor is any data presented to substantiate the statement.

The PAR index[3,17] provides a summary score for a malocclusion and was used to measure occlusal outcomes in this study. The goal of this study was to test the following hypotheses comparing the SE and LPE groups:

1. There is no difference in pre-active-treatment PAR score.

2. There is no difference in final PAR score.

3. There is no difference in months of active-treatment.

4. There is no difference in active–treatment chair-time.

5. There is no difference in number of active-treatment visits.

6. There is no difference in cost to the patient.

In this study, time point zero (T0) is the time when pre-serial extraction diagnostic records were taken for the SE group. Time point one (T1) is the time when pre-active-treatment diagnostic records were taken for the SE and LPE groups. Time point two (T2) is the time when final, post-active-treatment records were taken. Another goal of this study was to determine relationships between the following variables within the SE and LPE groups:

1. PAR T0, PAR T1, and PAR T2

2. PAR T0 and treatment time measures (months of treatment, number of visits, chair-time [minutes])

3. PAR T1 and treatment time measures

4. PAR T2 and treatment time measures

5. Treatment time measures and “no-show” rates

6. Braces-free months and months followed pre-active-treatment

SECTION II

REVIEW OF LITERATURE

A. Overview

Serial extraction consists of the sequential extraction of primary and permanent teeth at specific times during the development of the dentition. The goal of this specific extraction sequence is to create space into which permanent teeth can erupt and assume a more normal alignment. SE was first mentioned in the literature in 1743 by Robert Bunon[4] and again in 1851 by Linderer,[18] although the technique was not widely implemented until 1947, when independent papers published by Kjellgren[5] and Hotz[6] received wide attention by the profession.

Hotz[6] described the basic sequence of what he originally called "active supervision of dental eruption" or "guidance of dental eruption by means of extractions." He later shortened the phrase to "guidance of eruption."[19] Although Hotz argued that his term was superior to Kjellgren's term "serial extraction," the treatment protocol and goals were almost identical and therefore the terms are synonymous. Over time, the term "serial extraction" became widely accepted whereas the term "guidance of eruption" fell from use, despite Hotz's attempt in 1970 to revive it.[19]

The treatment described by Hotz in his original article[6] consisted of the sequential extraction of primary and permanent teeth in crowded dentitions using the following criteria: 1) Eruption of the four maxillary and mandibular incisors; 2) Pont's index showing 6 or more mm of arch constriction; and 3) Radiographic visualization of developing permanent premolars and determination of their position.

Pont's index was described in a 1909 German article[20] and reviewed in a further study by Nimkarn et al.[21] The index assumed that there was a constant relationship between the sum of the maxillary incisors and the width of the dental arch for an ideal dentition. According to Hotz, if the width of the maxillary inter molar dimension was 6 or more mm narrower than the expected value, SE was indicated. Pont's index, however, has been show to be invalid by Joondeph et al,[22] Worms et al,[23] and Nimkarn et al.[21]

Despite Hotz's inability to develop valid measurable criteria, it is clear that "a marked lack of space" is an indication for SE. His treatment recommendations are best summed up by an excerpt from his article:

> "The plan of treatment is as follows:
> 1. Extraction of deciduous canines, which is generally sufficient to lead to spontaneous correction of the position of the incisors.
> 2. Premature extraction of the deciduous first molars, to provoke early eruption of the first premolars.
> 3. Extraction of the first premolars immediately upon their appearance or even before in the cases I shall mention later."[6]

The "case he mentions later" described a scenario where the maxillary second premolar had erupted before the maxillary canine. In these cases, he recommended leaving the maxillary first premolar in place "sometime longer". The goal here was to maintain space for the canine and prevent the "forward shifting of the second premolar and first molar".

Hotz emphasized that his ultimate goal in alleviating crowding with SE was to prevent, if possible, or at least reduce the amount of future orthodontic treatment that would be necessary. The prospects of "treating the masses" shines through in his introduction and his guidance of eruption treatment method was presented as a simple way for general dentists to deal with the problem of severe crowding.[6]

Although Kjellgren[5] used the term "serial extraction," his treatment protocol was very similar to Hotz's guidance of eruption. One difference is that Kjellgren recommended extracting first premolars only when they were almost completely erupted and the permanent canines were one-half erupted. Apparently he was more concerned than Hotz about the mesial drift of second primary molars, second premolars, and first permanent molars, prior to the eruption of the permanent canines. In contrast, Hotz[6] was only concerned with this in the maxilla in cases where it was obvious that maxillary second premolars were erupting ahead of maxillary canines. Kjellgren[5] mentioned that SE without active orthodontic treatment could, at times, yield similar results to LPE with active orthodontic treatment; however he did not state this to be his goal. He balanced this claim by stating that SE is "many times a suitable preliminary treatment in cases which are planned to get finally treated by appliances and it will often greatly facilitate this later treatment."

In the 1950s through 1970s, subsequent articles about SE focused on the technique as well as its benefits and shortcomings. Most of these articles consisted of case reports and clinical opinions that contained very little data or evidence supporting the statements regarding SE. Some commonly mentioned benefits of SE compared to LPE include: greater long term (post retention) stability;[12-14,24,25] limited or reduced time necessary for fixed appliance therapy;[6,12-14]and reduced damage[14] and discomfort[12] to patients.

B. Long Term Stability

Little[7,11] described a quantitative method to assess dental crowding, the "irregularity index," and using this index found no significant difference in long term stability between SE and LPE treatment. This index has been widely used since Little demonstrated its reliability and validity in 1975.[26] The SE group consisted of 30 cases treated by serial extraction of deciduous teeth and first premolars followed by comprehensive orthodontic treatment and retention. The matched control group (LPE) consisted of 30 cases treated by extraction in the permanent dentition followed by comprehensive orthodontic treatment. At ten years post retention, 22 of the 30 SE cases were considered to have "clinically unsatisfactory" lower anterior alignment. In 28 of the 30 SE cases, incisor irregularity increased from the posttreatment to the postretention stages; 8 had minimal increase (<3.5 mm), 17 had moderate increase (3.5-6.5 mm), and 3 had severe increase (>6.5 mm) in the incisor irregularity index. No statistically or clinically significant differences in the incisor irregularity index were found between the SE and LPE groups.

McReynolds and Little[9] compared the long term postretention records (minimum ten years) of 14 second premolar SE patients to 32 late second premolar extraction patients. The SE patients were either missing lower second premolars or had them extracted before eruption. A minimum of one year of physiologic drift was allowed to occur prior to the start of comprehensive orthodontic treatment. For the late premolar extraction group mandibular second premolars were extracted only after all permanent teeth anterior to the first permanent molars had erupted. All patients had comprehensive orthodontic treatment and they all had at least 2 years of retention with a fixed mandibular canine-to-canine retainer following removal of appliances. There was no statistically significant difference in the postretention irregularity index between the two groups. No correlations were found between pretreatment and posttreatment incisor alignments for either test group and "no associations were found during or after treatment between alignment and any other variable" such as mandibular intermolar width, mandibular intercanine width, and mandibular arch length. This study failed to support the clinically based concept that SE yields more stable mandibular incisors than LPE when both are followed by comprehensive orthodontics.

Woodside et al[27] compared the stability of a group of 22 serial extraction patients (most without orthodontic treatment) to a control group of untreated normal patients. However, the post retention time (the length of time between the end of treatment (T2) and the long term (T3) measurements) was not clear, because only skeletal ages were given. The minimum postretention period was 5 years with a range of 5 to 10.5 years in skeletal age. The study found no differences in incisor irregularity, crowding, intermolar width, or intercanine width between the two groups. They did find a significant decrease in arch length in the SE group which could be attributed to physiologic drift of the teeth after extraction of the premolars. This study, however, does not give much information about the efficacy of SE followed by comprehensive fixed orthodontic treatment. Most patients in North America who undergo SE also have comprehensive fixed orthodontic treatment.

Haruki and Little[28] compared two groups of patients; an "early treatment group" (n = 47) that had premolar extractions and fixed orthodontic appliance treatment to a "late treatment group" (n = 36) that had premolar extractions followed immediately by fixed appliance treatment while in the full permanent dentition. The early treatment group had either a minimum of one primary tooth present or a lack of sequential fully erupted permanent teeth when appliances were placed. The goal with this first group was to align maxillary and mandibular incisors while waiting for remaining teeth to erupt after which comprehensive orthodontic mechanics would continue. Cases that were allowed a period of physiological drift, such as with serial extraction, were excluded from this study, therefore the results of this study do not give us direct evidence about the long term stability of SE. However, the results are applicable when thinking about the grand scheme of treatment timing and long term stability. The average pretreatment age for the early group was 11 years 3 months and 13 years 4 months for the late group. Time from pretreatment to posttreatment was 3 years 2 months for the early group and 2 years 11 months for the late treatment group. Despite these small differences in timing and the fact that there was no statistically significant difference in posttreatment results, the authors found a significant difference ($p<0.01$) in mandibular irregularity index at the postretention visit (minimum of 10 years). The early treatment group showed an irregularity index of 3.09 mm while the late group showed 4.15 mm. However, there was no significant difference in the change of mandibular irregularity from pretreatment to postretention time points between early and late groups.

Haruki and Little also point out that their early treatment group showed better long term stability than the SE group from the study by Little et al[7] described. Haruki and Little[28] conclude that "perhaps the key to improved stability is early extraction plus anterior alignment, rather than early extraction followed by physiologic drift." This comparison is counter-intuitive, however interesting, and would warrant further study to directly compare the long term stability of SE and physiologic drift with that of SE and active 2x4 fixed appliance treatment to align incisors.

C. SE versus LPE: Results and Treatment Time

Although the 1990 study by Little et al [7] focused on long term lower incisor stability following SE, treatment times were mentioned briefly. They reported an average active orthodontic treatment time of 12 months for the SE group and "nearly twice that time for the late extraction treated cases". No further data was given about these active-treatment times. A conclusion drawn by the authors was that although the SE group had a shorter fixed appliance treatment time, they required a longer overall observation time.

Wagner and Berg[1] compared treatment outcome and duration of treatment between SE and LPE groups. Orthodontic records from the University of Saarland, Germany and one nearby private practice were used to generate two samples consisting of 20 patients per group. The authors used the PAR index[3,17] to measure treatment outcomes and the PAR weightings as determined by DeGuzman et al.[2] They also chose to express the occlusal improvement as percentage reduction of PAR, and reported a statistically different reduction in PAR scores between the SE group (88%) and the LPE group (77%). All cases were either improved or greatly improved according to criteria established by Richmond et al.[3,17] The ratio of greatly improved to improved cases was 60:40 for the SE group and 35:65 for the LPE group.

In their study, unfortunately, there were no post-physiological drift records for the SE group so it was not possible to compare the PAR scores of the two groups immediately prior to treatment with fixed appliances. Records were gathered at two times, before treatment when initial diagnostic records were taken (T1), and after active-treatment (T2). For both groups, the time from T1 to T2 was the "overall observation time." For the SE group, this included the period of physiological drift after extractions. Time segments measured in this study[1] included physiological drift time, fixed appliance treatment time, and total observation time. The SE group had an average of 1.4 years in fixed appliances as compared to 2.3 years for the LPE group. The total treatment time for the SE group was 6 years compared to 3.6 years for the LPE group. The average number of appointments for the SE group appeared on a bar graph to be 43 as compared to 36 for the LPE group. All of these timing differences were reported as statistically significant. Furthermore, the severity of the initial malocclusion, the initial PAR score, was not significantly correlated with treatment time, and improvement in occlusion was not significantly correlated with the duration of treatment. The authors' well stated conclusion was:

> In the serial extraction group a comparably higher reduction in PAR score was registered in spite of a markedly shorter period with fixed appliances. However, the overall duration of treatment was significantly longer and the number of appointments significantly higher.

Unfortunately, the small sample size of only 20 patients per group and the lack of diagnostic records prior to the initiation of fixed orthodontic treatment for the SE group compromised the value of their findings.

Ringenberg[10] reported on several previously unpublished masters theses. One of these by Smolen[29] compared retrospective data on 49 SE patients to that of 28 LPE patients. There were no significant inter-group differences (p=0.05) for 30 cephalometric measurements done either pretreatment or post treatment. The fact that the two groups had similar cephalometric outcomes suggests that SE and LPE have similar effects on growth and development in these measured locations. The major difference between groups was the amount of time in active-treatment, with the LPE group requiring 19 months compared to 12.7 for the SE group.

D. Serial Extraction without Later Orthodontic Treatment

SE was originally developed as an alternative to fixed orthodontic treatment and a way to treat a large number of patients with as little intervention as possible. Hotz[6] mentioned that the general dentist should be the one to diagnose and treatment plan SE without involvement of an orthodontist and explicitly stated:

> I should like to make all orthodontic treatment by means of applances [sic] unnecessary. The "guidance" of eruption is a method for the general practitioner, and particularly for the school dentist, for it is this latter who examines children at an age favourable to such procedure. The decision to undertake such treatment must be made when the child is $8^1/_2$ years of age, and it is for the school dentist to make such a decision. The specialist will be able to devote himself more exclusively to those which demand more essentially orthodontic treatment, and which have need of this experience.

Through experience and published case studies[13,15,30] practitioners found that in many cases SE improved the occlusion, but in order to achieve a more ideal orthodontic standard, fixed orthodontic treatment was required after the eruption and drifting of teeth. Despite this paradigm shift regarding the general goals and protocols for SE, there remains an indication for SE without follow-up orthodontic treatment. In areas of the world where resources or access to care are limited, SE alone, or with limited interceptive orthodontic treatment may produce occlusions that are acceptable or at least better than what would otherwise develop.

It is important to understand the effects of serial extraction on the maxilla, mandible, and dentoalveolar units following SE. Several studies have investigated the specific effects of SE on maxillofacial and dental growth and development.

E. Effects of Serial Extraction

Ringenberg[10] also reported on a previously unpublished masters thesis by Whitney[31] using lateral cephalograms to compare 51 SE patients and 23 controls before SE and 23 months later. The SE group had deciduous canine, first deciduous molar and permanent first premolar extractions whereas the control group teeth were allowed to erupt without intervention. Eighteen landmarks, 21 linear and 9 angular measurements were studied and compared. On the initial cephalometric radiographs, the groups were largely similar with the exception of two differences. The SE group had maxillary and mandibular central incisors that were 5 degrees more upright than that in controls and the mandibular canines had erupted 2.8 mm more in the SE group than in controls.

Ringenberg attributed the first difference to the fact that initial lateral cephalometric radiographs were taken after extraction of primary canines in the SE group. He attributed the second difference to the fact that the SE group was on average 10 months older than the control group when the first lateral films were taken. Following the 23-month observation period, there were no significant inter-group differences in cephalometric measurement changes regarding maxillary or mandibular growth.

As expected, there were significant differences regarding the dentition. The SE group had uprighting of maxillary incisors by 6.7 degrees and mandibular incisors by 3.4 degrees. In the SE group the maxillary canine moved 3.9 mm distally compared to 0.6 mm mesially in the control group. The mandibular canine behaved similarly but to a lesser magnitude, erupting 1.8 mm distally in the SE group and 0.2 mm mesially in the control group. During the 23-month observation period, SE treatment seemed to very slightly accelerate canine eruption in the maxilla but retard it in the mandible. Molars moved forward in the SE group twice as much in the maxilla and five times as much in the mandible as they did in the control group.

Ringenberg[10] reported on Croce's thesis,[32] which focused on the changes in overbite during the drifting phase of SE patients. This study found no difference in the distances between apices of the maxillary and mandibular incisors and no difference between vertical skeletal measurements. Although this study was only reviewed briefly in two paragraphs, the author concluded that the deepening of the bite after serial extractions was due primary to incisor uprighting rather than supereruption.

Ringenberg[10] also reviewed a thesis by Dannelly[33] that compared soft tissue profiles of 44 SE patients to that of 22 controls. The SE group showed relatively more retrusive lips following extractions, although "when treatment was carried to completion and extractions in the control group had been performed, likeness in profiles was again established in the experimental and control groups." No statistical analysis was reported.

Glauser[34] compared skeletal and dental cephalometric measures of two groups of 40 Navajo Indian patients. Group one had serial extraction with either no subsequent orthodontic treatment or "in conjunction with a simple appliance." Group two did not have serial extraction. Patients were placed into either group depending on the treatment recommendations of six orthodontists, indicating it is likely that the groups had variable amounts of tooth size arch length deficiency. Initially there were no significant differences between the groups for measured skeletal and dental variables. Forty months later there was no significant difference between groups in change of each of the skeletal variables measured. There were significant differences between the groups in all of the cephalometric dental variables at 40 months with the exception of occlusal plane angle.

These differences showed that the SE group ended up with more upright maxillary and mandibular incisors. These dental differences, however, were not considered problematic by the authors mainly because they did not lead to a deep overbite or lack of "esthetic fullness" in the lower face profile.

This same study[34] also compared group differences in the change in distance (measured on an oblique cephalometric radiograph) between the canine and first molar. In the mandibular arch, both SE and non-extraction decreased in this distance over time, but the SE group had a 4.22 mm greater decrease, of which 75% was attributed to canine movement and 25% to molar movement. In the maxillary arch, the canine to first molar distance in the SE group decreased by 2.42 mm but in the non-extraction group the distance increased by 2.33 mm due to the eruption path of the permanent canine. This 4.75 mm difference was due to 71.6% canine movement and 28.4% molar movement.

Glauser[34] measured residual spacing in the SE patients and found that in the maxilla, 7% had spaces of 1 mm or more and 1% of spaces found were larger than 2mm. In the mandibular arch 43% of cases had spaces of 1mm or more, and 14% of spaces present were more than 2 mm. It was noted that, despite these spaces, no periodontal problems or excessive tipping were detected. Although described as not excessive, SE mandibular molars tipped forward 1.36 degrees more than non-extraction mandibular molars and SE maxillary molars tipped forward 3.27 degrees more than non-extraction maxillary molars. Both of these inter-group differences were statistically significant. It is evident from this study that serial extraction can be used without comprehensive orthodontics to treat severe crowding without major negative consequences. The author makes an interesting point that "the tendency for deep overbite is very meager among the ethnic group" and that this may have "undoubtedly had a great deal to do with the outcome."[34] This study does, however lack measures such as crowding, PAR score, irregularity index, arch length, or esthetic judgments.

Persson et al[35] compared initial and long term postretention records of 44 patients who had premolar extractions in the mixed or early permanent dentition to treat crowding with an untreated normal control group consisting of 29 individuals. The study group did not consist of true serial extraction cases, in that the mean pretreatment age was 10 years 6 months and some subjects were in full permanent dentition. The study still warrants consideration as four first premolars were extracted and subsequent drifting was allowed to occur. At recall (average age 30 years 4 months) there was no significant difference in crowding between groups. After approximately 20 years, crowding of incisors in the extract-and-drift group equaled that of the untreated normal group.

The authors also did not find a statistically significant difference between groups with respect to lower incisor angulation (L1 – NB) at the long term follow up. Persson et al [35] mention that this is in contrast to the uprighting found by Ringenberg.[10] This was not a valid comparison for two reasons. First and most importantly, Ringenberg found an uprighting effect of incisors after a 23-month observation period whereas Persson measured it approximately 20 years later. Secondly, the timing of records and extractions differed. In 17 out of the total 168 segments there was incomplete closure of extraction spaces. Ten of these were in the mandible and 7 in the maxilla.

Perhaps more significant is the comparison of overall malocclusion scores [35] for Persson et al's[35] extract-and-drift group to those of a group studied by Sadowsky and Sakols.[36] The latter group was treated with comprehensive orthodontics and recalled for long term postretention records (minimum of 12 years, mean 20 years). No significant differences in initial or follow-up malocclusion were found. The orthodontically-treated group was important for two reasons. First, 66% of these patients were deemed Angle's Class II and second, they were treated by a mix of extraction and non-extraction therapies. Despite the treatment rendered in the treatment group, the occlusal outcomes 20 years later were no better than outcomes for a group of Angle's Class I crowded patients that had extractions without orthodontic treatment. It might be argued that the treated groups had better occlusions during the years immediately following treatment and might be considered a significant benefit favoring such intervention.

Papandreas et al[37] reported on a group of 32 patients who had first premolars extracted at a mean age of 10.4 years followed by an approximate 2.5-year observation period. They found that the lower incisors became more upright by an average of 1.76 degrees and the irregularity index decreased by an average of 2.16. Lower molars tipped mesially 1.9 degrees/year; the molar cusp moved mesially 1.2 mm/year and the apex 0.6 mm/year. Also, the following measures changed at these rates: the lower incisor incisal edges moved distally by 0.39 mm/year; arch depth decreased by 1.65 mm/year; intercanine width increased by 0.59 mm/year; intermolar width decreased by 0.77 mm/year; and overbite increase by 0.34 mm/year.

These 32 patients were not compared to a control group, but with an extract-and-drift group of 20 patients who had extract-and-drift treatment after the eruption of their full dentition at an average age of 14.2 years followed by a 0.8-year observation period. These patients experienced significantly faster uprighting and distal movement of lower anterior teeth. They also had faster rates of canine expansion, molar constriction, deepening of the bite, and decrease in irregularity index. The authors allude to the fact that comparing rates of dental drifting for groups that had two significantly different lengths of time for drifting to occur "assumes continuity of pattern." This was done so changes could be compared statistically. This, of course, could make the comparison between these two groups invalid if there in fact is not a continuity of pattern. It seems plausible that the highest rate of tooth drifting occurs in the first few months after extractions and could possibly slow down as extraction spaces heal.

Yoshihara et al[38] studied dental changes in the mandibular arch following SE without orthodontic treatment in a group of 31 (17 male, 14 female) Japanese patients. Initial records were taken before extraction of primary canines (T1=avg. age 8.74 years), progress records after the extraction of first premolars (T2=avg. age 11.91 years), and final records at the end of the observation period (T3=avg. age 14.73 years). Although no controls were identified, the authors found that the mandibular irregularity index decreased throughout the study and the rate of change was more pronounced at the beginning of the study.

They also found that the mandibular molars drifted forward more between T2-T3 than T1-T2, and incisors moved and tipped distally more between T1-T2 than T2-T3. Although the net mandibular molar movement was to the mesial from T2-T3, there was an uprighting effect on this tooth while drifting during this time. This seems plausible because between T1-T2 the primary canines are removed and space is present for the incisors to move distally, whereas between T2-T3 the second premolars and first molars drift mesially into the space created by the extraction of the first premolars. The eruption of mandibular second premolars also might play a role in uprighting the first molars. In this Asian population, there were different net and annual molar movements for different molar classifications. Class III molars moved the most and the fastest while Class II molars moved the least and the slowest. This investigation showed more variability than previous similar studies, but in summary the following negative (1-4) and positive (5) correlations were found: 1) annual change in canine tipping and arch length discrepancy; 2) annual change in canine tipping and annual change in the irregularity index; 3) annual change in canine movement and arch length discrepancy; 4) annual change in canine movement and annual change in the irregularity index; 5) arch length discrepancy and annual change in the irregularity index.

In a companion study, Yoshihara et al[39] studied the maxillary dental casts of 32 female subjects at similar time points. The objectives were to quantify the effects of SE without treatment on irregularity index and arch length discrepancy, and also to study the relationships between tooth width, arch length, (tooth size) arch length discrepancy, and irregularity index, all in the maxillary arch. As in their other study[40] of mandibular arches, Yoshihara et al found here that the maxillary irregularity index decreased throughout the study and more so in the beginning of the study.

Many studies have attempted to quantify a relationship between orthodontic treatment with LPE and changes in the soft tissue profile. A review of this subject is beyond the scope of this literature review. One study, however, attempted to compare the effects of SE and LPE on soft tissue profiles.

Wilson et al[8] compared three groups of extraction patients and measured changes in several cephalometric measures. Study groups were as follows:

1. Group A: SE with no treatment (n=28)

2. Group B: SE with treatment (n=30)
3. Group C: Late extraction with treatment (n=30)

Orthodontic records were taken pre-extraction, post-extraction and late post-extraction for group A; pre-extraction, post-extraction before appliances, posttreatment, and postretention for group B; and pretreatment, posttreatment, and postretention for group C. The only significant difference found among groups was the posttreatment position of the mandibular incisors in the late extraction group and the SE with treatment group. The labial point of the mandibular incisors was positioned significantly more to the lingual in Group C. No differences were found between the groups for any of the soft tissue measurements.

Kennedy et al[41] compared root resorption and alveolar bone heights for three groups: a late first premolar extraction group treated with a conventional banded edgewise appliance, a serial extraction group treated with the same orthodontic treatment, and a serial extraction group that had no subsequent appliance therapy. They found that the last group had significantly less root resorption than the other two. Also, the two groups that underwent appliance therapy had more apically positioned bone heights on the distal of the canine than that for the SE without treatment group suggesting possible alveolar bone loss.

F. Economics of Serial Extraction

As previously mentioned, Dale briefly mentioned that SE may be less costly to the patient than LPE.[12,16] No reasoning was given nor has this suggestion ever been substantiated by data. In fact there is a general lack of published cost-effectiveness research within the specialty of orthodontics. In 2000, Richmond[42] made a call for this type of research in orthodontics and editorialized:

> The assessment of clinical performance is important at the individual, practice, institutional and national levels. It is a challenge not only to deliver high standards of care, but also to deliver the care at the lowest unit cost.

Cost-effectiveness, in medicine as well as dentistry, is simply about the accountability of inputs and outcomes. In other words, what resources are spent on treatment (i.e. money, time etc), what is gained from treatment (i.e. health, occlusion etc.) and how do these compare to alternative inputs and outputs? In 2004, Richmond[43] wrote:

> Cost-effectiveness is one of the techniques of economic evaluation, which involves assessing the outputs and outcomes of orthodontic care, relative to the level of inputs and need to arrive at an indicator of the relative efficiency of orthodontic care.

Assessing costs for calculating cost-effectiveness analyses is difficult in orthodontics for several reasons. First, there are almost unlimited combinations of fee structures and payment plans, which can vary greatly between practitioners. Second, there are myriad orthodontic treatment techniques and mechanics used in practice today. Between and among these techniques there are many different timing protocols.

Regardless of the difficulties, it is important to know the perspective when measuring cost-effectiveness. Most discussions regarding the economics of treatment modalities in orthodontics focus on the perspective of the orthodontist. Often times there are comparisons of these costs in order to calculate which treatment is most efficient and reduces operating costs. Very seldom are these analyses discussed from the perspective of the patient. Although sometimes what is less expensive for the orthodontist may be less expensive for the patient, this is not necessarily true. There are three types of costs that can be considered: direct costs, indirect costs, and intangibles. According to

Richmond et al[43] direct costs include material costs, pharmaceutical costs, costs of staff time, and "transport costs and out-of-pocket expenses" paid by other organizations and patients and families being treated. Indirect costs are more difficult to measure and include "losses to society incurred as a result of receiving the treatment such as, loss of production, education, domestic responsibilities, social and leisure activities." The intangibles are the most difficult to measure and include things such as pain and suffering, anxiety, and quality of life.

Assessing the outcomes for cost-effectiveness analyses is very difficult in orthodontics because of the many factors and variables that are affected. Orthodontics presents a wide range of opinions on what severity of malocclusion needs treatment, and what constitutes an acceptable orthodontic treatment outcome. Just as problematic is the question of how to quantify these malocclusions and results. Also, improvement of a patient's occlusion occurs very slowly over a long period of time, during which (in most orthodontic practices) few measurements documenting improvement are taken. This makes it difficult to establish specific improvement intervals. In other words, it is difficult to establish what amount of improvement is occurring during a given time span. The various reasons stated for the difficulty of performing cost-effectiveness analyses in orthodontics should be a motivating factor for, not a deterrent to, future research of this type.

Despite the difficulties, there have been several widely used tools to objectively measure treatment changes in orthodontics. One valid and reliable test for measuring the overall state of occlusion is the Peer Assessment Rating index.[3,17] It is a summary numerical score that represents occlusal anomalies, and deviations from normal, in a given patient. It is based upon assessments of different aspects of the patient's dental occlusion and alignment including: maxillary and mandibular anterior alignment, overjet, overbite, midlines, and buccal occlusion. A higher score represents a more severe malocclusion. A publication by Richmond et al[17] determined, using a panel of 74 orthodontists, that a minimum PAR score reduction of 30 percent was required for a case to be considered "improved" (on a scale of "Worse-no different," "improved," and "greatly improved"). It would seem that PAR index scores could be useful in studying cost-effectiveness of different orthodontic treatments. Richmond, in his well reasoned 2000 editorial,[42] warned against the use of percentage PAR reduction for cost-effectiveness analyses, stating:

> …the use of percentage reduction is questionable both scientifically and statistically in assessing cost-effectiveness. For instance, a PAR score change from 50 to 5 (case 1) and 10 to 1 (case 2) both represent a 90 per cent reduction in PAR score. However, case 1 showed a change of 45 PAR points and case 2 only 9 PAR points. If both treatments cost 90 [dollars], using cost per PAR reduction the cost effectiveness would be 1 [dollar] per percentage reduction in PAR score. This would not represent the effectiveness of treatment and [is] arguably inappropriate.[42]

A better but not perfect approach is to measure effectiveness by absolute PAR reduction. Breaking down the overall cost of treatment in terms of cost per patient visit was also recommended in Richmond's editorial. This cost per visit may or may not be equal to the cost per unit reduction in PAR, which depends on many different factors including the type of healthcare system and the standards in a given area where treatment is rendered.

SECTION III

MATERIALS AND METHODS

A. General Study Design

Post orthodontic treatment review forms from a multi-office, two-doctor private practice in Vancouver, Canada were screened using general inclusion criteria described below to identify suitable serial extraction and late premolar extraction cases. All patients consecutively treated with fixed appliances between January, 1990 and March, 2006 were screened for acceptability. The post treatment review forms included a one page summary of the diagnosis and treatment rendered for each individual patient and included: Angle's molar classification, presence of crossbites, extractions performed, timing of extractions, appliances used, and congenitally missing teeth. If the case initially passed the general inclusion criteria after screening these forms, treatment charts and plaster models were located and examined further to confirm or deny inclusion in the study. General inclusion criteria, general exclusion criteria, serial extraction criteria, and late premolar extraction criteria were established to select and differentiate the SE and LPE samples. The general inclusion criteria were: four permanent dental units extracted, one in each quadrant; full fixed appliance treatment following extractions; and complete pre and post treatment records available. The general exclusion criteria were: more than one tooth in posterior crossbite; Angle Class II molar relationship beyond ½ cusp Class II; Angle Class III molar relationship beyond ¼ cusp Class III; adult patients 21 years or older; missing teeth (other than 3rd molars); previous orthodontic treatment; incomplete records; or active Phase I orthodontic tooth movement (passive Nance arches and lower lingual holding arches were acceptable). The SE group inclusion criteria were: premolars extracted either before the permanent canines erupted or enucleated if the canines erupted first and a minimum of one year of physiological drift following extractions. The LPE group inclusion criteria were: patient in full permanent dentition (partially or fully erupted) and fixed appliances placed no longer than 3 months after extractions. Stringently employing these criteria, 51 SE and 49 LPE patients were identified who satisfied these criteria and were included in the study.

Records of SE patients included a complete treatment chart as well as plaster models and panoramic radiographs taken at the following times:

1. just prior to the extraction or enucleation of first premolars (T0).
2. just prior to the initiation of fixed orthodontic treatment (T1).
3. after the completion of fixed orthodontic treatment (T2).

Records of LPE patients included a complete treatment chart as well as plaster models and panoramic radiographs taken at the following times:

1. just prior to the initiation of fixed orthodontic treatment (T1).
2. after the completion of fixed orthodontic treatment (T2).

The study protocol assumed that all LPE patients would have been prescribed serial extractions if they had been referred to the practice early enough. The practitioners often prescribed serial extractions in cases with severe crowding in the mixed dentition after considering dental development and confirming a lack of extreme anterior-posterior and vertical skeletal discrepancies. However, not all patients were referred to the practice early enough in their dental development for the initiation of serial extraction treatment. These patients who were referred late but still had severe crowding were prescribed LPE. It was not possible to match the SE and LPE groups on initial malocclusion or crowding because the LPE group did not have initial records taken at the same age as the SE group. SE and LPE records at T1 did not allow matching, because at this point the SE group had already had extractions and drifting of teeth while the LPE group had received no treatment.

The two orthodontists providing treatment did not routinely perform a mixed dentition analysis such as the Tanaka Johnston or Moyers analyses. They did, however, state that the majority of cases that received serial extractions were estimated to have at least 8 mm of crowding by visual estimation. Other cases also may have received SE because of a combination of crowding and dento-alveolar protrusion. Cases with severe anterior-posterior discrepancies, decreased vertical face heights, and excessive deep bites were not considered suitable candidates for SE. Since these vertical dimension criteria were not quantifiable with the available records, they were not used in the inclusion and exclusion criteria.

Beyond the inclusion and exclusion criteria given previously, the equality of SE and LPE groups was based on the assumption that the practitioners used consistent treatment planning criteria in deciding when to extract teeth, regardless of the dental development of the patient. This seems to be a reasonable assumption.

The SE group average age at T0 was 7.58 years (standard deviation [SD] 2.09) and contained 19 male and 32 female patients. The LPE group average age at T1 was 11.95 years (SD 3.87) and contained 19 male and 30 female patients. Panoramic radiographs were examined to determine the presence of all permanent teeth other than third molars. Dental ages were determined according to the method of Demirjian,[44] based on measurement of root development of all teeth in the lower left quadrant (minus third molars) on a panoramic radiograph. Each tooth root is given a score and these scores are converted and summed to give a composite dental age. This was done to compare dental ages of the groups at T1. All scoring of films was done by a single examiner (the author) and thirteen panoramic films were chosen randomly for re-measurement to test for intra-examiner reliability.

PAR index scores were measured and calculated for SE models at T0, T1, and T2 and for LPE models at T1 and T2. All measurements were made by one examiner (the author) who was calibrated to the training set housed within the UNC, School of Dentistry, Department of Orthodontics. Intra-examiner reliability was first tested on a PAR reliability training set within the same institution. In addition, from the models used in the study, thirty-eight sets were chosen at random for re-measurement.

PAR index scores were first calculated using the original weighting system developed by Richmond et al.[3] These PAR scores were used for determining correlations with the treatment timing variables (see below paragraph). PAR index scores were also calculated using alternative weightings validated by DeGuzman et al[2] for comparison with a study done by Wagner and Berg.[1]

Patient treatment charts were reviewed by one examiner (the author) and the following information for each patient was entered into an Excel (Microsoft Office Excel 2003, Microsoft Corporation, Redmond, WA) spreadsheet: patient's date of birth; each date of visit to the orthodontist; length (minutes) of chair-time scheduled for each orthodontic visit; type of procedure scheduled at each orthodontic visit, i.e. exam, recall, etc.; tooth extraction referrals written, if any, at each visit. Because return letters from general dentists and oral surgeons confirming extractions were not present in most charts, an assumption was made for every patient that extractions were performed the day the referral letter was written.

For the SE group, means were calculated for pre-active-treatment and active-treatment months of treatment, number of visits to the orthodontist, and minutes of chair-time. For the LPE group, means were calculated similarly with the exception of months of pre-active-treatment. For this measurement, median was used because of the presence of several outliers. These were patients who came to the practice for an initial exam and did not return as scheduled for records or a consultation for several years.

Race of the patient was not recorded because this information was not explicitly recorded in the chart. It was evident from the patient records and last names, however, that the sample represented a racially diverse group, including Asian, African and Caucasian patients.

B. Statistical Analyses

Unpaired t-tests and Wilcoxon signed rank tests were used to compare the means of the two treatment groups for parametric and non-parametric data, respectively. For analysis of categorical data, a Fisher's exact test was used. To assess the bivariate relationship between variables, Pearson correlation and Spearman correlation were used for parametric and non-parametric data respectively and multiple linear regression with pairwise interaction for treatment group and the selected explanatory variable of interest was used to examine whether the bivariate relationships were the same in the two treatment groups and to compare the adjusted means of the two treatment groups after controlling for the effect of the explanatory variable. Level of significance was set at 0.05 for all analyses. Decision modeling with deterministic and probabilistic sensitivity analysis was employed in the cost-effectiveness analysis. This model is described below in detail.

C. Cost-effectiveness Model

1. Overview of the Model

A decision model was constructed to evaluate the cost-effectiveness of serial premolar extraction (SE) compared to late premolar extraction (LPE) for the data collected from the sample previously described. To the extent possible, we followed the recommendations of Gold et al[45] on economic modeling practices. A hypothetical cohort of children age 8 years of age with the inclusion and exclusion criteria previously described was run through the model; patients were followed for a total of five years through removal of all fixed appliances. Individuals who entered the model were treated with either SE or LPE.

2. Data Source

Data gathered from the 51 SE and 49 LPE cases studied was analyzed, specifically the number of pre-active-treatment observation visits and months of active-treatment time (months in fixed appliances). Treatment fees were derived from published fee schedules.[46,47] Base-case parameters are presented in Table 1 along with ranges tested in sensitivity analyses.

3. Clinical Scenarios

To estimate costs and outcomes for the serial extraction alternative, we developed three different treatment fee structure scenarios that may be used by orthodontists in private practice. These fee structure scenarios include: “SE-Interceptive”, where one interceptive fee is charged for the SE period (high extreme fee); “SE-Observation”, where a per-visit fee is charged during the observation period (moderate fee); and “SE-Discounted”, where no interceptive or observation fees were charged during observation and a “discounted” active-treatment fee is given (low extreme). In this model the discount of the active-treatment fee in the third fee structure scenario is given because of the orthodontist’s supposition that this period of treatment will be easier and/or faster. In the SE-Interceptive fee scenario, patients are charged a fixed interceptive fee for the SE period for pre-extraction planning and follow-up during the extraction and drifting process; we assumed that this fee would be spread across the SE period. In the per visit scenario (SE-Observation fee structure), patients are charged a fee for each of the visits made during the SE period (X=8.42, SD=1.82). This measurement of 8.42 extra pre-active-treatment observation visits for the SE group is based on the assumption in the model that the LPE group had only three pre-active treatment appointments; initial exam, records, and treatment consult. In the final SE scenario (SE-Discounted), patients are charged nothing except extraction fees during the SE period, and the orthodontist “discounts” the active-treatment fee because of the anticipation that it will be easier or shorter in duration. For all three SE scenarios, pre-SE diagnostic procedures and records were required. ADA codes from the 1999 Survey of Dental Fees[46] were used to calculate a sum for a diagnostic records fee; they were:

1. comprehensive oral exam- 00150

2. diagnostic casts- 00470
3. diagnostic photographs- 00471
4. panoramic film- 00330
5. lateral cephalometric film (no fee code is listed in the survey, therefore the fees given with those for a panoramic film were used)

For SE scenarios, these procedures are repeated just prior to installation of fixed appliances; for the LPE scenario, these procedures occur once. In this study, data collection was not recorded past the removal of fixed appliances. Since many cases had final records taken at a separate visit after removal of fixed appliances, fees for final records were not used in the calculations.

4. Costs

The analysis takes the dental care system perspective, which includes the direct dental and orthodontic costs for all parties. In this model, all costs were adjusted to 2006 US dollars using the Consumer Price Index (CPI) to account for inflation -1999-2006: 1.21.[48] Cost estimates are presented in Table 1. The general CPI was used rather than the medical care component of the CPI because the increase in published orthodontic fees between 1999 and 2006 followed the former more closely than the latter. According to the 1999 Survey of Dental Fees by the American Dental Association,[46] the United States national average fee for comprehensive orthodontic treatment of the adolescent dentition was $3890. According to the general CPI, $3890 in 1999 had the same buying power as $4707 in 2006.[48] In 2006 a JCO orthodontic practice study[47] reported that the United States average comprehensive orthodontic treatment fee was $4700. In contrast, the medical care component of the CPI[48] shows that $3890 in 1999 was worth the same as

$5162.03 in 2006. Since the fee for a comprehensive orthodontic treatment case is usually considered a benchmark for how orthodontic fees in general increase with inflation, the general CPI was used to inflate all of the individual dental and orthodontic fees from 1999 dollars to 2006 dollars.

For patients having SE, extraction of teeth was assumed to occur at ages 8.5, 10.5, and 12 years, with four teeth removed at each time point. In the case of LPE, we assumed that four teeth would be removed at age 12 years. It is important that each hypothetical patient in this model was analyzed from the same beginning time point of 8 years of age. The extraction fee for each set of teeth is derived as the cost of extracting a single tooth (ADA code 07110), plus three times the cost of extracting an additional tooth (ADA Code 07120). Extraction fee data from general practitioners was used. The costs were derived from the 1999 Survey of Dental Fees;[46] the 50th percentile of the fee was used for the base-case analysis (Table 1), with the 25th and 95th percentiles used in the sensitivity analysis. In the case of the observation fee, we used fee estimates for ADA code 08690 (Orthodontic treatment, alternative billing to a contract fee).

5. Outcomes

The outcome used in the model was time spent without appliances (i.e., braces-free months [BFM]). Time in appliances was calculated as the time between placement of separators and the final removal of bands and brackets from teeth as defined previously. BFMs were derived as the differences in time with appliances between two groups.

6. Analysis

For each treatment fee structure scenario, the model calculated time in appliances and costs during the study time horizon. The study compared the performance of the two treatments using an incremental cost-effectiveness ratio (ICER), defined as the additional costs for the more expensive treatment divided by the reduction in time spent in fixed appliances associated with that treatment (i.e., braces-free months). The three serial extraction fee scenarios (i.e., SE-Interceptive, SE-Observation, and SE-Discounted) were compared to late premolar extraction individually to estimate which yielded the best value for money.

One-way sensitivity analyses were conducted to assess the effect on the ICER of varying individual baseline estimates within plausible ranges; tornado diagrams are used to present these results. Probabilistic sensitivity analysis was conducted by simultaneously varying parameters over predefined probability distributions. Costs were approximated by lognormal distributions, the number of observation visits by the Poisson distribution, and time in fixed appliances using the normal distribution* (Table 18).[49] Values from each probability distribution were randomly selected during each of 1,000 Monte Carlo iterations. These values were then used to calculate the costs and outcomes associated with each of the scenarios and the incremental cost-effectiveness ratios comparing each SE strategy to LPE; simulated cost and outcome values also were used to calculate net benefit.

The incremental cost-effectiveness ratios (ICERs) first are presented in an ICER plane (Figure19) in which the incremental difference in costs (i.e., the additional costs for each SE strategy compared to the LPE strategy) are plotted against the difference in outcomes (i.e., in this case the time in fixed appliances or braces-free month).[50] The upper right quadrant represents a tradeoff between costs and improved outcomes; in this quadrant the SE strategies cost more but result in less time in fixed appliances. In this

* Distribution fit to raw data using Crystal Ball distribution fitting macro.

case, whether one selects the SE strategy will depend on how much one is willing to pay for an additional month without braces (i.e., the threshold ratio or λ). The lower right quadrant represent situations in which the SE strategy dominates LPE (i.e., SE costs less and results in earlier removal of fixed appliances). The upper left-hand quadrant represents situations in which the SE strategy is dominated because LPE costs less and results in earlier removal of fixed appliances). In these two quadrants it is most appropriate to implement the strategy that costs less and provides a better outcome (in this case, less time in fixed appliances.)

We also calculated cost-effectiveness acceptability curves (CEAC) comparing each of the four strategies simultaneously (Figure 20). The proportion of times that a strategy is cost-effective (i.e., preferred) is plotted against various willingness to pay thresholds (i.e., how much the decision maker is willing to pay to avoid an additional month with fixed appliances). The CEAC for this project was calculated using a net-monetary benefit frame work[49,51] to address the fact that ICERs with the same sign can have very different meanings. With a CEAC, one can compare multiple interventions to determine the probability that an intervention is preferred at a given willingness-to-pay threshold, and to simultaneously compare interventions. The highest curve for a given willingness to pay (in this case $ per BFM) is the most cost-effective strategy. Such methods are commonly used to represent uncertainty in economic evaluations of health care technologies and to assist policy-making decisions.[49,52]

The decision model was constructed using Microsoft Excel 2002 version (Microsoft Corporation, Redmond, WA), and all sensitivity analyses were conducted using Crystal Ball 7.1.2 (Desicioneering Inc., Denver, CO). All costs and outcomes were discounted at

3% per annum (Gold et al[45]), with sensitivity analyses conducted using discount rates ranging from 0% to 10%.

SECTION IV

RESULTS

A. Reliability

Intra-examiner reliability for measurement of the dental ages yielded an r-score of 1.0. Results were identical for the second set of measurements likely because the scoring rules set by Demirjian[44] are concrete and easily followed with little room for interpretation.

PAR scoring intra-examiner reliability calculated from the training set housed within UNC was r= 0.924. In addition, PAR scoring reliability testing done on the models from the study yielded and r-score of r=0.967.

There was no significant difference in dental age (according to Demirjian's[44] method) at T1 (pre-active-treatment) between the SE (14.02 years) and LPE (13.97 years) groups (p=.8366). Therefore, the effect of dental development on treatment time variables (i.e. months of active-treatment, chair-time in active-treatment, and number of visits in active-treatment) for the active-treatment segment should be equal in both the SE and LPE groups.

Table 2 and Figures 1-2 show the weighted PAR index score measurements with weightings validated by Richmond et al.[3] There was no significant difference in the final PAR scores between SE and LPE groups suggesting that the final occlusal outcomes were similar in each group. However, there was a statistically significant difference in PAR scores at T1 (p <.0001) with the LPE group having a higher mean PAR score than the SE group. There was also a significant difference in percentage PAR reduction between T1 and T2 between the groups ($p=0.0095$). Figure 2 shows PAR scores as a function of time for the groups throughout the study.

B. Categorical Outcomes Based on Reduction of PAR Score

Figure 3 is a nomogram plotting the PAR score before treatment (T0 for SE and T1 for LPE) against PAR after treatment (T2 for both groups). The two lines on the graph divide the data points into three categories of change in PAR score; worse or no different, improved, and greatly improved. An individual case that has less than a 30% reduction in PAR after treatment is categorized in the worse or no different category. If a case experiences 30% or more reduction in PAR it is categorized as improved; unless it has a 22 or greater absolute point reduction whereby it is categorized as greatly improved.[17] For the SE group, 0 cases finished worse or no different, 41 cases (80%) were improved, and the remaining 10 cases (20%) were greatly improved. For the LPE group, 0 cases finished worse or no different, 17 cases (35%) were improved, and the remaining 32 cases (65%) were greatly improved (Table 3). These categorical outcomes are significantly different between groups ($p<0.001$).

C. Categorical Outcomes Based on Final PAR Score

Figure 3 and Table 4 show that for the SE group, 48/51 patients had a final PAR score of 5 or below, placing them in the "almost ideal" occlusion category. The remaining 3/51 patients had a PAR score of between 6 and 10, placing them in the "acceptable" occlusion category. For the LPE group, 44/49 patients had a final PAR score of 5 or less ("almost ideal") and 5/49 finished with a PAR between 6 and 10 ("acceptable.") None of the patients in either group finished in the "less acceptable" occlusion category. These categorical outcomes of final PAR score were not significantly different between groups (p= 0.483). It is important that these categories are for absolute final PAR score only and do not take the change in PAR score into account. These categories of post treatment PAR score were first described by Richmond et al[3] and were used by Tulloch et al[53] in the UNC Class II clinic trial.

Figures 5 and 6 are analogues to Figures 3 and 4, the difference being that PAR scores were calculated with weightings validated by Deguzman et al[2] in order to compare them to a previous study by Wagner and Berg.[1] These weightings changed several outcomes but only slightly.

Component PAR index scores for the anterior segments in Table 7 and Figure 7 compare the incisor irregularity segment of PAR for the SE and LPE groups at each time point during the study. A major reason for the extraction of teeth, either SE or LPE, is the elimination of anterior irregularity which often accompanies tooth-size arch-length discrepancies. There was no significant difference in final upper or lower anterior PAR index scores for SE and LPE groups (p= .8412 for upper, p=.7391 for lower). As was the case for total weighted PAR, there was a significant difference (p<0.001) in T1 mean upper and lower anterior PAR scores for SE and LPE groups.

Un-weighted raw PAR index score measurements for each individual component (i.e. overjet, overbite etc.) at pretreatment and posttreatment records are shown in Table 8 and Figures 8-11. These are shown for quantitative and visual comparison with a previous study by Wagner and Berg.[1] Statistical comparisons between studies are not possible because the previous study[4] did not include standard deviations of the individual component PAR scores.

D. Treatment Timing Factors

Table 9 and Figure 12 show three treatment timing factors (months of treatment, numbers of visits, and minutes of chair-time) for each test group. Treatment timing measurements are listed for both the total treatment time as well as active treatment time. Between the SE and LPE groups, there was a statistically significant difference between the mean values ($p < 0.05$) for the following timing factors:

1. total treatment time (months)
2. active-treatment time (months)
3. total number of visits
4. active-treatment number of visits
5. active-treatment chair-time (minutes).

The only treatment timing factor which did not show a significant difference between groups was mean total treatment chair-time (minutes).

The SE group was followed for a median span of 62.5 months before active-treatment was begun while the LPE group was followed for a median of 4.2 months for a difference of 58.2 months. The SE group also had an average of 6.4 more pre-active-treatment appointments and 123 more minutes of pre-active-treatment chair time than the LPE group (Table 10). During this pre-active-treatment time, for the SE group, teeth were extracted and drifting and eruption occurred. The SE group had an average of 31.9 months (SD 8.89, range 14-51 months) of physiological drift time between the extraction or enucleation of premolars and the initiation of active orthodontic treatment.

For the LPE group, pre-active-treatment time represented the time between initial exam and the initiation of orthodontic treatment and included one appointment for each of the following tasks; initial exam, diagnostic records, and patient consult. Median was used rather than mean to represent the months following pre-active-treatment because of the presence of several outliers in the LPE group that skewed the mean. While the SE group was followed for a much longer time period before active-treatment, they enjoyed a mean of 4.24 fewer months of active-treatment compared to the LPE group. This benefit of SE compared to LPE is given the term "braces-free months."

E. Correlation Between Variables

Correlation between the following PAR scores was investigated using Spearman correlations (Table11):

1. PAR T0 with PAR T1 for the SE group
2. PAR T0 with PAR T2 for the SE group
3. PAR T1 with PAR T2 for the LPE group.

The only variables with significant correlation were PAR at T0 and PAR at T1 for the SE group (p=.0010).

Correlation between PAR at T0 and the following variables for the SE group were investigated using Spearman correlation (Table 12):

1. months of active-treatment
2. total number of visits
3. number of visits in active-treatment
4. total chair-time (minutes)

None of these above were significantly correlated with PAR at T0.

Linear regression models were developed to test the association of PAR at T1 with the following variables between the SE and LPE groups: (Tables 13 and 14, Figures 13-15):

1. months of active-treatment
2. number of visits in active-treatment
3. active-treatment chair-time

The slope of the relationship between PAR T1 and months of active-treatment differed significantly for the two treatment groups ($p=.0014$) (Table 14). For the SE group, the estimated change in months of active-treatment was 0.27 months as PAR T1 increased by 1 unit (Table 14, Figure 13). For the LPE group, the estimated change in months of active-treatment was approximately 0 months as PAR T1 increased by 1 unit (Table 14, Figure 13). After controlling for the effect of PAR T1 on months of active-treatment, there was a statistically significant difference in months of active-treatment between the SE and LPE groups ($p=.0003$).

The slope of the relationship between PAR T1 and number of visits in active-treatment differed significantly for the two treatment groups ($p=.0305$) (Table 14, Figure 14). For the SE group, the estimated change in number of visits in active-treatment was 0.11 as PAR T1 increased by 1 unit. For the LPE group, the estimated change in number of visits in active-treatment was approximately 0 as PAR T1 increased by 1. After controlling for the effect of PAR T1 on number of visits in active-treatment, there was a statistically significant difference in number of visits in active-treatment between the SE and LPE groups ($p<.0001$).

PAR T1 was not significantly correlated (p-value= .8500) with active-treatment chair-time (minutes) (Table 14, Figure 15).

Linear regression models were developed to investigate the association of PAR T2 with the following variables (Table 15):

1. months of active-treatment
2. total months of treatment
3. number of visits in active-treatment
4. total number of visits
5. total chair-time (minutes)
6. active-treatment chair-time (minutes)

The threshold for a significant association was set at $r>0.30$ and $p\leq 0.05$. None of the above were significantly correlated (Table 15).

Pearson correlation coefficients were calculated to determine the strength of association between % PAR reduction and following variables for both groups (Table 16):

1. Total months of treatment with % PAR reduction
2. Months of active-treatment with % PAR reduction
3. Total number of visits with % PAR reduction
4. Number of active-treatment visits with % PAR reduction

The threshold for a significant association was set at $r>0.30$ and $p\leq 0.05$. None of the above were significantly correlated.

The Pearson correlation coefficient was also calculated to determine the strength of association between months of drift (driftodontics) and braces-free months for the SE group (Table 17). For this pair of variables to be considered significantly correlated, significance was set at $r>0.30$ and $p\leq 0.05$. An r-score of 0.20 was calculated therefore they are not significantly correlated.

There was no significant difference for no-show rate of SE and LPE groups during the total treatment time span ($p= .6715$) or the active-treatment time span ($p=.6635$) (Wilcoxon rank test).

F. Cost-effectiveness Model Results

1. Base-Case Results

Compared to late premolar extraction, each of the three serial extraction approaches costs more- $1,639 for the SE-Interceptive fee scenario, $887 for the SE-Observation fee scenario, and $443.40 for the SE-Discounted fee scenario- but results in a shorter time period with fixed appliances (Table 18). Thus, compared to LPE, these strategies cost $403.73, $218.71, and $109.21 per BFM respectively.

2. Sensitivity Analysis Results

One-way sensitivity analyses are shown visually in a series of tornado charts (Figures 16-18). In one-way sensitivity analyses, we varied each of our input parameters one at a time across the range described in Table 1. For the comparison of the SE-Interceptive fee strategy with LPE, the cost of the interceptive fee is the primary driver of the ICER with pre-SE diagnostic records fees and the extraction fees also influencing the conclusion that would be drawn. In the case of SE-Observation fee versus LPE comparison, the conclusions drawn are driven by the similar sorts of fees, but not the per visit observation fee. Likewise, the same costs are driving the model for the SE-Discounted fee versus LPE. The other parameters had relatively little effect in comparison with these parameters and are not shown in the charts. The effect of the discount for the active-treatment fee (given because of the supposed easier and or shorter treatment) has little effect.

Figure 19 presents the ICER plane for the probabilistic sensitivity analyses. Compared to each of the three SE approaches, LPE dominates (i.e., costs less and results in shorter time in fixed appliances) approximately 21% of the time; that is, LPE is the preferred treatment. The remainder of the time the particular SE scenario is cost-effective but whether one would choose it depends upon how much one is willing to pay to have an additional BFM. The lines in the ICER plane show the thresholds for $100/BFM and $300/BFM. For example, the number of points below the threshold divided by 1000 tells you what proportion of the time the particular SE strategy is cost-effective compared to LPE if one is willing to pay up to $100 to avoid an additional month with braces.

Figure 20, the cost-effectiveness acceptability curve, uses a net-benefit framework (benefits net of costs--where one is looking for the strategy that has the greatest net benefits) to compare the four strategies simultaneously depending upon how much the decision maker is willing to pay to have an additional braces-free month. LPE is superior to each of the three SE scenarios across reasonable willingness to pay thresholds; even when the threshold is set to more than $100,000 per BFM, LPE remains cost-effective or preferred 71% of the time; even at maximal willingness to pay levels, SE-Observation Fee and SE-Discounted are preferred only 18.6% and 8.2% of the time.

SECTION V

DISCUSSION

A. General Discussion

The orthodontists in the private practice sampled consisted of one senior doctor and one junior associate. For the first 3 to 6 months after the junior doctor joined the practice there was joint treatment planning of all cases. For the following approximate 15 months there was collaboration between doctors regarding treatment planning of any cases that were deemed complex or borderline. There is no data available to indicate consistency in treatment planning after this point. Matching of SE and LPE groups based on initial malocclusion was not possible because the LPE group did not have initial records taken at the same age as the SE group. This may have created residual selection bias and it may have been possible to account for this statistically.

All of the SE cases studied had extractions followed by an average of 31.9 months of eruption which allowed drifting of the remaining dentition into the extraction spaces. This phenomenon is commonly referred to as "driftodontics"[37] and in this study the duration ranged in time from 14-51 months. In the sample there were variations from the classic serial extraction pattern. Some patients presented with one or more exfoliated primary teeth; for example, one or more missing primary canines. Extractions were prescribed to suit each individual patient and there were many different combinations of extractions and timing of these extractions. The average dental age for the SE group at T0 was 10.95 with a small SD of 0.88 indicating similarity in dental developmental age at the initiation of extractions. Dental age is more appropriate than chronological age due to the large variation in dental development relative to age in years from patient to patient. It is important to note that Demirjian's[44] method is based on root development and not eruption of the crown. Nevertheless, this was a way to confirm that patients in the study were of similar dental developmental age.

The PAR index scoring system was chosen for use in this study because it is a reliable and valid[3] way of representing the overall state of a patient's occlusion. Although tooth-size arch-length discrepancy, i.e. crowding, is the specific occlusal component that is targeted by SE or LPE treatment, there are other aspects of occlusion that must be managed with each case. Therefore, an index that measured the total state of each occlusion was chosen. Nevertheless, since anterior crowding and irregularity are important in these samples, the upper and lower anterior segment of the PAR index (measuring the contact point displacements of the contacts from mesial of canine to mesial of canine) were recorded and analyzed separately.

One possible shortcoming of the use of the PAR index in this study is the way it measured anterior contact point displacements in the mixed dentition. The convention of not recording contact points of primary anterior teeth was followed. The following discussion applies to the SE group because none of the SE cases had permanent canines erupted at T1. If there is no permanent canine erupted, by convention, there is no contact point displacement possible for the distal of the permanent lateral incisor. There is, however, a scenario in the PAR index to account for a severe lack of space for unerupted canines, first premolars, and second premolars. If permanent canines, first premolars, or second premolars are unerupted, they are assigned an average width, in mm. These are 8, 7, 7, 7, 7, 7 mm for the upper canine, first premolar, second premolar, and lower canine, first premolar, and second premolar respectively. In the maxilla, for example, the total of these upper three unerupted teeth is 22 mm and if the space available between the mesial of the permanent first molar and the distal of the lateral incisor is less than or equal to 18 mm, the canine is considered impacted and is given a score of 5. If, however the space is 19mm, it does not receive a score of 5 and it may receive a score of 0. Importantly, it was very common in the SE group to find a permanent lateral incisor with a large rotation or displacement and therefore a significant contact point discrepancy with a primary canine (which is not recorded, by convention); despite this significant discrepancy, if spacing was 19mm or more between the distal of the lateral incisor and the mesial of the first molar a score of 0 was recorded. This "all-or-nothing" scoring system for unerupted permanent canines in the SE group may not be comparable to contact point displacements measured between the permanent lateral incisors and permanent canines as in LPE cases. In future studies where upper or lower anterior crowding or irregularity is important in mixed dentition cases, a modification of PAR might be used where contact points between permanent lateral incisors and primary canines are recorded. Despite these shortcomings in the index, total weighted PAR scores were calculated and analyzed because this index is widely used in the orthodontic literature and well known as an overall measure of a patient's state of occlusion.

One reason for using the PAR index was to repeat the previous study by Wagner and Berg.[1] The authors report that there was no statistically significant difference in pre treatment PAR scores for the SE and LPE groups but a significantly higher final PAR score for the LPE group than for the SE group ($0.001<p<0.01$). This is in contrast to the present study which found a significant difference in pre treatment PAR scores (p<0.0001) and no significant difference in final PAR scores (p=0.2652). The present study found that all cases in each group were improved or greatly improved by treatment. Wagner and Berg[1] found this as well, however the percentages of patients falling into these categories was different in their study (Table 5, Figure 6). When comparing these two studies it is interesting that the ratios of improved to greatly improved occlusions are somewhat reversed for both groups between studies. In the present study, using the weightings validated by DeGuzman,[2] the ratio of not improved to improved to greatly improved SE occlusions was 4:78:18 while Wagner and Berg found a ratio of 0:40:60. In the present study, the ratio of improved to greatly improved LPE occlusions was 39:61 while in the previous study it was reversed at 65:35. Although Wager and Berg[1] found that numbers of patients in each PAR improvement category were different between groups, this difference was reported as insignificant. No p-value was given for this comparison. In the present study, the difference between groups was significant (p<0.001). The most likely reason for this inter-study difference is that the LPE group in the current study began with a higher PAR score and thus had more room for improvement. If the orthodontist has a standard of when a case is “finished,” it is likely that he will strive to reach this goal regardless of the initial occlusal condition. There are some exceptions to this; however Class I SE and LPE patients are most times similar enough that a common final occlusal standard is achievable.

The fact that 48/51 (94%) of SE and 44/49 (90%) of LPE patients finished with a PAR score of 5 or less, using the original weightings from Richmond et al,[3] is evidence to support this (Figure 3 and Table 3). SE and LPE groups were not significantly different (p=0.483) in this categorical measurement. The cutoff mentioned of 5 or less PAR points was first validated by Richmond et al[3] and later used by Tulloch et al[53] in the UNC Class II clinic trial. These guidelines categorize a final PAR score of 0-5 as near ideal, 6-10 as acceptable, and 11 and higher as less acceptable. When applying the PAR weightings of DeGuzman et al,[2] the PAR scores change such that 38/51 of SE and 30/49 of LPE patients finished with a near ideal score of 0-5 (Figure 5 and Table 6). With this change in PAR weighting factors there was still no significant difference between groups for the amount of patients finishing in the near ideal, acceptable, and less acceptable categories (p=0.318). The data on the diagram published by Wagner and Berg[1] showed no significant inter-group difference for the amount of patients finishing in these same categories (p=0.318). These categorical measurements of final PAR scores are likely more meaningful than comparison of means when considering the outcome of treatment because the difference between a PAR score of 2 and 5, while significantly different statistically, may not be clinically significant.

Because the same PAR weighting factors and x-y axis scales were used, the PAR nomogram from the present study (Figure 5) can be visually compared to a similar nomogram previously published in the German Wagner and Berg[1] study. They graphically represent the values reported above.

Figure 2 shows the changes in PAR scores for the SE and LPE groups at all time points during the study period while Figure 7 shows the same changes for anterior segment PAR scores. It is obvious that anterior segment and total weighted PAR scores for the LPE group at T1 are much higher than that for the SE group ($p< .001$ for anterior segment PAR, $p< .0001$ for total PAR). This can be explained by the fact that the SE patients had early intervention via extractions of several primary teeth, some times as many as 12 over the course of several years, thereby allowing the remaining permanent teeth to erupt over basal bone in a less crowded state. The LPE group, on the other hand, had no extractions which forced the remaining permanent teeth to erupt into a more crowded state resulting in a higher PAR score compared to the SE group at T1. Hypothetically, it is possible that if the LPE group received extractions of primary canines, first primary molars, and first premolars at the appropriate times they would not have developed similar occlusal relationships and PAR scores as the SE group at T1. The ideal way to study SE and LPE change over time would be to develop a prospective randomized clinical trial with groups matched for crowding and malocclusion at an early age then randomly allocated to SE or LPE groups to track changes in crowding and malocclusion with time.

Figures 8-11 compare the individual component PAR scores of the present study with those from Wagner and Berg's[1] study while the raw PAR scores for each are shown in Table 7. Although statistical comparisons are not possible due to the lack of standard deviations for component PAR scores in the previous study, it is obvious that T0 PAR scores are at least somewhat similar for the SE groups in each study. The major exceptions are the upper and lower anterior displacement components. The investigator who performed the PAR scoring in the previous study was not calibrated to the scorer in the current study and it is also possible that the different scorers had different systematic error in their measurement of the casts.

There are myriad additional factors that could account for the differences in PAR scores and PAR score changes between studies. Some examples include different diagnostic criteria for the treatment planning of SE and LPE, differences in treatment modalities, mechanics, and finishing standards, and different PAR scoring calibrations of examiners. Despite all of these possible factors and differences in PAR scores between groups and studies, it seems most important that in both studies, there was no significant inter-group differences in the number of patients finishing in the near ideal, acceptable, and less acceptable categories which suggests that both treatment modalities enable the orthodontist to achieve similar occlusal results. This leads to the question of, "how did they get there?" If both modalities can be used to produce similar occlusal results, then other factors differentiating the treatments, such as timing and cost, must be examined.

Figure 12 represents the treatment timing factors for each group. The midline of each graph represents the start of active-treatment which was defined as the date at which separators for orthodontic bands were placed since it represents the beginning of active tooth movement and all of the possible events associated with it. Everything to the right of this midline represents active-treatment timing factor measurements while everything to the left represents pre-active-treatment timing factor measurements. This figure visually shows how the SE patients required more treatment time (months), number of visits, and chair-time (minutes) in the first segment of treatment (i.e. before fixed appliances). If a shorter time in fixed appliances, i.e. braces-free months, is seen as a benefit of SE, then it can be argued that, for each of the three factors, early "investment" before active-treatment "paid-off" in the form of less time (months, minutes, # of visits) with the orthodontist while in fixed appliances. It is reasonable to assume braces-free months is a benefit to the patient; most orthodontists would agree that, as a rule, all else being equal, patients strive to minimize their time in fixed appliances. Therefore, it can be assumed that the majority of patients would prefer pre-active-treatment time (either months, minutes, number of visits) over active-treatment time with appliances.

The fact that the SE group enjoyed, on average, 4.24 braces-free months would suggest that, all other factors being equal, this treatment would be preferable. All other factors, however, are not equal. The SE group required a significant ($p < .0001$) 45.9 months longer total treatment span. Although this is a long time period, it could be argued that, from the patient's perspective, the number of orthodontic visits rather than the treatment span (months) is most important. This is because each patient visit requires that patient and parent to disrupt their normal daily routines such as work and school to travel to the orthodontist's office. The SE group had a statistically significant ($p < .0001$) average of 4.2 fewer active-treatment visits, however they had a statistically significant ($p = .0133$) average of 4.2 more total treatment visits. Stated differently, SE patients invested 8.2 more pre-active-treatment visits than LPE patients in order to save 4.2 visits while in fixed appliances (Figure12).

SE and LPE groups did not have a significantly different mean total treatment chair-time in minutes. The SE group was scheduled for more chair-time before fixed appliances than the LPE group but made up for this with less active-treatment chair-time. Scheduled chair-time likely has less significance from the patient's perspective than from the orthodontist's. Since many orthodontists strive to minimize chair-time in order to maximize efficiency and productivity, the fact that both SE and LPE patients were scheduled for the same amount of total chair-time would suggest that, from a practice management standpoint, that both treatments are equally efficient to the orthodontist.

The present study was designed to investigate the patient perspective; therefore this concept was not developed further

The mean difference in active orthodontic treatment time (in months) between SE and LPE, i.e. "braces-free months" for SE, was reported in other studies as follows: 10.8 by Wagner and Berg,[1] approximately 12 by Little et al,[7] 6.3 by Ringerberg,[10] and 4.2 for the present study. With available data, it is impossible to determine why mean active-treatment time was so different between studies. These differences may be due to variable samples, diagnostic criteria, treatment planning, treatment mechanics, and the definition of orthodontic finishing standards. It is notable that cases are capable of being finished beyond the limits of what PAR scoring can detect. It is possible, for example, for a case with a PAR score of 2 to have excessive buccal root torque on the maxillary buccal segments. An orthodontist may spend several months correcting such a problem and not improve the PAR score as a result. In other words, orthodontists who tend to focus on detailed orthodontic finishing procedures may increase the active-treatment time for their patients without seeing increased occlusal benefit in terms of PAR score. A more sensitive instrument, such as the ABO scoring method, would be necessary to measure some of these improvements not specifically measured by the PAR index.

The result that PAR T0 was significantly correlated with PAR T1 for the SE group ($p<0.001$) suggests that "driftodontics" has a somewhat predictable effect on improving a patient's occlusion. A lack of correlation between PAR T0 and PAR T2 for the SE group and PAR T1 and PAR T2 for the LPE group suggests that, with this sample, the initial malocclusion has little bearing on how a given case finished.

The fact that PAR T1 was significantly correlated with months of active-treatment and number of visits in active-treatment (p-values < .05) for both SE and LPE groups is logical. It would seem plausible to an orthodontist that an increase in severity of malocclusion would lead to an increase in time needed to achieve an acceptable set standard final occlusion via orthodontic treatment. For the SE group, for example, the ratio of increase in PAR points to increase in months of active-treatment is 1:0.27. For the LPE group, however, an inverse relationship is found between PAR T1 and months of active-treatment.

The fact that none of the treatment timing factors were significantly correlated with PAR T2 for either group is expected because of the similarity of PAR T2 scores.

Many orthodontists suggest that increased treatment time produces patient "burnout" and decreased patient compliance. One measure of compliance is the number of "no-shows" or the number of appointments broken by a patient. Assuming prolonged treatment times cause patient burnout, one might assume since the SE group had a longer total treatment time that they also had a higher no-show rate. There was, however, no significant difference in no-show rate between SE and LPE groups during either the total treatment span (p= .6715) or the active-treatment span (p= .6635). This suggests that, in the population studied, the longer total treatment time of the SE group did not lead to patient burn out in the form of higher no-show rates.

B. Cost-effectiveness Discussion

In any two given private practices, the fees paid by the patient have potential to vary greatly both among and between SE and LPE treatment modalities. There is almost an endless combination of fee structures that could be used by an orthodontist to charge for SE or LPE, but it was necessary to choose several realistic fee structures representing a range of different combinations. This is the reason for the low extreme, moderate, and high extreme fee structure categories.

The cost-effectiveness model used in this study makes some important general assumptions. One is that the data we found supporting the similar occlusal outcomes of SE and LPE are correct. Another is that the periodontal and specifically gingival health of both SE and LPE groups are equal. A logical argument could be made that SE leads to better overall gingival health because it may prevent teeth from erupting in unattached gingival as often happens in severely crowded LPE cases. It was not possible to investigate this outcome in the present study. Another argument in favor of SE is the natural alignment of the incisors and possible alleviation of severe crowding that may lessen the chance of psychosocial problems due to an abnormal looking dentition. It was not possible to investigate this outcome either in the present study. A third assumption is that it is actually the treatment of SE itself, and not some other undetected or confounding variable, that yields the braces-free months as compared to LPE.

The major clinically significant benefit of SE over LPE measured in this study is the braces-free months it provides in most cases. The important question is whether parents would be willing to pay more money ($100/BFM, $300/BFM) in order to provide their child braces-free months. Although most adolescent orthodontic patients would most likely wish to have their fixed appliances removed as soon as possible, the willingness of their parents to pay for this is unknown. Based on our model, the best case scenario for the patient and parent would be an orthodontist who offers the SE-Observation fee structure. If this parent was willing to pay up to $100 per BFM, one could predict that SE would be cost-effective compared to LPE only 24% of the time. If the parent was willing to pay up to $300 per BFM then SE is likely cost-effective 58% of the time. The parent will pay for these BFMs regardless of whether or not the extractions and drifting actually do reduce their future active-treatment time in their given case as compared to waiting and choosing LPE.

For the SE-Discounted scenario, the parameter of percentage discount on the active-treatment orthodontic fee was altered in the model from a minimum of 0% to a maximum of 10%. Even at this 10% discount, SE was still more costly than LPE to the patient because the added fees for extractions and diagnostic records overcame the savings on the active-treatment fee. If, however, the orthodontist was willing to discount the active treatment to a higher percentage, then this would increase the likelihood of SE being cost-effective. In fact, it would be possible for an orthodontist to discount the SE active-treatment fee in order to make the total fee for the patient, including additional extractions and records, equal to the total fee of LPE.

SECTION VI

CONCLUSIONS

Based on the sample studied, several conclusions can be made about SE versus LPE. They are:

- SE and LPE yield clinically similar occlusions as measured by the PAR index.
- The SE group had a much longer total treatment time, however patients in this group enjoyed an average of 4.2 BFMs (i.e. a shorter active-treatment time).
- The SE group had more total visits to the orthodontist but fewer visits while in active-treatment compared to the LPE group.
- Both SE and LPE groups were scheduled for the same number of minutes of chair-time throughout the span of treatment.
- PAR T0 was significantly correlated with PAR T1 for the SE group ($p<0.001$) suggesting that the extraction and drifting of teeth has a somewhat predictable effect on improving a patient's occlusion, however the length of time allowed for this drifting is not significantly associated with the number of braces-free months.
- An increase in PAR T1 is associated with an increase in the months of active-treatment for the SE group but not for the LPE group.
- Even at maximal willingness to pay levels, SE is preferred over LPE the following percentages of times
    - 18.6% (observation fee)

- 8.6% (discounted ortho. fee)
- 2.2% (interceptive fee)

Table 1: Base-case parameter estimates and ranges used in sensitivity analysis

Parameter (Fee code)	Base-Case*	Range**	Source(s)
Comprehensive Oral Exam Fee (00150)	$229.92	$186.35-$677.65	ADA Fee Guide, 1999
Diagnostic Casts Fee (00470)			
Panoramic Radiograph (00330)			
Cephalometric Radiograph Fee (no code listed)			
Diagnostic Photos Fee (00471)			
Extraction of 1st Tooth (07111)	$90.76	$78.66-$151.26	ADA Fee Guide, 1999
Extraction of 3 additional teeth (07120)	$250.49	$210.55-$348.50	ADA Fee Guide, 1999
Interceptive Care Fee(08060)	$1,815.13	$1028.57-$3,509.24	ADA Fee Guide, 1999
Observation Visit Fee (08690)	$100	$50-$150	ADA Fee Guide, 1999
Comprehensive Orthodontics Fee (08080)	$5,058.15	$4,658.82-$6,655.46	ADA Fee Guide, 1999
Discount due to ease of care	10%	0% - 10%	Author assumption
# of observation visits	8.42	4.8-12.0	Present Study Data
# sets of extractions (4 teeth each)			
SE	3	---	Fixed by # of teeth
LPE	1	---	Fixed by # of teeth
Time in fixed appliances (in months)			
SE	21.3 (4.8)	20-22.6	Present Study Data
LPE	25.3 (4.0)	24.2-26.4	Present Study Data

* For costs, the 50th percentile of the fee was used for the base-case analysis, for visits and time in braces; the mean value was employed as the base-case value

** For costs, the 25th and 95th percentiles are presented as the range for sensitivity analysis. Otherwise, the range is represented by the 95% confidence interval

Note: All costs are presented in 2006 US dollars. SE: serial extraction, LPE: late premolar extraction.

Table 2: Comparison of PAR scores between groups

		Treatment Group		
		SE	LPE	p-value
	N	51	49	
PAR at T0	Mean	18.24	n/a	
	Std	10.09	n/a	
PAR at T1	Mean	15.18	31.27	<.0001
	Std	8.35	11.68	
PAR at T2	Mean	2.61	3.08	.2652
	Std	2.05	2.18	
Total % Reduction in PAR	Mean	82.01	88.6	.0095
	Std	15.6	8.46	

Table 3: Comparison of %-improvement categorical PAR score outcomes between groups

		Improved PAR	Greatly Improved PAR
Group	SE (n=51)	41	10
	LPE (n=49)	17	32

Table 4: Comparison of final PAR score categorical outcomes between groups

		Final PAR	
		0-5 (ideal)	5 - 10 (acceptable)
Group	SE (n=51)	48	3
	LPE (n=49)	44	5

Table 5: Comparison of % improvement categorical PAR score outcomes between study by Wagner and Berg[1] (W & B) and present study (with weightings by DeGuzman[2])

	Group	Worse- No Different	Improved	Greatly Improved
Present Study	SE (n=51)	2	40	9
	LPE (n=49)	0	19	30
W & B	SE (n=20)	16	3	1
	LPE (n=20)	12	4	4

Table 6: Comparison of final PAR score categorical outcomes between study by Wagner and Berg[1] (W & B) and present study (with weightings by DeGuzman[2])

		Final PAR		
	Group	0-5 (ideal)	5 - 10 (acceptable)	11+ (less acceptable)
Present Study	SE (n=51)	38	12	1
	LPE (n=49)	30	18	1
W & B	SE (n=20)	16	3	1
	LPE (n=20)	12	4	4

Table 7: Anterior segment PAR scores

		Treatment Group		
		SE	LPE	p- value
Upper Ant. PAR at T0	N	51	0	
	Mean	5.20	-	
	Std	4.01	-	
Lower Ant. PAR at T0	N	51	0	
	Mean	2.24	-	
	Std	2.64	-	
Upper Ant. PAR at T1	N	51	49	
	Mean	3.49	5.49	.0001
	Std	2.60	3.38	
Lower Ant. PAR at T1	N	51	49	
	Mean	1.27	4.84	<.0001
	Std	1.59	2.50	
Upper Ant. PAR at T2	N	51	49	
	Mean	0.14	0.12	.8412
	Std	0.40	0.33	
Lower Ant. PAR at T2	N	51	49	
	Mean	0.07	0.06	.7391
	Std	0.27	0.24	

Table 8: Segmental raw PAR scores of present study (Pres.) vs. Wagner and Berg[1]

	SE T0		SE T1		SE T2		LPE T1		LPE T2	
	Pres.	W&B	Pres.	W&B	Pres.	W&B	Pres.	W&B	Pres.	W&B
Upper Ant Segment	5.2	7.5	3.49	n/a	0.14	0	5.49	4.4	0.12	0.2
Lower Ant Segment	2.24	6.3	1.27	n/a	0.08	0.1	4.84	4.9	0.06	0.4
Overjet	1.08	1.8	0.94	n/a	0.84	0.1	2.63	1.5	0.1	0.3
Overbite	0.76	0.7	0.75	n/a	0.2	0.1	0.8	0.8	0.2	0.3
Midline	0.24	0.4	0.2	n/a	0	0.1	0.2	0.2	0	0.1
Right Bucc occl	1.2	1.1	1.33	n/a	0.8	0.3	1.43	1.5	0.9	0.6
Left bucc occl	0.98	0.8	1.16	n/a	0.84	0.4	1.31	1.1	0.98	0.6

Table 9: Total treatment and active-treatment time factors

			Treatment Group			
	Time Span		SE	LPE	ALL	p-value
Treatment Time (months)	Total Treatment	Mean	83.2	37.3	60.7	<.0001
		Std	15.7	17.2	28.3	
	Active Treatment	Mean	21.2	25.3	23.2	<.0001
		Std	4.8	4	4.8	
Number of Visits	Total Treatment	Mean	33.6	31.3	32.5	.0133
		Std	4.9	3.9	4.6	
	Active Treatment	Mean	22.2	26.4	24.2	<.0001
		Std	4.2	3.9	4.5	
Chair time (minutes)	Total Treatment	Mean	915	872	894	.0664
		Std	113.2	115.7	115.8	
	Active Treatment	Mean	692	772	731	.0005
		Std	109.9	115.7	119.3	
		N	51	49	100	

Table 10: Pre-active-treatment time factors and braces-free months (BFM)

		Treatment Group			
		SE	LPE	ALL	p-value
Months Followed Pre Active-Tx	N	51	49	100	
	Median	62.53	4.27	39.16	<.0001
	IR	20.09	9.21	59.44	
# Pre-Active-Tx appointments	N	51	49	100	
	Mean	11.43	4.95	7.30	<.0001
	Std	1.82	0.76	4.43	
Pre-Active-Tx Chair-Time (min)	N	51	49	100	
	Mean	223.24	100.00	162.85	<.0001
	Std	31.27	0.00	65.78	
Braces-Free Months	N	51	0	51	
	Mean	4.24	n/a	4.24	
	Std	4.81	n/a	4.81	
Physiological Drift Time (months) Before Active-Tx.	N	51	0	51	
	Mean	31.9	n/a	31.9	
	Std	8.89	n/a	8.89	

Table 11: Correlation between PAR scores at different time points using Spearman correlation

Variable		Treatment Group	Spearman Correlation	p-value
PAR T0	PAR T1	SE	0.446	.0010
PAR T0	PAR T2	SE	0.275	.0512
PAR T1	PAR T2	LPE	-0.084	.5675

Table 12: Correlation between PAR T0 and treatment timing factors using Spearman correlation

Variable		Treatment Group	Spearman Correlation	P-value
PAR T0	Act. Tx. Time	SE	0.177	.2136
PAR T0	Total Visits	SE	-0.005	.9704
PAR T0	Act. Tx. Visits	SE	-0.113	.4294
PAR T0	Total Chair-time	SE	-0.128	.3718

Table 13: Linear regression testing association of PAR T1 with months of active-treatment, total visits in active-treatment, and chair-time in active-treatment

Model	Source	DF	F statistic	p-value
1	PAR T1	(1, 97)	14.13	.0003
	Tx. Group	(1, 97)	17.24	<.0001
	PAR T1*Tx. Group	(1, 97)	10.77	.0014
	Comparison Test for Group Difference	(2, 97)	8.79	.0003
	Dependent variable: Act. Tx. Time (mos)			
2	PAR T1	(1, 97)	2.82	.0966
	Tx. Group	(1, 97)	16.03	.0001
	PAR T1*Tx. Group		4.82	.0305
	Comparison Test for Group Difference	(2, 97)	11.07	<.0001
	Dependent variable: # Act.-Tx. Visits			
3	PAR T1	(1, 97)	0.04	.8500
	Tx. Group	(1, 97)	8.37	.0047
	Dependent variable: Act. Tx. Chair-time			

Table 14: Parameter estimates for linear regression

Model	Association between (A) & (B)		Tx. group	Estimates		Association between (A) and (B) is different according to treatment group?*	Association between (A) and (B) is statistically significant?**
	(A)	(B)		Intercept	Slope		
1	Act. Tx. Time (mos)	PAR T1	SE	17.24	0.27		
	Act. Tx. Time (mos)	PAR T1	LPM E	26.00	-0.02	Yes	NA
2	# Act.- Tx. Visits	PAR T1	SE	20.50	0.11		
	# Act.- Tx. Visits	PAR T1	LPM E	28.65	-0.07	Yes	NA
3	Act.- Tx. Chair-time (min)	PAR T1	SE	695.49	-0.22		
	Act.- Tx. Chair-time (min)	PAR T1	LPM E	779.67	-0.22	No	No

* Whether two groups have same slopes?

** If slope is the same for two groups, Is the slope statistically non-zero?

Table 15: Linear regression testing the association of active-treatment (months), total treatment time (months), active-treatment number of visits, total treatment number of visits, total treatment chair-time, and active-treatment chair-time with PAR T2

Model	Source	DF	F statistic	p-value
1	Act. Tx. Time (mos)	(1, 97)	0.72	.3974
	txgrp	(1, 97)	1.88	.1731
2	Tot. Tx. Time (mos)	(1, 97)	0.01	.9297
	txgrp	(1, 97)	0.33	.5671
3	Act.-Tx. # of Visits	(1, 97)	1.19	.2788
	txgrp	(1, 97)	2.24	.1376
4	Tot. tx. # of Visits	(1, 97)	0.67	.4167
	txgrp	(1, 97)	0.78	.3796
5	Tot. Tx. Chair-time (min)	(1, 97)	0.60	.4391
	txgrp	(1, 97)	0.91	.3415
6	Act.-Tx. Chair-time (min)	(1, 97)	0.77	.3813
	txgrp	(1, 97)	1.83	.1797
Dependent variable: (PAR T2)				

Table 16: Pearson correlation coefficients for both groups combined

	Pearson Corr. Coeff's (r-value)	p-value
% PAR red. with Tot. Tx. Time (mos)	0.0182	0.0018
% PAR red. With Act. Tx. Time (mos)	0.1183	0.1246
% PAR red. With # Tot. Tx. visits	0.2906	0.0046
% PAR red. With # Act. Tx. visits	0.3187	0.0785

Table 17: Pearson correlation coefficients* for SE group

	Pearson Correlation Coefficient (r-value)
Months of drift and braces-free months	0.2055

* because r<0.30, p-value was not calculated

Table 18: Economic evaluation, base-case results**

			Incremental		
Alternative	Costs	Time in braces (mos.)	Costs	BFM	Ratio
SE-Interceptive	$6,479	21.3	$1,639.15	4.06	$403.73
SE-Observation	$7,188	21.3	$887.97	4.06	$218.71
SE-Discounted Ortho.	$5,312	21.3	$443.40	4.06	$109.21
LPE	$4,856	25.3			

**All costs are in 2006 US dollars. Outcomes measured in time in active-treatment, with difference between scenarios reported as braces-free months (BFM). Incremental cost-effectiveness ratio reported as additional costs per braces-free month.
SE: serial extraction, LPE: late premolar extraction, Ortho: orthodontic fee

Figure 1: PAR (with +1 and -1 standard deviation) of SE and LPE at all time points

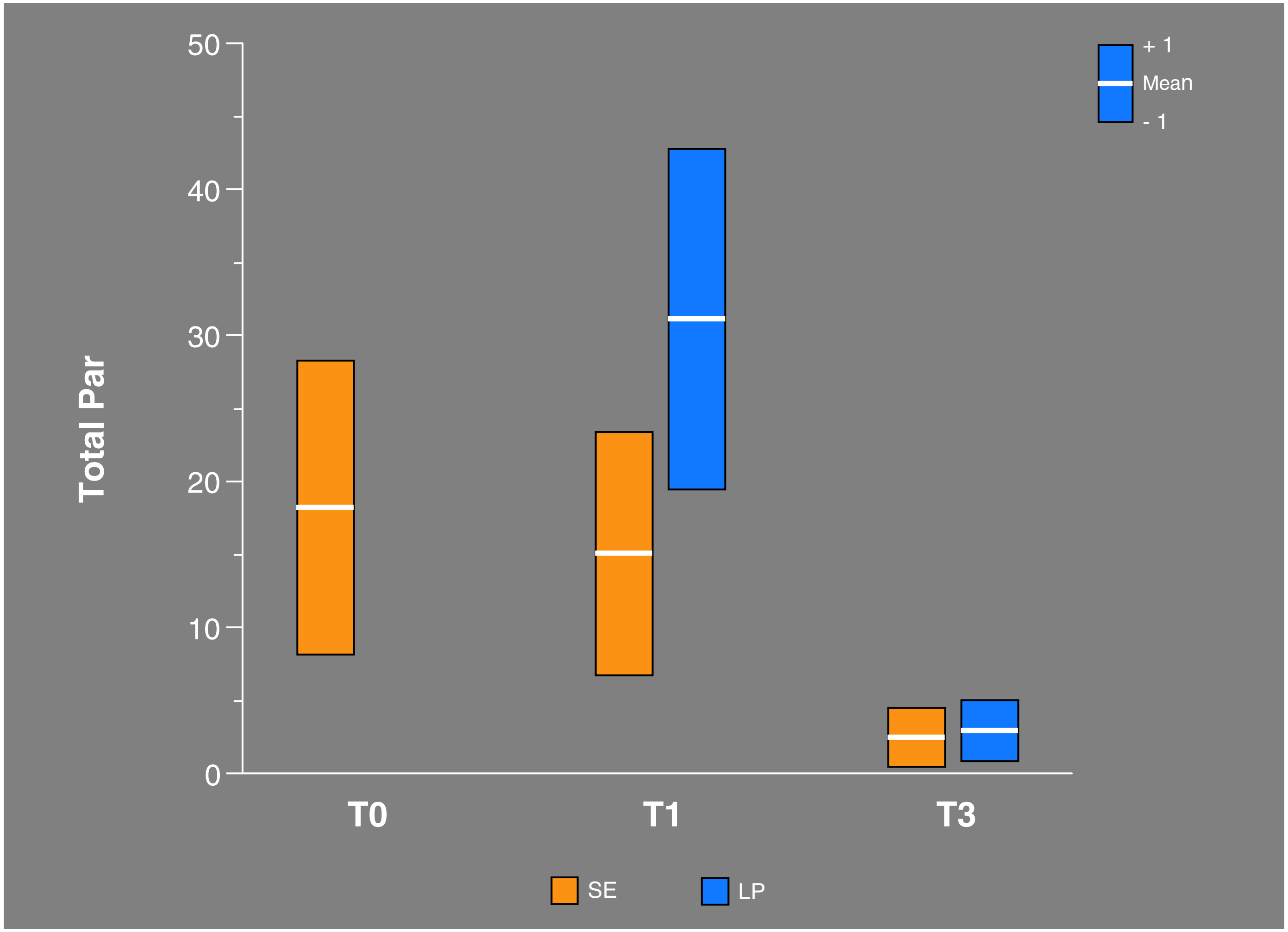


Figure 2: Change in mean PAR score over time for SE and LPE groups

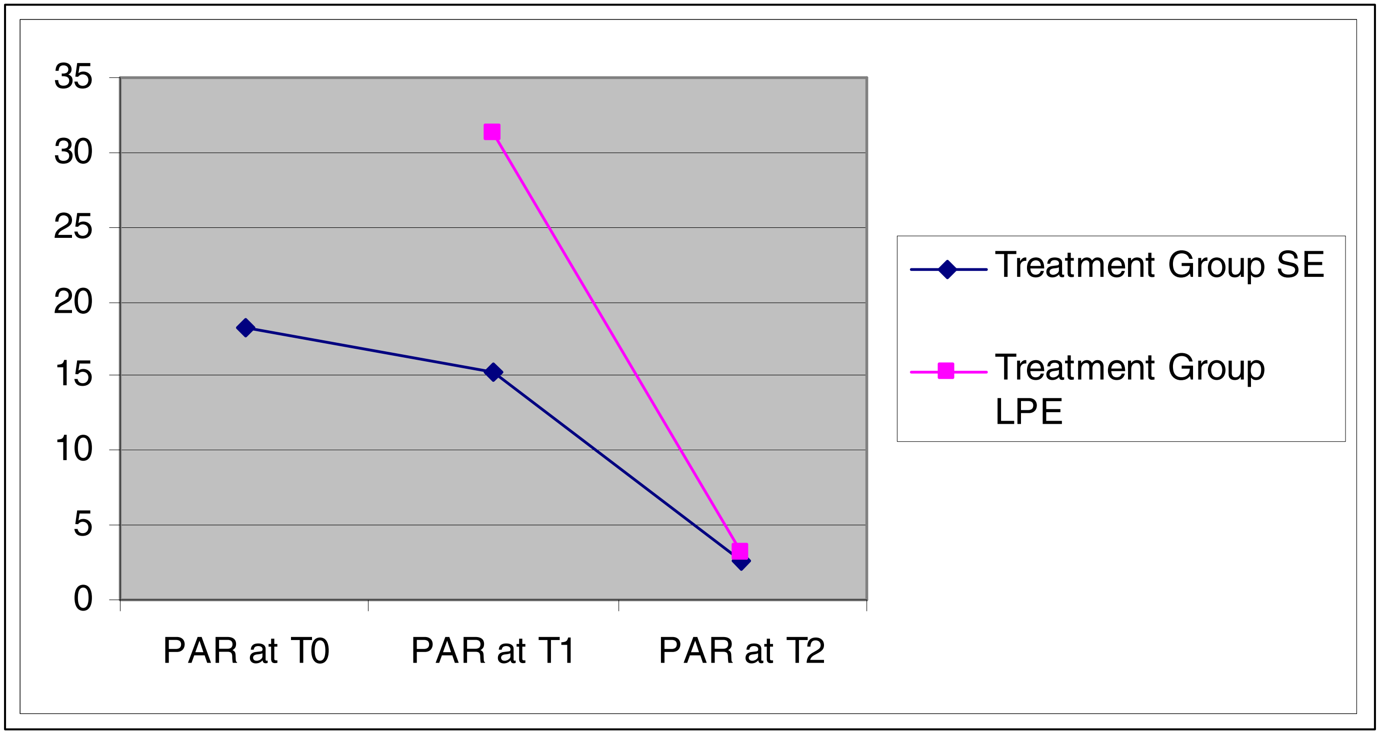


Figure 3: Nomogram plotting PAR before treatment vs. PAR after treatment with original PAR weightings according to Richmond et al[3]

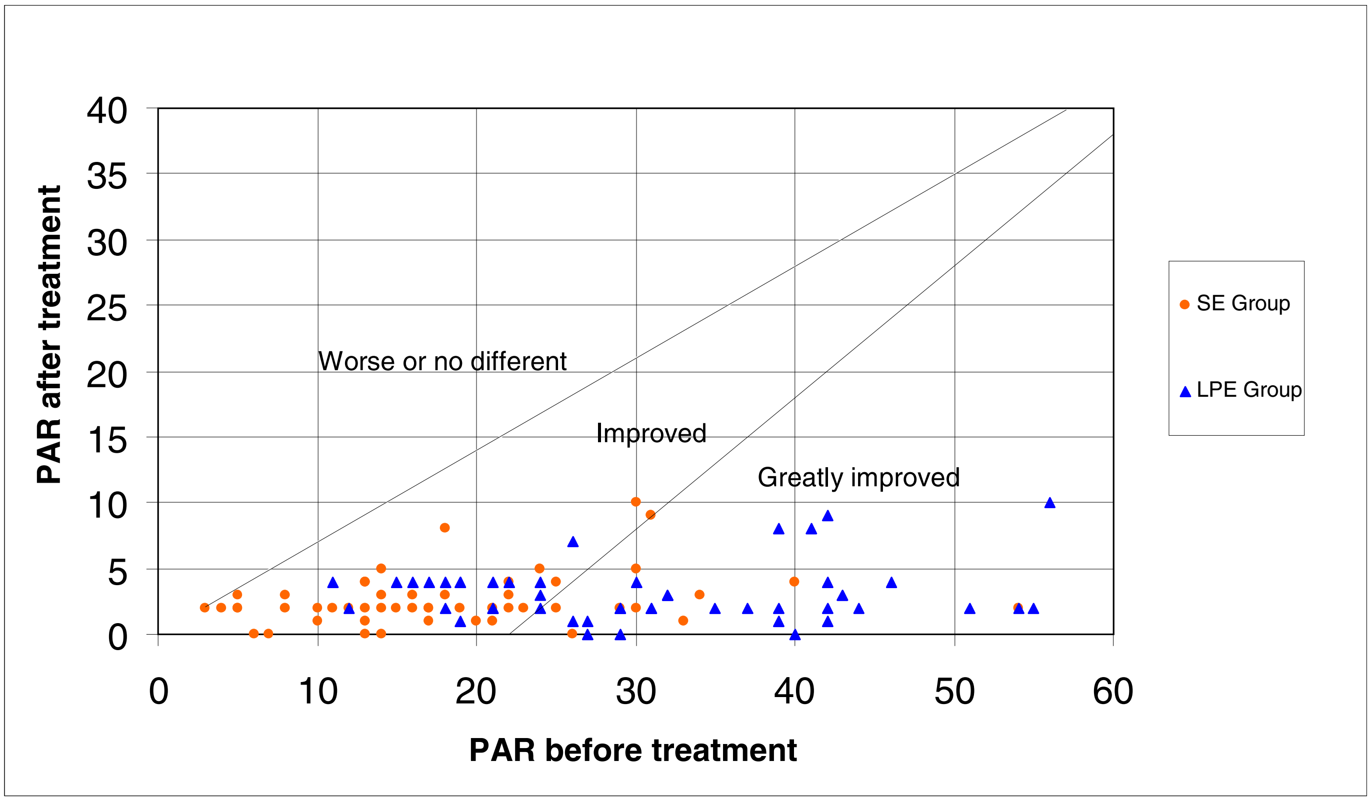
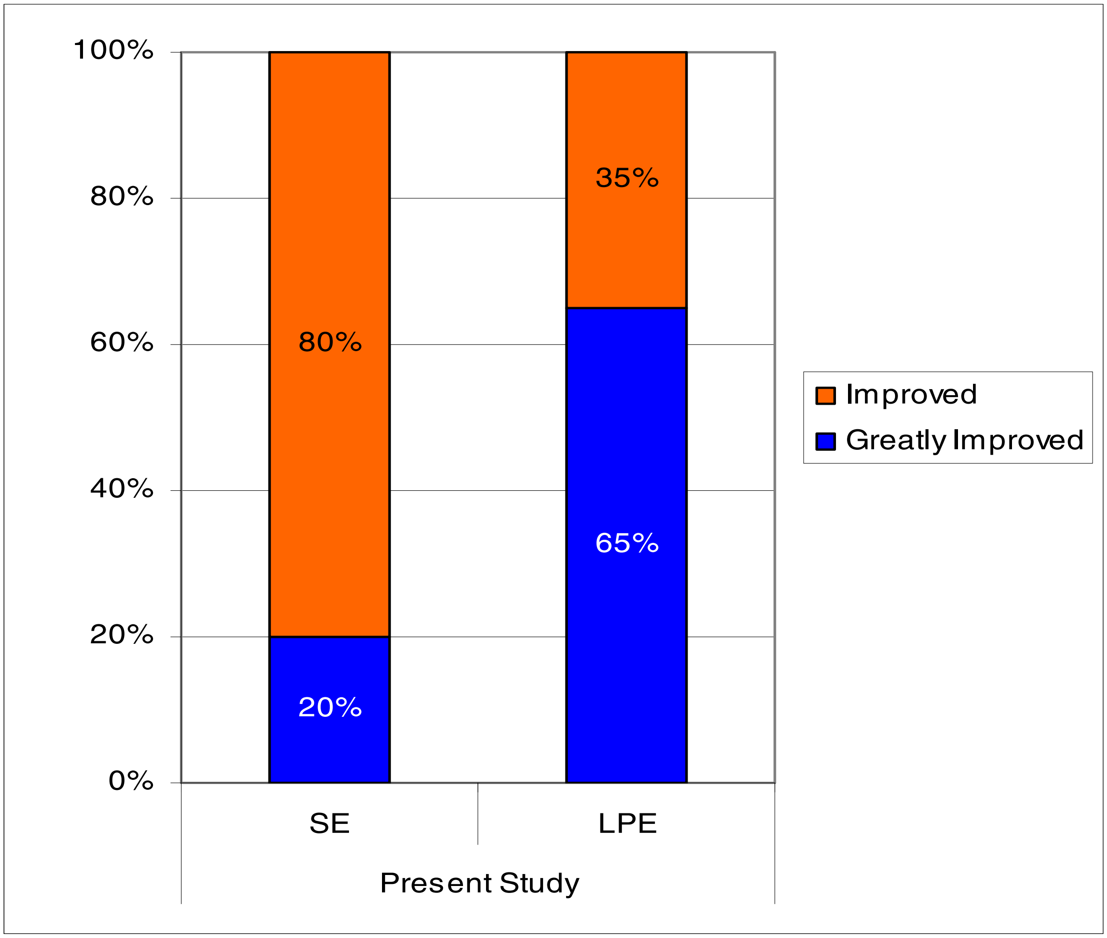


Figure 4: Percent improvement of PAR categorical outcomes*



*Calculations used original PAR weightings according to Richmond et al.(1992); Improved= minimum of 30% reduction of PAR score; greatly improved= reduction of 22 or more PAR points

Figure 5: Nomogram plotting PAR before treatment vs. PAR after treatment with PAR weightings according to DeGuzman et al[2]

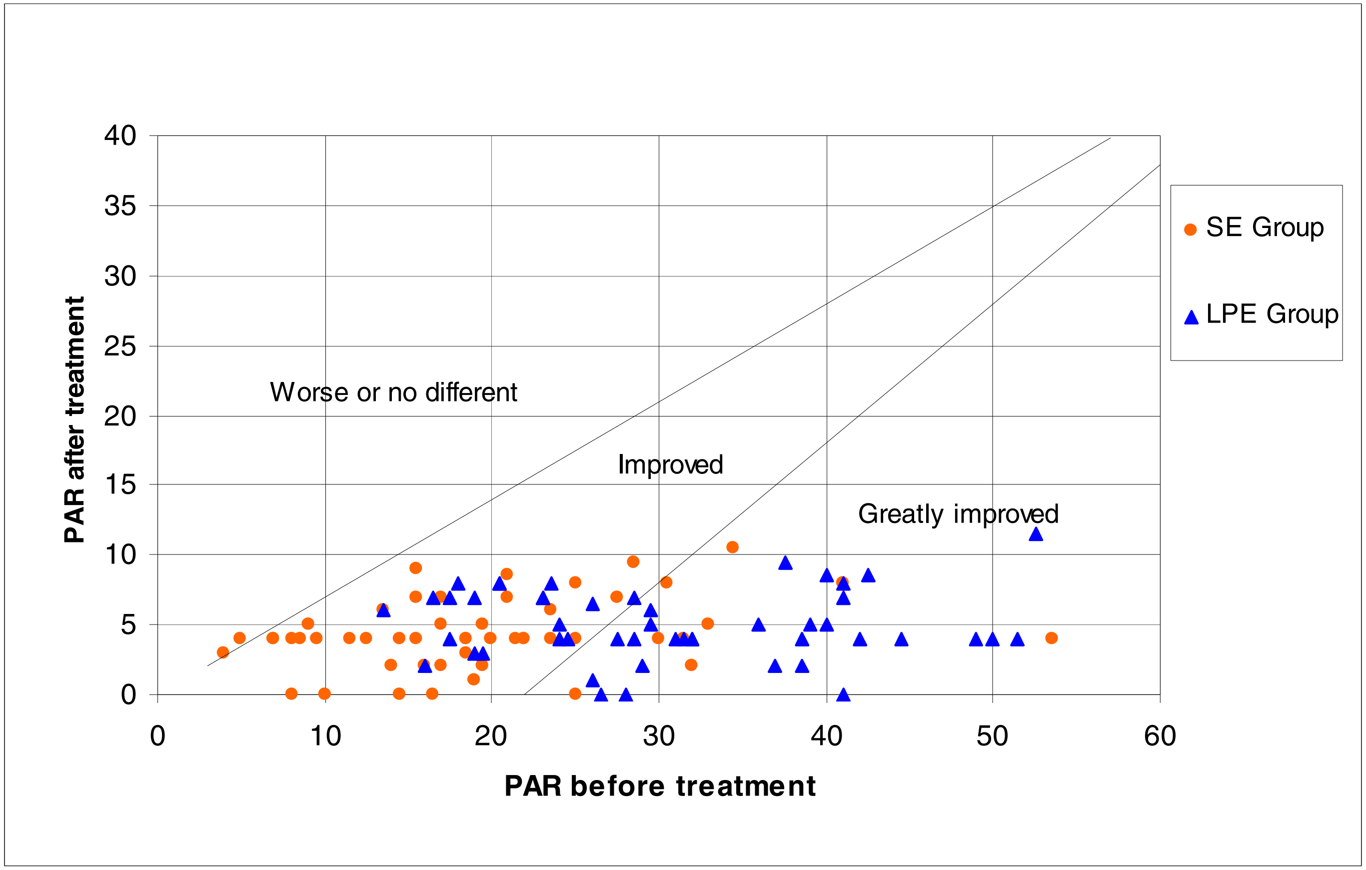
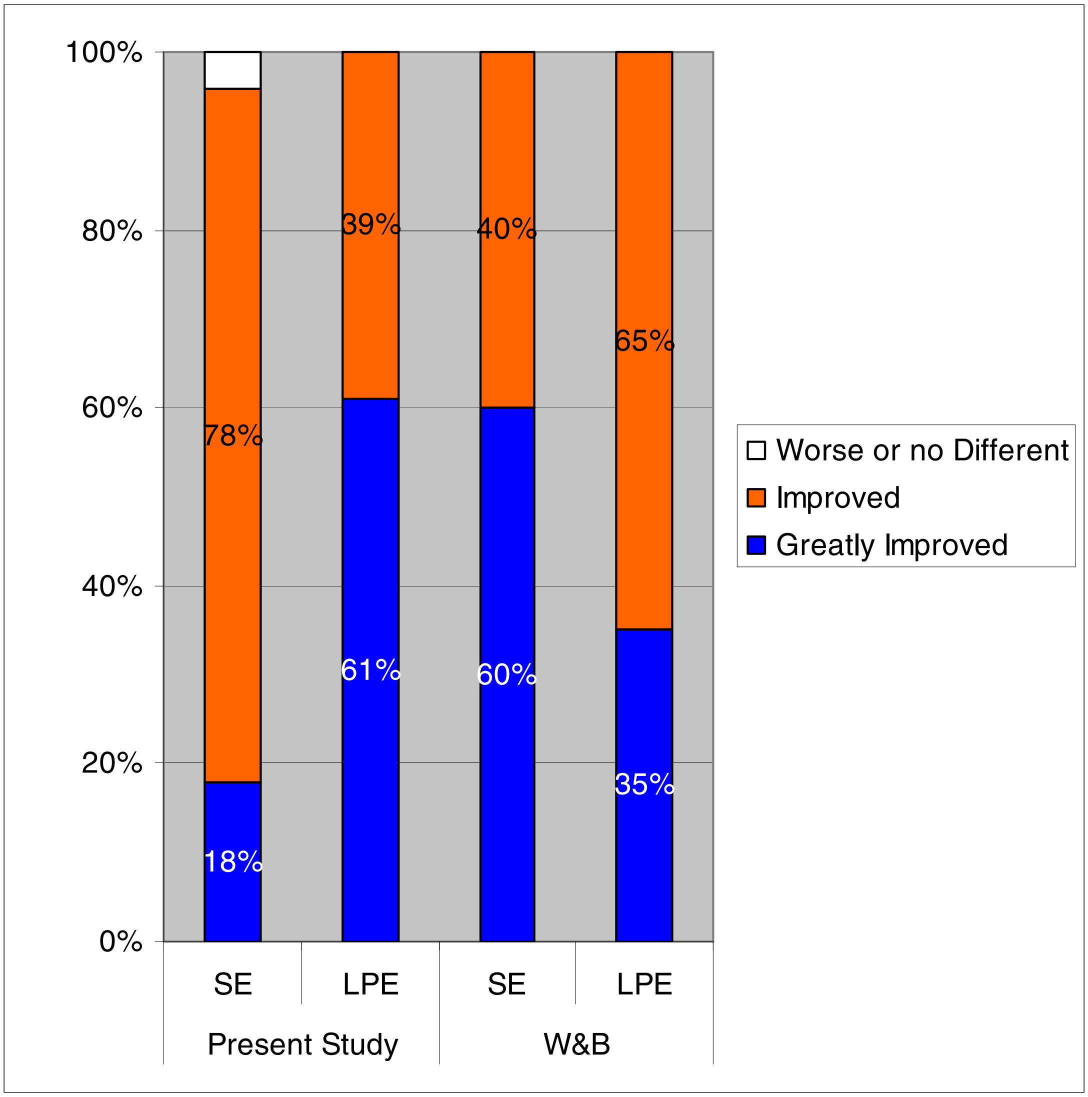


Figure 6: Percent improvement of PAR categorical outcomes of Wagner and Berg[1] compared to the present study*



*Calculations used PAR weightings according to DeGuzman et al.(1995); Improved= minimum of 30% reduction of PAR score; greatly improved= reduction of 22 or more PAR points

Figure 7: Change in mean anterior segments PAR score over time for SE and LPE groups

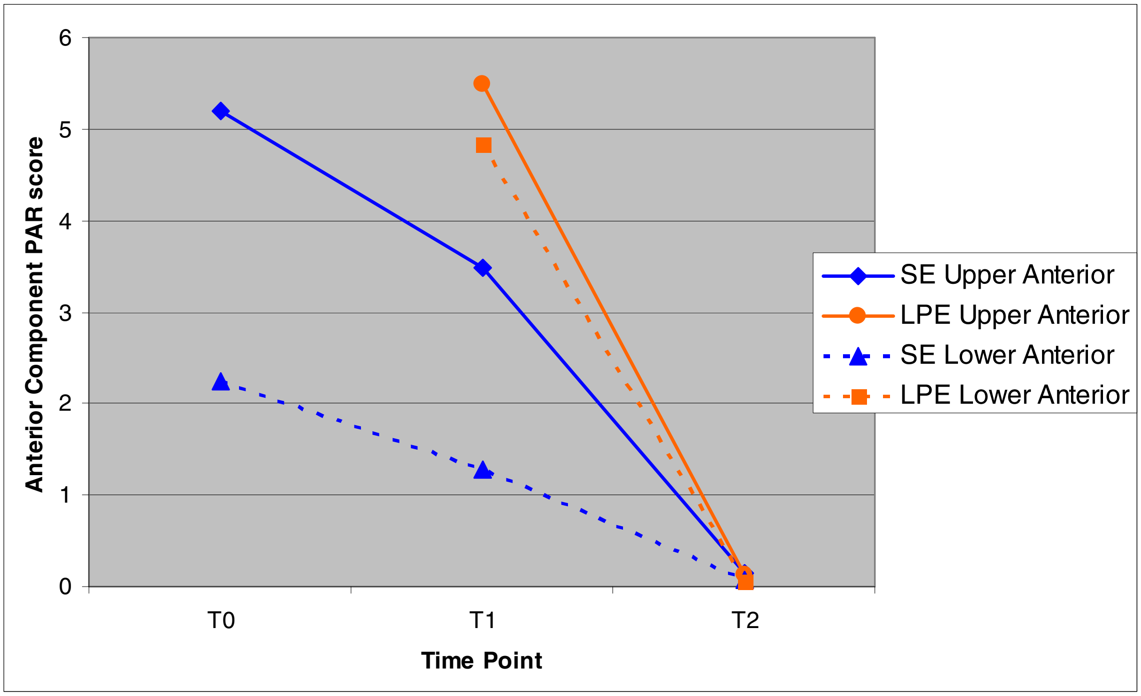


Figure 8: Pre-serial extraction (T0) component PAR scores for SE group of present study vs. Wagner and Berg[1]

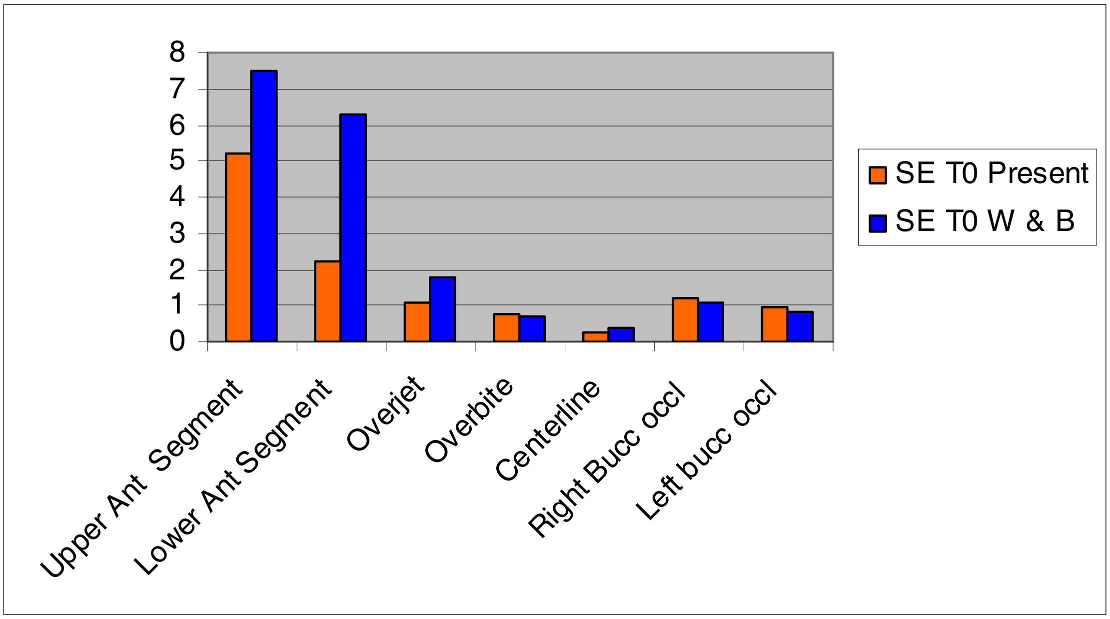


Figure 9: Final, posttreatment (T2) component PAR scores for SE group of present study vs. Wagener and Berg[1]

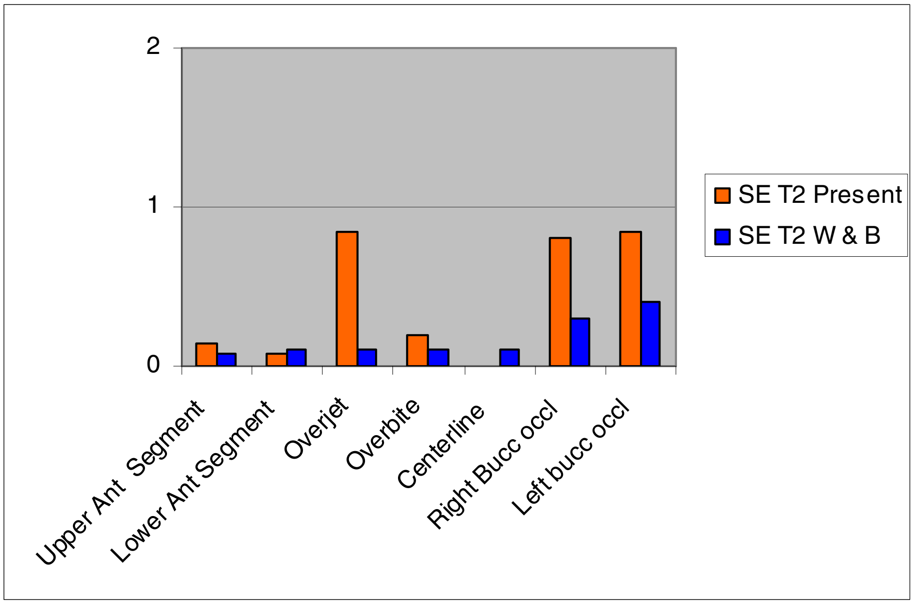


Figure 10: Pretreatment (T1) component PAR scores for LPE group of present study vs. Wagner and Berg.[1]

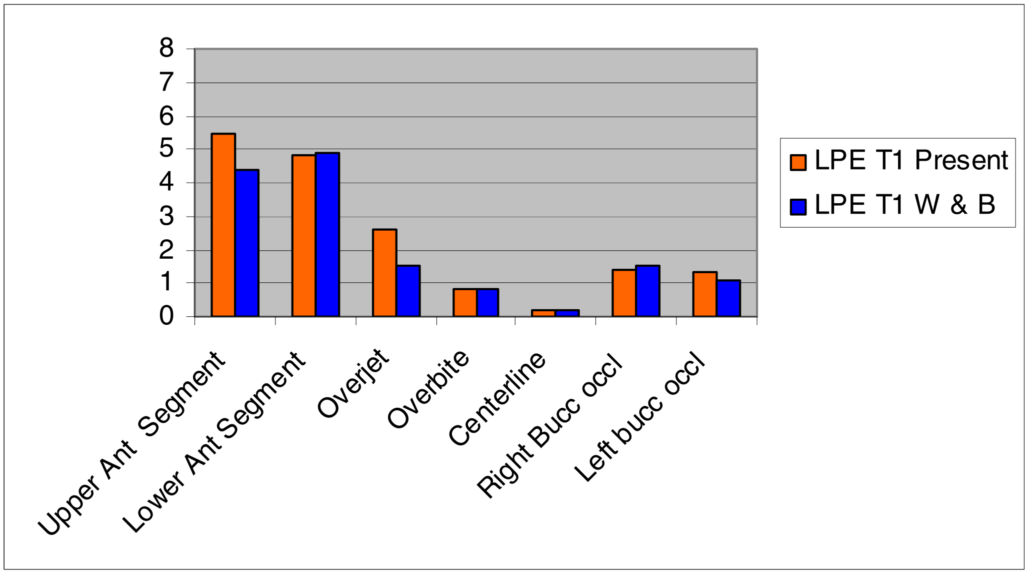


Figure 11: Posttreatment (T2) component PAR scores for LPE group of present study vs. Wagner and Berg.[1]

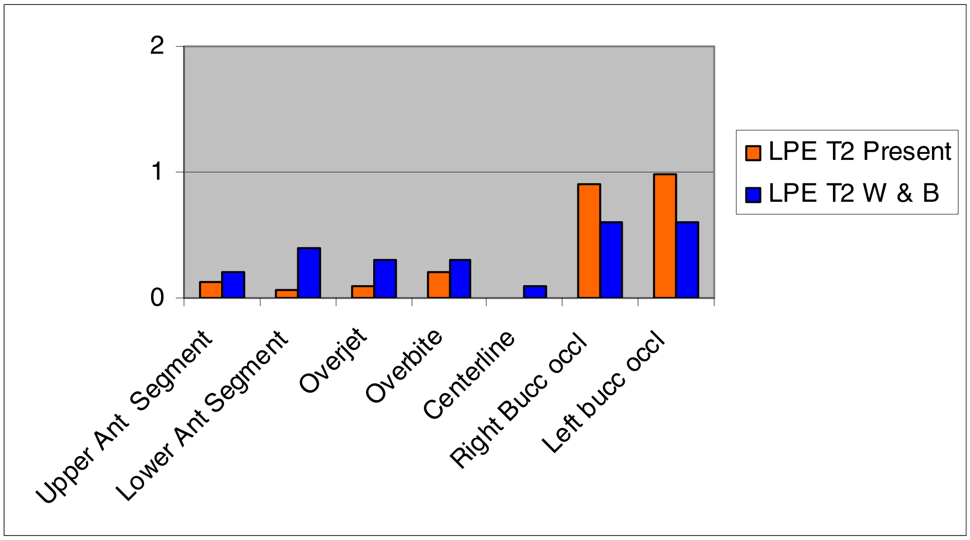
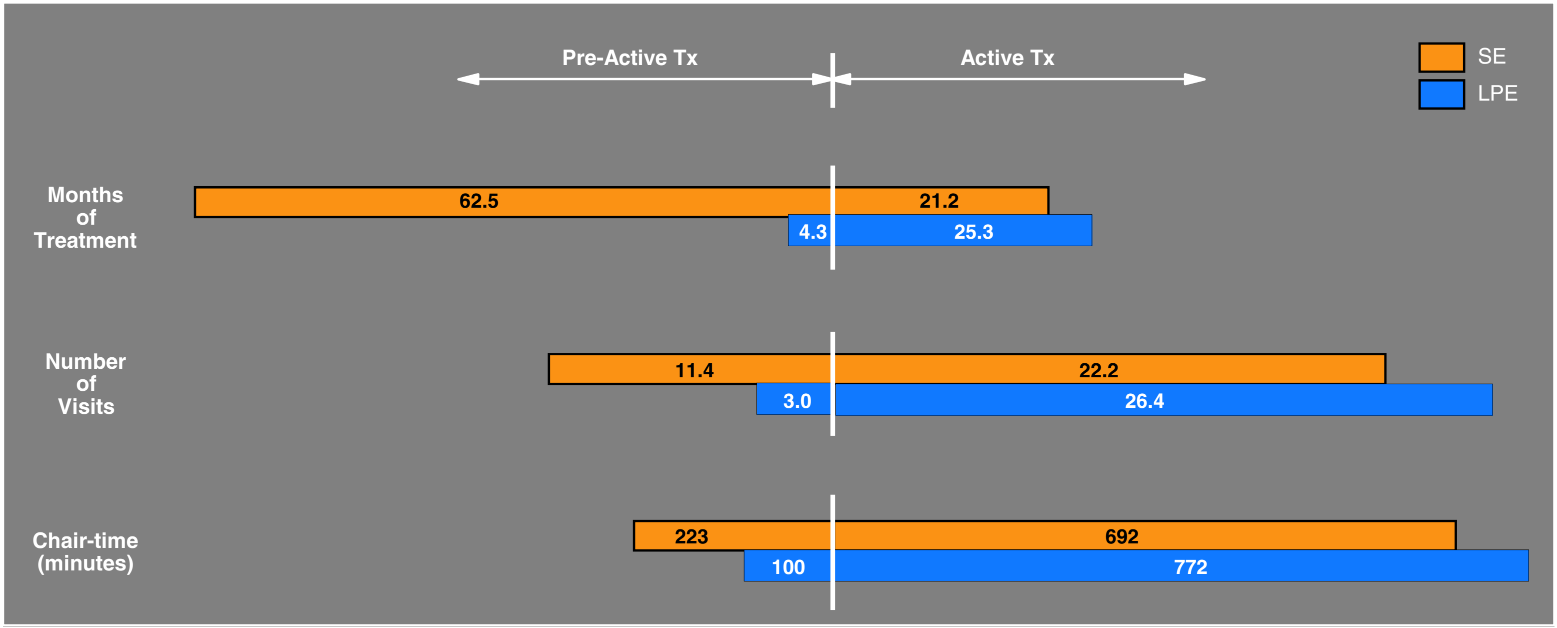


Figure 12: Diagrammatic representation* of time factors of treatment for SE and LPE groups.



* The center line represents the beginning of active treatment.

Figure 13: Plot of PAR T1 across months of active-treatment for SE and LPE groups

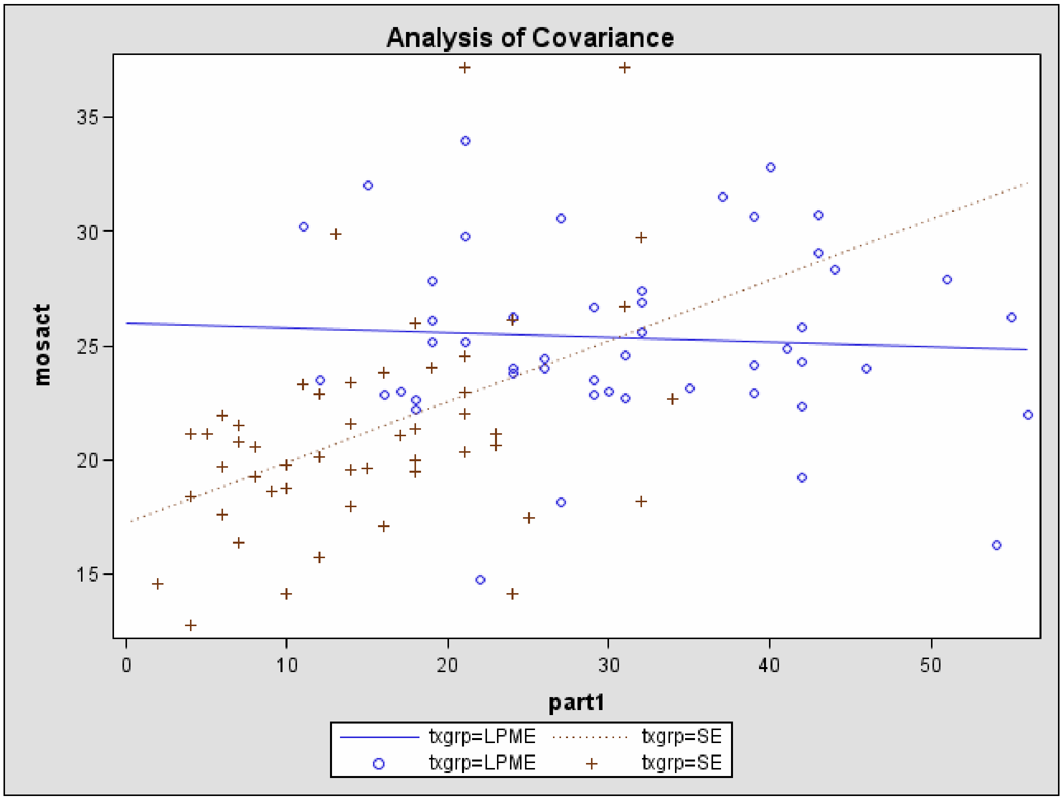


Figure 14: Plot of PAR T1 across total number of visits in active treatment for SE and LPE groups

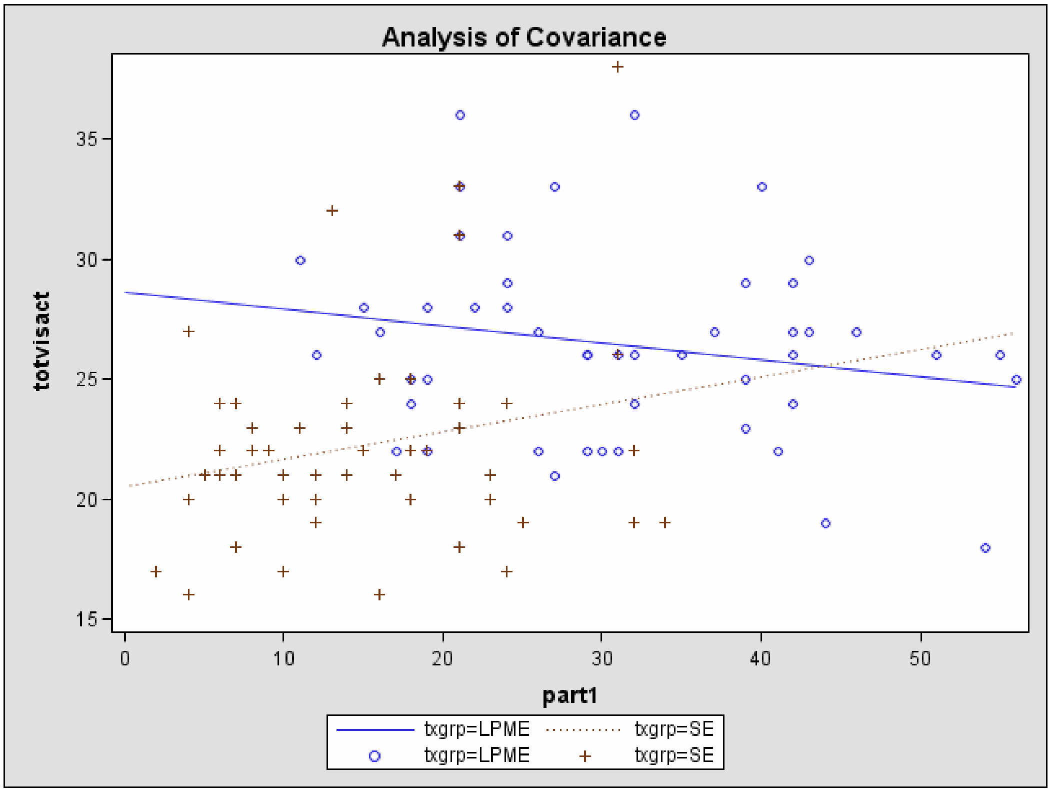


Figure 15: Plot of PAR T1 across active treatment chair-time (minutes) for SE and LPE

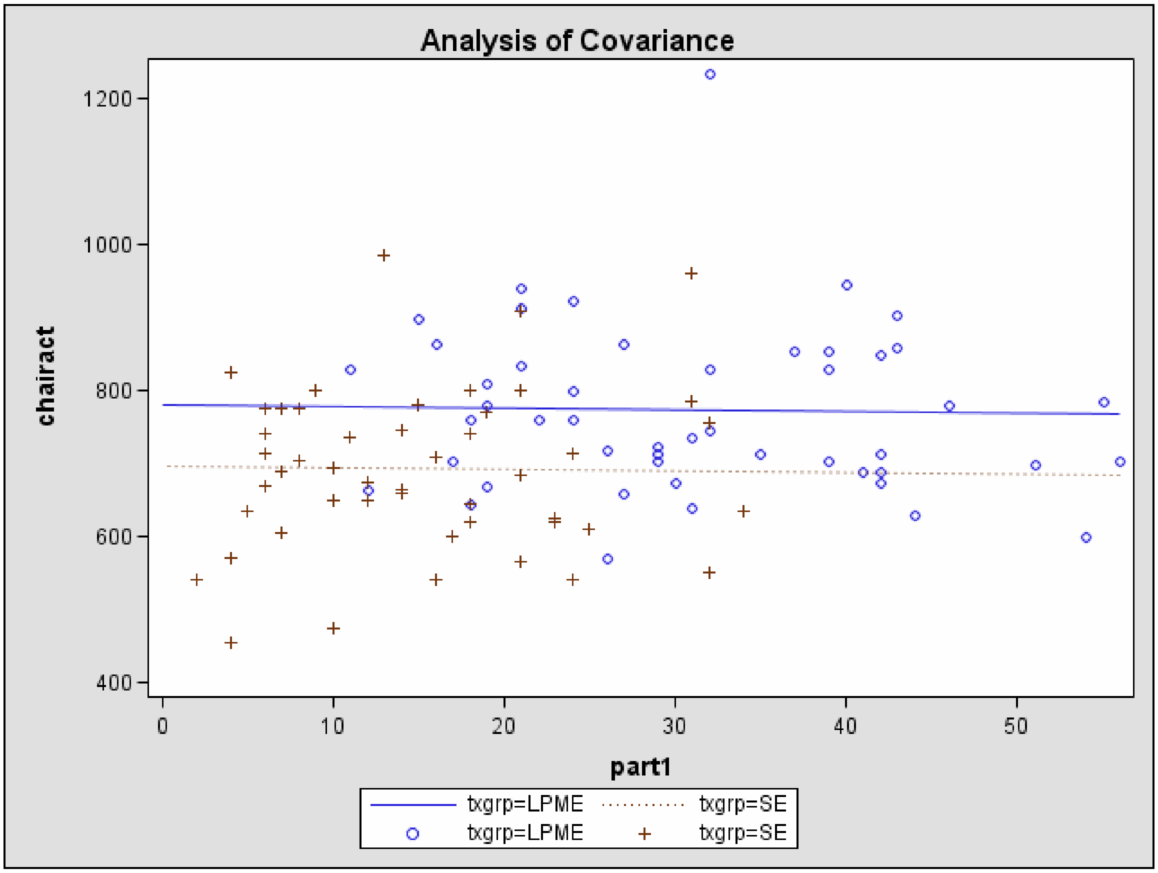


Figure 16: Tornado Chart comparing SE-Interceptive Scenario with LPE

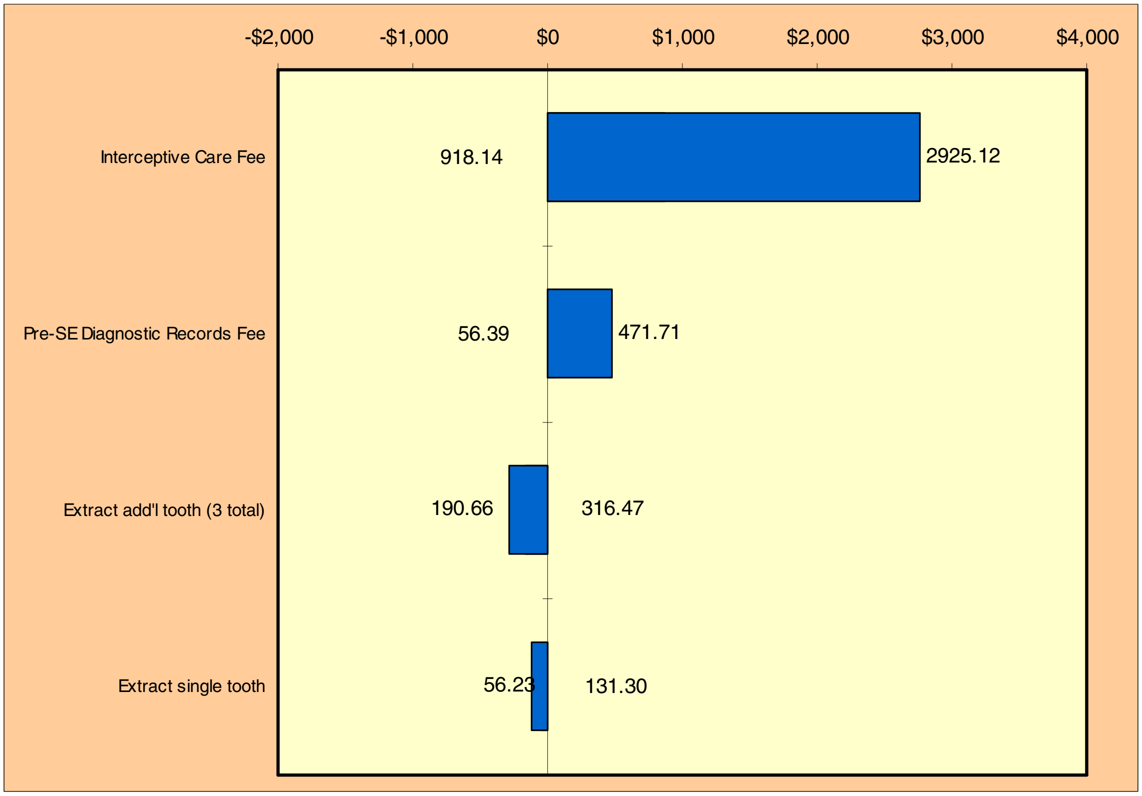


Figure 17: Tornado Chart comparing SE-Observation Scenario with LPE

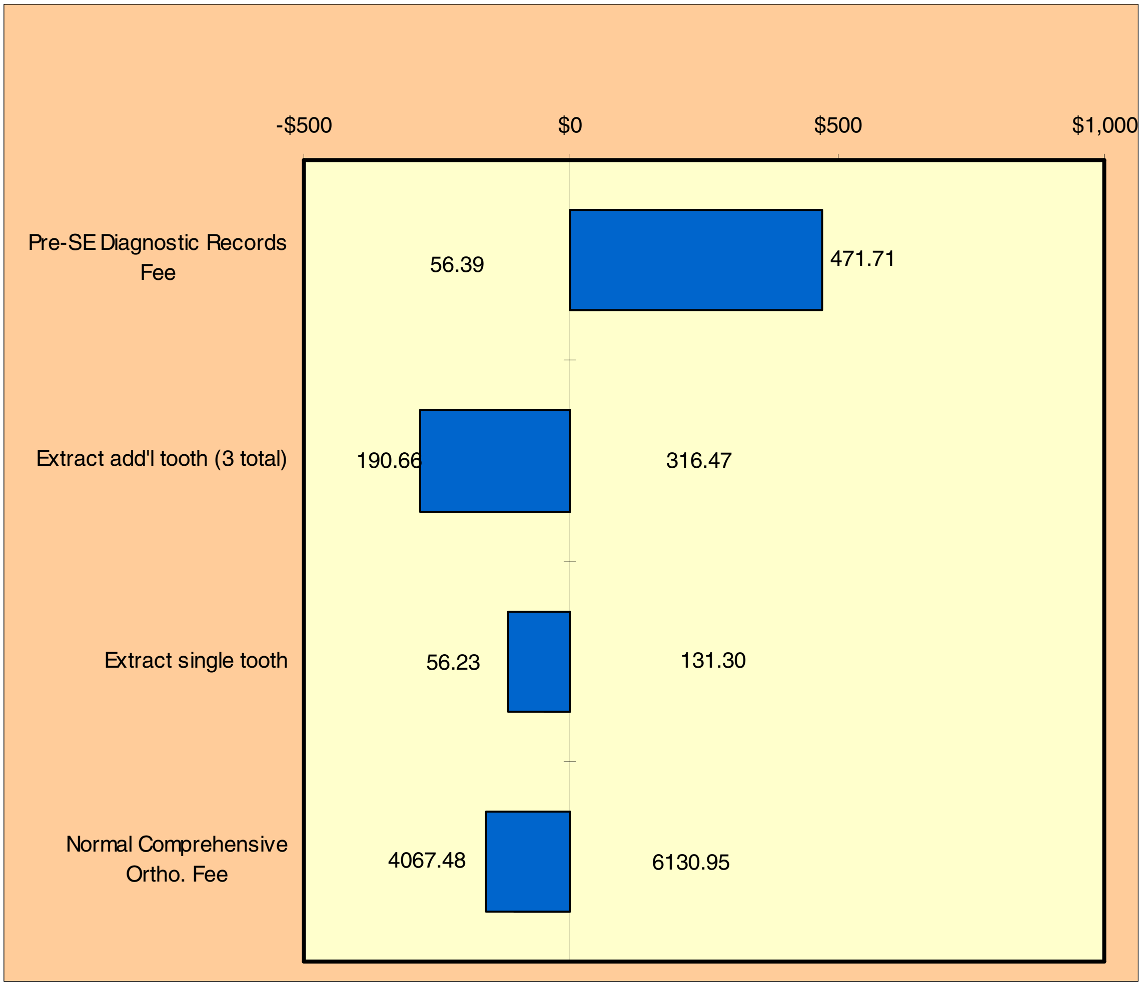


Figure 18: Tornado Chart comparing SE-Discounted Scenario with LPE

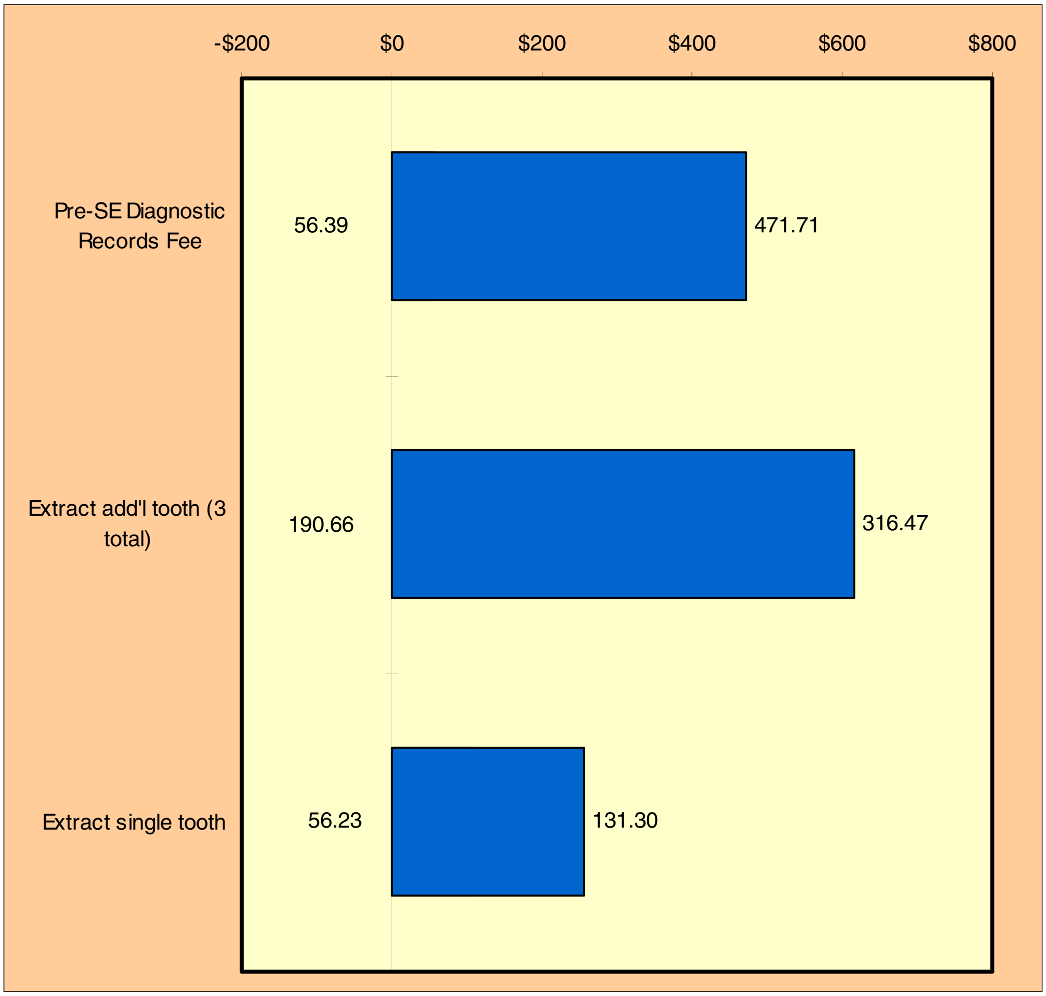


Figure 19: Incremental Cost-Effectiveness Ratio (ICER) plane comparing SE options to LPE

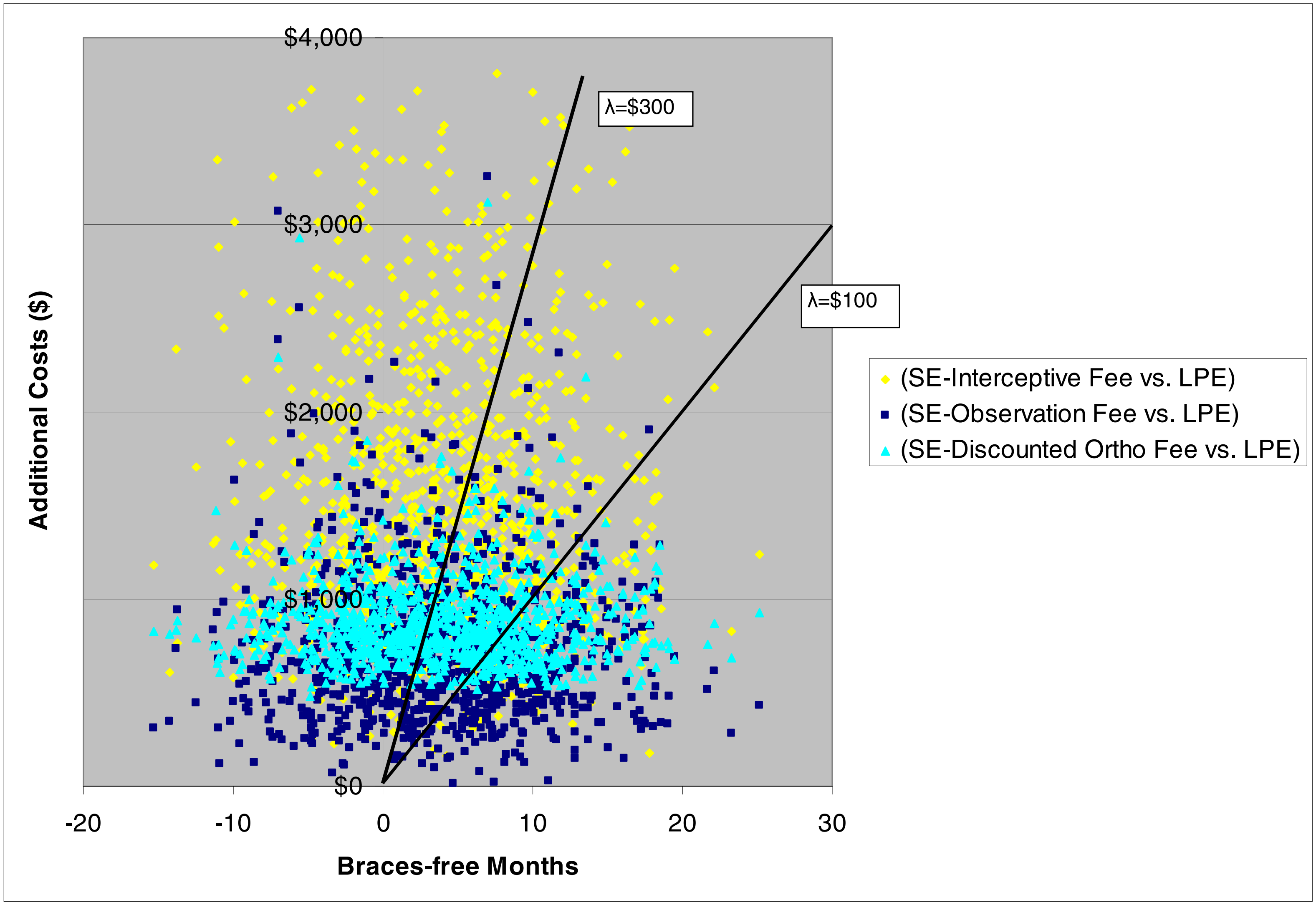
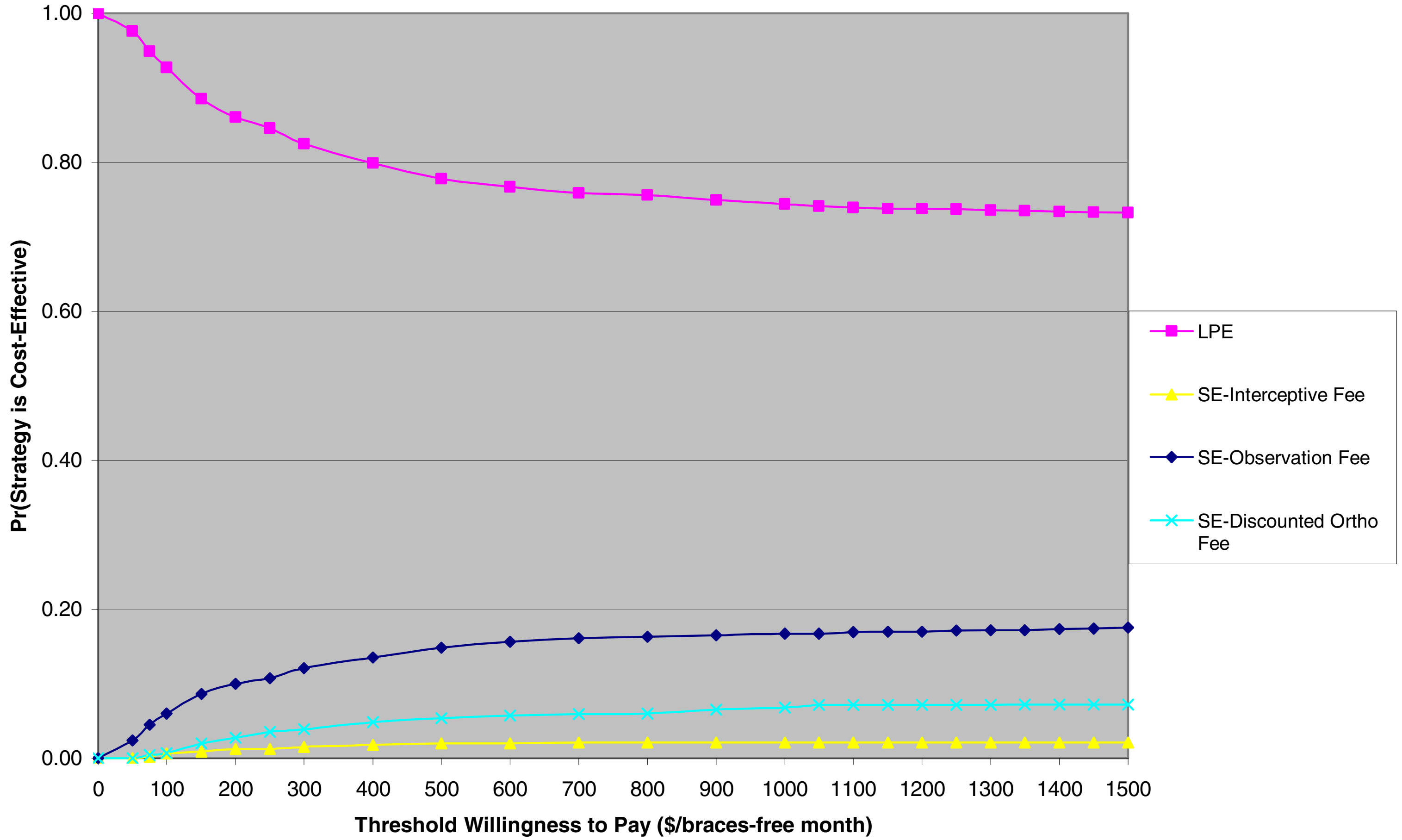


Figure 20: Cost-Effectiveness Acceptability Curves for comparing all Strategies to reduce Time in Braces



REFERENCES

1. Wagner M, Berg R. Serial extraction or premolar extraction in the permanent dentition? Comparison of duration and outcome of orthodontic treatment. J Orofac Orthop 2000;61:207-216.

2. DeGuzman L, Bahiraei D, Vig KW, Vig PS, Weyant RJ, O'Brien K. The validation of the Peer Assessment Rating index for malocclusion severity and treatment difficulty. Am J Orthod Dentofacial Orthop 1995;107:172-176.

3. Richmond S, Shaw WC, O'Brien KD, Buchanan IB, Jones R, Stephens CD et al. The development of the PAR Index (Peer Assessment Rating): reliability and validity. Eur J Orthod 1992;14:125-139.

4. Bunon R. Essay sur les maladies des dents, ou l' on propose les moyens de leur procurer une bonne confirmation des la plus tendre enfance, et d'en assurer la conservation pendant tout le cours de la vie 1743.

5. Kjellgren B. Serial extraction as a corrective procedure in dental orthopedic therapy. Trans Eur Orthod Soc 1947-8:134-160.

6. Hotz RP. Active supervision of the eruption of the teeth by extraction. Trans Eur Orthod Soc 1947-48:34-47.

7. Little RM, Riedel RA, Engst ED. Serial extraction of first premolars--postretention evaluation of stability and relapse. Angle Orthod 1990;60:255-262.

8. Wilson JR, Little RM, Joondeph DR, Doppel DM. Comparison of soft tissue profile changes in serial extraction and late premolar extraction. Angle Orthod 1999;69:165-173; discussion 173-164.

9. McReynolds DC, Little RM. Mandibular second premolar extraction--postretention evaluation of stability and relapse. Angle Orthod 1991;61:133-144.

10. Ringenberg QM. Influence of serial extraction on growth and development of the maxilla and mandible. Am J Orthod 1967;53:19-26.

11. Little RM. The effects of eruption guidance and serial extraction on the developing dentition. Pediatr Dent 1987;9:65-70.

12. Dale JG. Serial extraction. nobody does that anymore! Am J Orthod Dentofacial Orthop 2000;117:564-566.

13. Norman F. Serial Extraction. Angle Orthod 1965;35:149-157.

14. Graber TM. Serial extraction: a continuous diagnostic and decisional process. Am J Orthod 1971;60:541-575.

15. Dewel BF. A critical analysis of serial extraction in orthodontic treatment. Am J Orthod 1959;45:424-455.

16. Dale JG, Brandt S. Dr. Jack G. Dale on serial extraction. 3. J Clin Orthod 1976;10:196-217.

17. Richmond S, Shaw WC, Roberts CT, Andrews M. The PAR Index (Peer Assessment Rating): methods to determine outcome of orthodontic treatment in terms of improvement and standards. Eur J Orthod 1992;14:180-187.

18. Linderer J. Die Zahnheilkunde nach ihrem neuesten Standpunkte 1851:202-210.

19. Hotz RP. Guidance of eruption versus serial extraction. Am J Orthod 1970;58:1-20.

20. Pont A. Der Zahn-Index in der Orthodontie. Zahnarztuche Orthopodie 1909;3:306-321.

21. Nimkarn Y, Miles PG, O'Reilly MT, Weyant RJ. The validity of maxillary expansion indices. Angle Orthod 1995;65:321-326.

22. Joondeph DR, Riedel RA, Moore AW. Pont's index: a clinical evaluation. Angle Orthod 1970;40:112-118.

23. Worms FS, TM Isaacson, RJ Meskin, LH. Pont's index and dental arch form. J Am Dent Assoc 1972;85:876-881.

24. Tweed C. Clinical orthodontics, Vol. 1. St. Louis: Mosby; 1966.

25. Riedel RA, Brandt S. Dr. Richard A. Riedel on retention and relapse. J Clin Orthod 1976;10:454-472.

26. Little RM. The irregularity index: a quantitative score of mandibular anterior alignment. Am J Orthod 1975;68:554-563.

27. Woodside DG, Rossouw PE, Shearer D. Postretention mandibular incisor stability after premolar serial extractions. Semin Orthod 1999;5:181-190.

28. Haruki T, Little RM. Early versus late treatment of crowded first premolar extraction cases: postretention evaluation of stability and relapse. Angle Orthod 1998;68:61-68.

29. Smolen G. A cephaloetric evaluation of class I serial extraction treatment. Unpublished thesis, St. Louis Univeristy, 1965.

30. Boley JC. Serial extraction revisited: 30 years in retrospect. Am J Orthod Dentofacial Orthop 2002;121:575-577.

31. Whitney D. A cephalometric evaluatoin of class I serial extraction treatment. Unpublished thesis, St. Louis University, 1965.

32. Croce R. A cephalometric evaulation of vertical dimension in serial extractoin, class I. Unpublished thesis, St. Louis University, 1966.

33. Dannelly W. A cephalometric evaluation of serial extractoin in soft tissue, class I. Unpublished thesis, St. Louis University, 1966.

34. Glauser RO. An evaluation of serial extraction among Navajo Indian children. Am J Orthod 1973;63:622-632.

35. Persson M, Persson EC, Skagius S. Long-term spontaneous changes following removal of all first premolars in Class I cases with crowding. Eur J Orthod 1989;11:271-282.

36. Sadowsky C, Sakols EI. Long-term assessment of orthodontic relapse. Am J Orthod 1982;82:456-463.

37. Papandreas SG, Buschang PH, Alexander RG, Kennedy DB, Koyama I. Physiologic drift of the mandibular dentition following first premolar extractions. Angle Orthod 1993;63:127-134.

38. Yoshihara T, Matsumoto Y, Suzuki J, Ogura T. Effect of serial extraction alone on crowding: relationship between closure of residual extraction space and changes in dentition. J Clin Pediatr Dent 2002;26:147-153.

39. Yoshihara T, Matsumoto Y, Suzuki J, Sato N, Oguchi H. Effect of serial extraction alone on crowding: relationships between tooth width, arch length, and crowding. Am J Orthod Dentofacial Orthop 1999;116:691-696.

40. Yoshihara T, Matsumoto Y, Suzuki J, Sato N, Oguchi H. Effect of serial extraction alone on crowding: spontaneous changes in dentition after serial extraction. Am J Orthod Dentofacial Orthop 2000;118:611-616.

41. Kennedy DB, Joondeph DR, Osterberg SK, Little RM. The effect of extraction and orthodontic treatment on dentoalveolar support. Am J Orthod 1983;84:183-190.

42. Richmond S. The need for cost-effectiveness. J Orthod 2000;27:267-269.

43. Richmond S, Phillips CJ, Dunstan F, Daniels C, Durning P, Leahy F. Evaluating the cost-effectiveness of orthodontic provision. Dent Update 2004;31:146-152.

44. Demirjian A, Goldstein H, Tanner JM. A new system of dental age assessment. Hum Biol 1973;45:211-227.

45. Gold M, Siegel, J.E., Russell, I., Weinstein, M.C. *Cost-effectiveness in Health and Medicine,* 1996.

46. 1999 Survey of Dental Fees: American Dental Association; 2000.

47. Keim RG, Gottlieb EL, Nelson AH, Vogels DS, 3rd. 2005 JCO Orthodontic Practice Study. Part 1: trends. J Clin Orthod 2005;39:641-650.

48. Consumer Price Index- All Urban Consumers: Bureau of Labor Statistics.

49. Briggs AH. Handling uncertainty in cost-effectiveness models. Pharmacoeconomics 2000;17:479-500.

50. Black WC. The CE plane: a graphic representation of cost-effectiveness. Med Decis Making 1990;10:212-214.

51. van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects and C/E-ratios alongside a clinical trial. Health Econ 1994;3:309-319.

52. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. Health Econ 2001;10:779-787.

53. Tulloch JF, Phillips C, Proffit WR. Benefit of early Class II treatment: progress report of a two-phase randomized clinical trial. Am J Orthod Dentofacial Orthop 1998;113:62-72, quiz 73-64.