Treatment times of Class II malocclusion: four premolar and non-extraction protocols

Guilherme Janson, Danilo Pinelli Valarelli, Fabrício Pinelli Valarelli and Marcos Roberto de Freitas
Department of Orthodontics, Bauru Dental School, University of São Paulo, Bauru, Brazil

Correspondence to: Dr Guilherme Janson, Department of Orthodontics, Bauru Dental School, University of São Paulo, Alameda Octávio Pinheiro Brisolla 9-75, Bauru, SP 17012-901, Brazil. E-mail: jansong@travelnet.com.br

SUMMARY The purpose of this study was to retrospectively compare the treatment times of Class II division 1 malocclusion subjects treated with four first premolar extractions or a non-extraction protocol and fixed edgewise appliances. Eighty-four patients were selected and divided into two groups. Group 1, treated with four first premolar extractions, consisted of 48 patients (27 males and 21 females) with a mean age of 13.03 years and group 2, treated without extractions, consisted of 36 patients (18 males and 18 females) with a mean age of 13.13 years. Group 2 was subdivided into two subgroups, 2A consisting of 16 patients treated in one phase and 2B consisting of 20 patients treated in two phases. The initial and final Treatment Priority Index (TPI), initial ages, initial mandibular crowding, and treatment times of groups 1 and 2 were compared with t-tests. These variables were also compared between group 1 and the subgroups with analysis of variance followed by Tukey’s tests.

The treatment times for groups 1 and 2 and subgroups 2A and 2B were 2.36, 2.47, 2.25, and 2.64 years, respectively, which were not significantly different. Treatment times with non-extraction and four premolar extraction protocols are similar.

Introduction

Class II malocclusions can require different types of treatment when severity of the antero-posterior discrepancy, crowding, age, and patient compliance are considered (Rock, 1990; Bishara et al., 1995). Options for correction of Class II malocclusions include headgear, fixed and removable functional appliances, and fixed appliances with Class II elastics, extractions, and orthognathic surgery (Proffit et al., 1992; Aelbers and Dermaut, 1996). Most often, extractions can involve two maxillary premolars (Cleall and BeGole, 1982) or two maxillary and two mandibular premolars (Strang, 1950). It is known that the number of teeth extracted and malocclusion severity can influence treatment time (Vig et al., 1998; Chew and Sandham, 2000). Because malocclusion severity is an inherent characteristic that cannot be controlled, efforts have been made to quantify the influence of extractions on the duration of orthodontic treatment (Fink and Smith, 1992; O’Brien et al., 1995). Investigations comparing treatment times between unspecified malocclusions treated with and without extractions demonstrated that the number of teeth extracted is positively correlated with treatment time (Vig et al., 1990; Fink and Smith, 1992; Skidmore et al., 2006). This correlation was confirmed in complete Class II malocclusions where it was demonstrated that treatment time is shorter with a two maxillary premolar extraction protocol than with a four premolar extraction protocol (Janson et al., 2006). However, it did not hold true when it was found that in complete Class II malocclusions, treatment time is also shorter with a two maxillary premolar extraction protocol than with a non-extraction protocol, suggesting that in this malocclusion, the antero-posterior discrepancy severity and the treatment protocol play a major role in treatment time (Janson et al., 2007, 2008). According to this rationale, it was speculated that four premolar extraction and non-extraction protocols would result in a similar treatment time in complete Class II malocclusion patients because these two protocols require similar anchorage reinforcement and patient compliance to correct the molar relationship. Therefore, the objective of this study was to test the following null hypothesis: complete Class II malocclusion treatment times are similar with four premolar and non-extraction protocols.

Materials and methods

The study protocol was approved by the Ethics Committee of the University of São Paulo.

The sample was retrospectively selected from the files of over 4000 treated patients of the Orthodontic Department at Bauru Dental School, University of São Paulo. Records of all patients who initially had complete bilateral Class II malocclusions (molar relationship; Andrews, 1975; Wheeler et al., 2002) and were treated with four premolar extractions...
or non-extraction and fixed edgewise appliances were selected and divided into two groups. Sample selection was based exclusively on the initial antero-posterior dental relationship, regardless of any other dentoalveolar or skeletal characteristic. Additionally, all had the permanent teeth erupted up to the first molars, with no tooth agenesis, impacted or supernumerary teeth, cleft lip and/or palate, and the groups were matched regarding age.

Group 1 consisted of 48 patients (27 males and 21 females), with a complete Class II malocclusion, treated with four premolar extractions, with an initial mean age of 13.03 years (range 10.67–18.33 years). All had a Class II division I malocclusion.

Group 2 comprised 36 patients (18 males and 18 females), with complete Class II malocclusions, treated without extractions, with an initial mean age of 13.13 years (range 9.40–16.04 years). All had a Class II division I malocclusion. Twenty of the 36 patients treated without extractions received two-phase treatment and used the combined headgear-activator appliance for 11.2 ± 6 months before treatment with fixed appliances. In view of the controversy concerning the influence of two-phase Class II malocclusion treatment on outcome and treatment time (Livieratos and Johnston, 1995; Tulloch et al., 1997; Beckwith et al., 1999), group 2 was divided into two subgroups. Subgroup 2A consisted of 16 patients treated in one phase (fixed appliances) and subgroup 2B comprised 20 patients treated in two phases (functional followed by fixed appliances).

The mechanics used with the fixed edgewise appliances included 0.022 × 0.028 inch brackets, associated with extraoral headgear to correct the malocclusion in the non-extraction patients, and to reinforce anchorage for the maxillary teeth in the extraction patients. Class II elastics were also used when applicable to aid in correcting the Class II antero-posterior relationship. No Tweed anchorage wire adjustments or temporary anchorage devices were included. The usual wire sequence began with a 0.015 inch twist-flex or 0.014 inch nitinol wire, followed by 0.016, 0.018, 0.020 inch, and finally a 0.021 × 0.025 or 0.018 × 0.025 inch stainless steel archwire (UniTek, Monrovia, California, USA). In the extraction protocol, the canines were initially retracted a small amount to allow space for levelling and aligning of the anterior teeth. The anterior teeth were retracted en masse with sliding mechanics in the rectangular archwire, after levelling and aligning. If a 0.021 × 0.025 inch wire was used, the dimensions of the rectangular wire were electrolytically reduced in the posterior segments to reduce the friction forces with the brackets and tubes. The canines and anterior teeth were retracted with elastomeric chain. Deep overbites were usually corrected by reversing and accentuating the curve of Spee of the mandibular and maxillary stainless steel archwires until overcorrection was obtained. This overcorrection was also maintained by accentuating and reversing the curve of Spee in the rectangular wire.

From the patient records, the following information was obtained: initial age, gender, date of treatment onset, date of treatment completion, and total treatment time. To evaluate the initial malocclusion severity compatibility of the groups, the Treatment Priority Index (TPI; Grainger, 1967) and the amount of mandibular crowding were blindly calculated on the pre-treatment dental study models of each patient. The final occlusal results of the groups also had to be compatible. Therefore, the final TPI of the groups was also calculated on the post-treatment dental study models. The TPI provides weighted subscores for overjet, vertical overbite or open bite, tooth displacement, and posterior crossbite, as well as summary scores reflecting the overall severity of the malocclusion. With the exception of rotation and displacement, all TPI components were measured along a continuous scale from positive to negative values. Thus, mandibular overjet and open bite were entered as negative overjet and negative overbite, respectively. A constant corresponding to the first molar relationship was added to the TPI score. Total TPI scores range from 0 to 10 or more, with higher scores representing more severe malocclusions (Corruccini and Potter, 1980; Corruccini and Whitley, 1981).

TPI components were defined as follows:

- Overjet: anterior distance from the most mesial part of the labial surface of the maxillary central incisor to the labial surface of the opposing mandibular incisor, measured perpendicular to the coronal plane.
- Overbite or open bite: with the dental models in centric occlusion, the amount of vertical overlap of the maxillary central incisor over the mandibular central incisor, taken as a ratio of the total crown height (cervical to incisal edge) of the mandibular incisor.
- Tooth displacement: the sum of the number of teeth noticeably rotated or displaced from ideal alignment, plus twice the number of teeth rotated more than 45 degrees or displaced more than 2 mm.
- First molar relationship: a constant comprising the severity of the malocclusion, based on the relationship between the maxillary and mandibular first molars.
- Posterior crossbite: buccolingual deviation in occlusion of the post-canine teeth. The measurement is positive for buccal crossbite (first molar positioned too far to the buccal side) or negative for lingual crossbite. Crossbite is also scored as the number of teeth deviating from ideal cusp-to-fossa fit or cusp-to-cusp relationship.
- Mandibular crowding of the initial dental study models was calculated as the difference between arch length (circumference from left to right first molars) and the sum of tooth widths from first molar to first molar, in millimetres. In a well-aligned arch, arch length was equal to the sum of the tooth widths. Negative values indicated crowding.
**Error study**

Twenty pairs of dental study models were randomly remeasured by the same examiner (DPV) for TPI evaluations. The casual error for each was calculated according to the formula of Dahberg (1940): $S^2 = \frac{\sum d^2}{2n}$, where $S^2$ is the error variance and $d$ is the difference between two determinations of the same variable. The systematic errors were evaluated with dependent $t$-tests (Houston, 1983) at a significance level of $P < 0.05$.

**Statistical analyses**

Means and standard deviations for each variable were calculated to enable characterization of the groups. Normal distribution evaluated with Kolmogorov–Smirnov tests showed that all variables were normally distributed in the groups. A chi-square test was used to evaluate compatibility regarding gender distribution in the groups and $t$-tests to compare the initial and final TPIs, initial age, mandibular crowding, and the treatment times of groups 1 and 2. Group 1 and subgroups 2A and 2B were compared with analysis of variance, followed by Tukey’s tests. The results were considered significant at $P < 0.05$.

**Results**

There were no systematic errors in intraexaminer TPI evaluation and the casual error was within acceptable levels (Table 1). The groups were compatible regarding gender distribution (group 1: 27 males and 21 females and group 2: 18 males and 18 females; $P = 0.323$), initial and final TPIs, and initial age (Table 2). However, group 1 had significantly more crowding than group 2. Despite this difference, treatment time was similar in both groups. Additionally, when comparing group 1 with subgroups 2A and 2B, treatment time was similar (Table 3).

Because the groups were significantly different regarding crowding, patients with moderate to severe crowding (crowding greater than 3 mm) in group 1 were excluded in order to match the groups regarding this variable. This resulted in subgroup 1 with 23 patients that were compared with group 2. The results demonstrated that all the other variables were matched and treatment time continued to be similar (Table 4).

**Discussion**

**Sample selection**

The subjects were selected on the basis of having complete bilateral Class II malocclusions, independent of the associated cephalometric skeletal characteristics. Thus, it would be expected that the skeletal characteristics would be evenly distributed. Usually, it is not the skeletal characteristics of a Class II malocclusion that primarily determine whether it should be treated with or without extractions but rather the dentoalveolar characteristics (Russell, 1994; Bryk and White, 2001; Janson et al., 2008). Nevertheless, the similarity in the initial malocclusion severity points towards cephalometric compatibility of the groups (Keeling et al., 1989).

**Group compatibility**

The statistically significant difference in the initial crowding between the groups might have influenced treatment time. For this reason, the groups were divided into compatible subgroups regarding this variable and compared again (Table 2).

**Treatment time**

Treatment time in the four premolar extraction and in the non-extraction groups was similar, even when group 1 was compared with subgroups 2A and 2B or when the groups were matched regarding the amount of mandibular crowding (Tables 2–4). The treatment time difference between groups

---

**Table 1** Results of intraexaminer errors.

<table>
<thead>
<tr>
<th>Variable</th>
<th>First measurement, $n = 20$</th>
<th>Second measurement, $n = 20$</th>
<th>$P$</th>
<th>Dahlberg</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Initial Treatment Priority Index</td>
<td>7.71</td>
<td>1.01</td>
<td>7.67</td>
<td>0.89</td>
</tr>
</tbody>
</table>

**Table 2** Results of $t$-tests between groups 1 and 2.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group 1, four premolar extractions, $n = 48$</th>
<th>Group 2, non-extraction, $n = 36$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Initial Treatment Priority Index</td>
<td>8.22</td>
<td>1.09</td>
<td>8.02</td>
</tr>
<tr>
<td>Final Treatment Priority Index</td>
<td>1.90</td>
<td>1.76</td>
<td>1.43</td>
</tr>
<tr>
<td>Initial age (years)</td>
<td>13.03</td>
<td>1.70</td>
<td>12.59</td>
</tr>
<tr>
<td>Initial mandibular crowding (mm)</td>
<td>3.48</td>
<td>2.69</td>
<td>0.44</td>
</tr>
<tr>
<td>Treatment time (years)</td>
<td>2.36</td>
<td>0.63</td>
<td>2.47</td>
</tr>
</tbody>
</table>

*$P < 0.05$. 
Table 3 Comparison between group 1 (four premolar extractions), subgroup 2A (one-phase non-extraction), and subgroup 2B (two-phase non-extraction), with analysis of variance and Tukey’s tests.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group 1, n = 48</th>
<th>Subgroup 2A, n = 16</th>
<th>Subgroup 2B, n = 20</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial Treatment Priority Index</td>
<td>8.22 (1.09)*</td>
<td>7.91 (1.78)*</td>
<td>8.11 (1.23)*</td>
<td>0.703</td>
</tr>
<tr>
<td>Final Treatment Priority Index</td>
<td>1.90 (1.76)*</td>
<td>1.35 (1.49)*</td>
<td>1.49 (1.66)*</td>
<td>0.447</td>
</tr>
<tr>
<td>Initial age (years)</td>
<td>13.03 (1.70)*</td>
<td>13.15 (1.38)*</td>
<td>12.14 (1.37)*</td>
<td>0.078</td>
</tr>
<tr>
<td>Initial mandibular crowding (mm)</td>
<td>3.48 (2.69)*</td>
<td>0.75 (1.81)*</td>
<td>0.19 (0.65)*</td>
<td>0.000*</td>
</tr>
<tr>
<td>Treatment time (years)</td>
<td>2.36 (0.73)*</td>
<td>2.25 (0.73)*</td>
<td>2.64 (0.89)*</td>
<td>0.221</td>
</tr>
</tbody>
</table>

Same letters horizontally indicate no significant difference.
*P < 0.05.

Table 4 Comparison between compatible groups regarding crowding (t-tests).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Subgroup 1, four premolar extractions, n = 23</th>
<th>Group 2, non-extraction, n = 36</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>SD</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td>Initial Treatment Priority Index</td>
<td>7.75</td>
<td>8.02</td>
<td>0.425</td>
</tr>
<tr>
<td>Final Treatment Priority Index</td>
<td>1.85</td>
<td>1.43</td>
<td>0.358</td>
</tr>
<tr>
<td>Initial age (years)</td>
<td>13.13</td>
<td>12.59</td>
<td>0.189</td>
</tr>
<tr>
<td>Initial mandibular crowding (mm)</td>
<td>1.13</td>
<td>0.44</td>
<td>0.063</td>
</tr>
<tr>
<td>Treatment time (years)</td>
<td>2.52</td>
<td>2.47</td>
<td>0.794</td>
</tr>
</tbody>
</table>

1 and 2 was 0.11 years (1.3 months), which can also be regarded as not clinically meaningful.

Reports stating that two-phase Class II malocclusion treatment increases treatment time suggest that this protocol would have influenced treatment time in the non-extraction group (Livieratos and Johnston, 1995; Tulloch et al., 1997; Beckwith et al., 1999). However, when group 2 was subdivided into one- and two-phase subgroups and compared with group 1, treatment time remained similar. It is interesting to note that subgroup 2B had a non-significant lower initial age than group 1 and subgroup 2A (Tables 3 and 4). A younger initial age within this range tends to facilitate treatment and therefore would imply a shorter treatment time (Harris et al., 1991; Tulloch et al., 1997), but because these patients were treated in two phases, treatment time was similar to the other group and subgroup. Reported longer treatment times for two-phase treatment included patients at an even younger age range (Vig et al., 1990; Gianelly, 1995). Although the findings of the comparisons with these subgroups showed that two-phase treatment did not influence treatment time in the non-extraction group, they have to be considered with caution due to the relatively small number of patients in these subgroups.

The amount of mandibular crowding did not influence treatment time between the groups, which continued to show similarity (Table 4). These results suggest that Class II molar relationship correction, which was necessary in both groups, is an important variable for the similarity in treatment time between the groups. Patients must comply when using extraoral headgear, a headgear-activator or Class II elastics to correct Class II antero-posterior discrepancies (Bryk and White, 2001; Janson et al., 2004). If the necessary compliance level is not achieved, the occlusal results are compromised (Janson et al., 2003) and treatment time will be increased (Chew and Sandham, 2000).

Previous studies that compared treatment times in extraction and non-extraction subjects but did not distinguish between Class I and Class II malocclusions concluded that the number of extracted teeth is positively correlated with treatment time (Vig et al., 1990; Fink and Smith, 1992). This correlation was confirmed in complete Class II malocclusions where it was demonstrated that treatment time is shorter with a two maxillary premolar extraction protocol than with a four premolar extraction protocol (Janson et al., 2006). However, when Class II treatment efficiency in two maxillary premolar extraction and non-extraction protocols was compared, it was concluded that treatment time with the non-extraction protocol was longer. It was also demonstrated that molar relationship correction influenced treatment time (Janson et al., 2007). It seems that there is a mechanical similarity between the four premolar extraction protocol and the non-extraction protocol, which consists of correcting the Class II molar relationship. Therefore, it seems that it is this aspect in the two protocols, which results in a similar treatment time.
The treatment times for both groups in this study were longer than those reported in the literature (Alger, 1988; Amditis and Smith, 2000). This may be explained by the fact that the patients were treated by postgraduate students, whereas in the other investigations, they were treated in private practice. It is speculated that treatment times are usually shorter in the hands of more experienced clinicians (McGuinness and McDonald, 1998; Chew and Sandham, 2000).

It should be emphasized that the present results were obtained using orthodontic mechanics without the use of temporary anchorage devices, which have significantly improved and enlarged treatment horizons (Gelgor et al., 2007; Lai et al., 2008). If they are capable of simplifying both treatment protocols investigated, the results would probably be similar. However, this will require further investigation. As this present study was retrospective, with its inherent limitations, future prospective studies would be required to confirm these results.

Clinical implications

The similarity in treatment time between these complete Class II malocclusion treatment protocols, in conjunction with previous reports (Janson et al., 2006, 2007), shows that treatment protocol plays a primary role in determining treatment time rather than the number of extractions in Class II malocclusion subjects. Correction of the Class II molar relationship seemed to be the common denominator, which contributed to the similar treatment time with both protocols. Therefore, selecting a two maxillary premolar extraction protocol, in which correction of a Class II molar relationship is not necessary, might fulfill the requirement of treatment in a shorter period of time (Kessel, 1963; Janson et al., 2007).

Conclusions

The null hypothesis was accepted because treatment times with four premolar extraction or with non-extraction protocols of complete Class II malocclusions were similar.

Funding

Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), project number 04/12445-8.

References


Amditis C, Smith L F 2000 The duration of fixed orthodontic treatment: a comparison of two groups of patients treated using edgewise brackets with 0.018” and 0.022” slots. Australian Orthodontic Journal 16: 34–39


Dahlberg G 1940 Statistical methods for medical and biological students. Interscience, New York


Kessel S P 1963 The rationale of maxillary premolar extraction only in Class II therapy. American Journal of Orthodontics 49: 276–293


Russell D M 1994 Extractions in support of orthodontic treatment. NDA Journal 45: 15–19


Strang R 1950 A textbook of orthodontics. Lea & Febiger, Philadelphia


