Long-term profile attractiveness in Class II Division 1 malocclusion patients treated with and without extractions

Lucas Marzullo Mendes,[a] Guilherme Janson,[a] Cintia Helena Zingaretti Junqueira-Mendes,[a] and Daniela Gamba Garib[b]
Bauru, São Paulo, Brazil

Introduction: This study evaluated the influence of Class II treatment protocols in profile attractiveness. **Methods:** Sixty-eight patients with initial full Class II Division 1 malocclusion, orthodontically treated an average of 15 years before, and matched by sex, age, time after treatment, orthodontic outcomes, and overjet, were compared. Three groups were formed, according to the treatment protocol: NE, nonextraction (n = 20; mean age 29.94 years, 15.62 years after treatment); 2E, 2-premolar extractions (n = 27; mean age 30.56 years, 15.09 years after treatment); and 4E, 4-premolar extractions (n = 21; mean age 32.29 years, 17.20 years after treatment). Cephalometric measurements and profile silhouettes were obtained from posttreatment and long-term posttreatment lateral cephalograms. With a 10-point numeric scale, 77 orthodontists and 77 laypeople rated profile attractiveness of each silhouette on a website. **Results:** The raters' posttreatment and long-term posttreatment scores, respectively, were: NE 4.76 and 4.32; 2E 5.35 and 5.08; and 4E 4.53 and 4.33. **Conclusions:** The posttreatment profile attractiveness was significantly higher in the 2E than in the 4E group, and in the NE group it was similar to the others. The long-term profile attractiveness in the 2E group was significantly greater compared with the NE and 4E groups. (Am J Orthod Dentofacial Orthop 2019;155:362-71)

The debate on dental extractions for orthodontic reasons began more than 100 years ago.[1,2] Besides problems such as severe crowding and dental protrusion, extractions may solve anteroposterior discrepancies, such as Class II or III malocclusions.[3] But the long-term stability of extraction treatment and its effects on facial profile still concern orthodontists.[4]

Regarding Class II malocclusion, nonextraction treatment has an efficiency similar to a 4-premolar extractions protocol, sharing the challenge of molar relationship correction.[5] The 2-premolar extractions treatment requires shorter treatment time than nonextraction or 4-premolar extractions because the first molars are kept in a full Class II relationship.[6]

[a]Department of Orthodontics, Bauru Dental School, University of São Paulo, Bauru, São Paulo, Brazil.
[b]Department of Orthodontics, Bauru Dental School, and Hospital for Rehabilitation of Craniofacial Anomalies, University of São Paulo, Bauru, São Paulo, Brazil.
All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest, and none were reported.
Address correspondence to: Lucas M. Mendes, Department of Orthodontics, Bauru Dental School, University of São Paulo, Alameda Octávio Pinheiro Brisolla 9-75, Bauru, SP, 17012-901, Brazil; e-mail, lucas@lucasmendes.com.br.
Submitted, February 2018; revised and accepted, April 2018.
0889-5406/$36.00

https://doi.org/10.1016/j.ajodo.2018.04.030

The soft-tissue effects of nonextraction and 2- and 4-premolar extractions are similar in the long term.[2,7,8] However, there is still concern with the long-term esthetic effects of premolar extractions on the facial profile.[4]

The lips are quite important in the context of facial attractiveness, which is also influenced by the aging process. It has been suggested that premolar extractions could impair lip support, contributing to a straight or "dished in" profile.[9]

Because nowadays esthetics is a general concern, it is worthwhile to clarify if extraction treatment really may esthetically impair the soft tissue profile, especially in the long term, which still lacks consistent scientific information. In this context, the present study aimed to compare profile attractiveness in patients with Class II Division 1 malocclusion treated with nonextraction or 2 or 4 premolar extractions at the posttreatment stage and in the long term after treatment.

MATERIAL AND METHODS

This study was approved by the Ethics in Research Committee at Bauru Dental School, University of São Paulo, number 033/2011.

Sample size was calculated as follows: To detect a difference of 1.0 mm in the distance from the upper

lip to the Steiner S-line (UL-Pog'Cm) with a standard deviation of 1.1 among the 3 groups, with an α error of 5% and a test power of 80%, 20 patients would be necessary in each group.[2]

Therefore, a retrospective sample of 68 phenotypically Mediterranean adults was selected according to the following inclusion criteria: adult (≥21 years of age); initial full cusp Class II Division 1 malocclusion; orthodontically treated without extractions or with 2- or 4-premolar extractions; initial and final orthodontic records in good conditions; long-term cephalograms taken ≥8 years after treatment; and no anterior tooth loss in the long-term period.

Patients were divided into 3 groups according to the treatment protocol: nonextraction (n = 20), 2 maxillary premolar extractions (n = 27), and 4-premolar extractions (n = 21).

Treatments were performed with fixed edgewise appliances, with 0.022 × 0.028-inch conventional brackets. Wire sequences including initial 0.015 Twist-Flex or 0.016-inch Nitinol, followed by 0.016, 0.018, 0.020, and 0.018 × 0.025 or 0.019 × 0.025-inch stainless steel archwires (3M Unitek, Monrovia, Calif) were used during treatment. Accentuated and reversed curves of Spee were used to correct deep overbites. In the extraction groups, for overjet and Class II canine correction the anterior teeth were retracted "en masse" with the use of rectangular archwires.

In the nonextraction and in the 4-premolar extraction groups, extraoral headgear was used to correct the Class II anteroposterior relationship. In the nonextraction group, some patients also used functional appliances only or associated with extraoral headgear. In the 2-premolar extraction group, Class II molar relationship was maintained with extraoral headgear during anterior retraction. Class II elastics helped to correct molar relationship in the nonextraction and 4-premolar extraction groups, and helped to preserve Class II molar relationship in the 2-premolar extraction group. For retention, a maxillary Hawley plate was recommended to be used full time, except during meals, for 6 months and an additional 6 months only at night. In the mandible, a canine-to-canine fixed retainer was installed to be used for 3 years.

To be compared, the groups had to present similar occlusal finishing. Therefore, to evaluate the occlusal finishing of the patients in each group, the Objective Grading System (OGS) index was used.[10]

The posttreatment (T1) and long-term posttreatment (T2) lateral head films were obtained in centric occlusion, with passive lip posture. All of them were scanned and the images digitally stored.

With the use of Dolphin Imaging 11.5 software (Dolphin Imaging and Management Solutions, Chatsworth, Calif), the following cephalometric variables were measured: Prn-Sn, Prn-Pog′Ls, UL-Pog′Cm, LL-Pog′Cm, UL-Pog′Prn, LL-Pog′Prn, UL-Pog′Sn, LL-Pog′Sn, Cm.Sn.Ls, Li.B′.Pog′, Prn-Nperp, ANS-Sn, A-A′, UL-Ulp, LL-Llp, B-B′, Pog-Pog′, N′.Sn.Pog′, N′.Prn.Pog′, N.A.P, ULPog′.PoO, Pog′Ls.NB, NS.Gn, SN.GoGn, and PoOr.GoMe.[11]

The magnification factors of the head films ranged from 6.0% to 9.8%, which were corrected by the software. The changes that occurred in the posttreatment period were calculated as T2 − T1.

Profile silhouette images were obtained from the digitized cephalograms with the use of Dolphin Imaging 11.5. To minimize the influence of the head position, all images were standardized with the Frankfort plane horizontally oriented. To exclusively evaluate profile attractiveness and eliminate the influence of features such as age, skin color, hair style, and other individual attributes, the contour of the soft-tissue profile of each patient was edited with the use of Photoshop CS5 for Mac software. The images were colored in black, aiming to reproduce only a shadow of the profile. The resulting image was a black profile on a white background with the Frankfort plane horizontally oriented (Fig 1).

Assessment of profile attractiveness was conducted through an internet site. Each evaluator received an username and password that provided personal and nontransferable unlimited access.[12]

A simple and straightforward brief explanation of the assessment page resources was provided, such as operation of the navigation arrows and how to begin and end the assessment. The evaluator was instructed to view all profiles before starting analysis of the sample.[4]

Sample randomization was automatically performed and the images displayed in randomly pulled pairs,[2] in a different order for each rater. They were not able to know which group the profile image came from or if it was a posttreatment or a long-term posttreatment profile image. Profile attractiveness was rated on a 10-point visual analog scale (VAS), where 1 represented "the most unattractive profile" and 10 represented "the most attractive profile." The evaluator could alter the scores at any time before submitting them.[12]

A sample size calculation helped to determine the adequate number of raters. To detect a difference of 0.6 point on a scale of 1 to 10 among the 3 groups of patients, with a standard deviation of 1.28 points,[13] at a significance level of 0.05 and a test power of 80%, 72 evaluators would be necessary. Therefore, 77 laypeople and 77 orthodontists were selected.

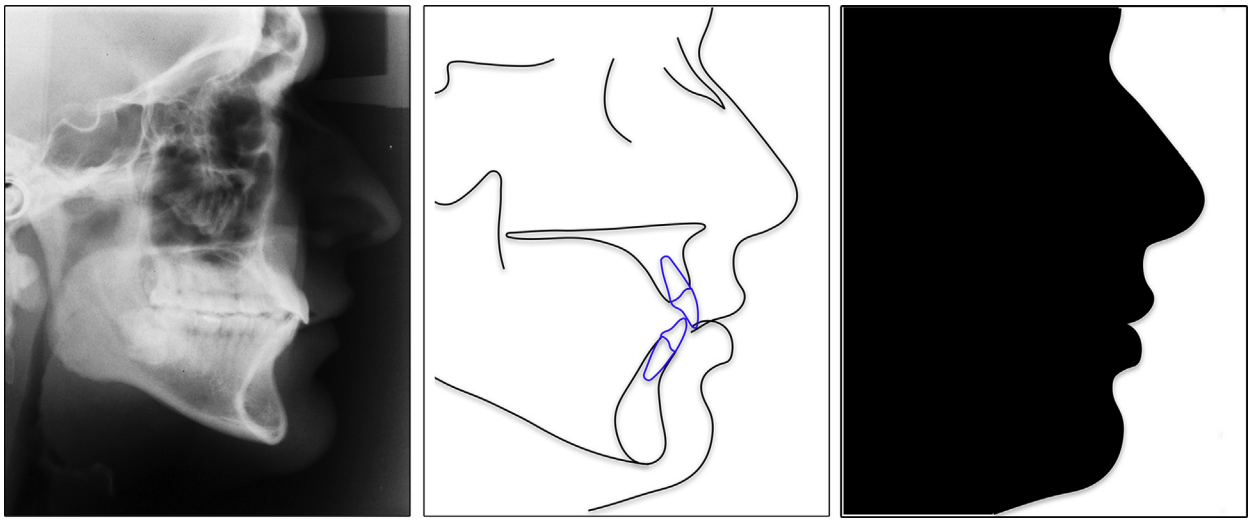

Fig 1. Steps performed to achieve a black shadow of the patient's profile.

Groups of laypeople and orthodontist raters were selected.[14] The laypeople were defined as persons without formal education in dentistry or dental hygiene. The orthodontists were dental surgeons with completed graduate degrees in orthodontics.

Two weeks after the first evaluation, 41 randomly selected cephalograms were retraced by the same examiner. In a different arrangement, the whole sample was reevaluated by 47 laypeople and orthodontists. The random errors were evaluated with the use of the Dahlberg formula, where S^2 is the error variance and d is the difference between 2 determinations of the same variable.[15] The systematic errors were evaluated with the use of dependent *t* tests.[16]

Statistical analyses

Normal distribution of all quantitative variables was evaluated by means of Shapiro-Wilk tests. The variables without normal distribution were long-term posttreatment time, nasal projection, subnasal thickness, upper lip thickness, lower lip thickness, facial convexity angle including nose, Steiner mandibular plane inclination, overjet, orthodontists' score, laypeople's score, and evaluators' age. All other variables had normal distribution.

Comparability among the 3 groups regarding sex was evaluated with the use of chi-square test. Regarding posttreatment age, long-term posttreatment age, postretention time, quality of occlusal outcomes, and posttreatment and long-term posttreatment overjets, comparability was evaluated with the use of analysis of variance (ANOVA). The posttreatment and long-term posttreatment cephalometric statuses as well as the changes in the posttreatment period were compared with the use of ANOVA. Sex and age distribution betweeen rater groups was evaluated with the use of chi-square and Mann-Whitney tests, respectively.

To evaluate if profile attractiveness was influenced by the different treatment protocols, type of rater, or their interaction, 2-way ANOVA was applied among the 3 groups of patients and the 2 categories of raters at the posttreatment and long-term posttreatment stages.

All tests were conducted with the use of Statistica 10.0 software (Statsoft, Tulsa, Okla). Results were considered to be significant at $P < 0.05$.

RESULTS

Cephalometric measurements showed random errors varying from 1.22mm (ANS-Sn) to 3.77° (Li.B'.Pog'). Only UL-Ulp, LL-Llp, UL-Pog'Prn, and Cm.Sn.Ls showed significant systematic errors. There were no random or systematic errors in the assessment of profile attractiveness.

The groups were similar regarding sex, posttreatment and long-term posttreatment age, postretention time, OGS index, and posttreatment and long-term posttreatment overjets (Table I). However, at the posttreatment stage, the 4E group had a significantly more vertical growth pattern than the nonextraction group (Table II). At the long-term posttreatment stage, the nonextraction group had a significantly greater nasal length than the 2E group (Table III). Changes in the long-term posttreatment period were similar among the groups (Table IV).

The lay group had significantly fewer men and more women, and were significantly younger than the orthodontist group (Table V).

Table I. Descriptive analysis and comparability among the 3 groups regarding sex distribution, posttreatment and long-term age, postretention time, OGS index, and posttreatment and long-term overjet

Variable	*NE group (n = 20)*	*2E group (n = 27)*	*4E group (n = 21)*	P *value*
Posttreatment age (y)	14.43 ± 3.48	15.30 ± 1.78	15.01 ± 1.40	0.448*
Long-term age (y)	29.94 ± 7.35	30.56 ± 5.74	32.29 ± 5.70	0.455*
Long-term posttreatment period (y)	15.62 ± 7.22	15.09 ± 4.79	17.20 ± 5.76	0.316*
Objective Grading System index	41.55 ± 17.61	41.00 ± 9.32	46.38 ± 12.78	0.338*
Posttreatment overjet	2.81 ± 1.41	2.42 ± 0.81	2.79 ± 0.89	0.350*
Long-term posttreatment overjet	3.47 ± 2.03	2.73 ± 0.86	3.30 ± 1.09	0.150*
Female	11 (55.00%)	13 (48.15%)	6 (28.57%)	0.202†
Male	9 (45.00%)	14 (51.85%)	15 (71.43%)	

Values are presented as mean ± SD or n (%).
*ANOVA; †Chi-square test.

Table II. Comparison among the 3 groups regarding the cephalometric variables at the posttreatment stage (T1)

Variable	*NE group (n = 20)*	*2E group (n = 27)*	*4E group (n = 21)*	P *value*
Nasal measures (mm)				
Prn-Sn	14.72 (1.45)	13.85 (1.54)	14.42 (1.87)	0.181†
Prn-Pog′Ls	7.19 (5.19)	5.63 (3.75)	5.37 (5.21)	0.404†
Distance from lips to esthetics lines (mm)				
UL-Pog′Cm	−1.03 (2.60)	−0.38 (1.87)	−0.12 (2.67)	0.455†
LL-Pog′Cm	−0.40 (3.19)	0.72 (2.22)	1.19 (2.73)	0.160†
UL-Pog′Prn	−4.11 (3.00)	−3.30 (2.18)	−3.15 (3.11)	0.488†
LL-Pog′Prn	−1.90 (3.46)	−0.71 (2.43)	−0.38 (3.15)	0.236†
UL-Pog′Sn	2.99 (2.24)	3.39 (1.63)	3.73 (2.26)	0.510†
LL-Pog′Sn	1.58 (2.86)	2.55 (1.99)	3.14 (2.26)	0.107†
Soft tissue angular measurements (°)				
Cm.Sn.Ls	108.22 (9.71)	111.04 (11.12)	108.65 (11.06)	0.615†
Li.B′.Pog′	114.52 (16.74)	118.30 (13.86)	122.20 (12.36)	0.238†
Soft tissues thickness (mm)				
Prn-Nperp	18.60 (1.67)	18.18 (1.74)	18.25 (1.66)	0.681†
ANS-Sn	14.95 (2.33)	15.25 (2.36)	14.58 (2.51)	0.635†
A-A′	15.03 (2.10)	14.99 (2.31)	14.81 (1.90)	0.942†
UL-Ulp	18.26 (3.13)	19.59 (2.13)	19.55 (3.21)	0.219†
LL-Llp	12.63 (1.59)	13.43 (1.62)	13.08 (2.00)	0.297†
B-B′	11.36 (1.40)	12.00 (2.10)	11.45 (1.20)	0.353†
Pog-Pog′	11.70 (1.86)	12.08 (2.03)	12.22 (1.73)	0.658†
Convexity measures (°)				
n′.Sn.Pog′	128.47 (4.50)	127.79 (3.28)	130.07 (4.27)	0.147†
n′.Prn.Pog′	160.14 (6.03)	157.33 (4.35)	159.66 (5.05)	0.132†
n.A.P	4.52 (6.06)	6.69 (4.45)	6.70 (5.97)	0.331†
ULPog′.PoO	73.97 (13.39)	71.13 (8.87)	67.73 (9.76)	0.179†
Pog′Ls.NB	10.10 (6.30)	12.24 (4.20)	11.71 (4.52)	0.340†
Growth pattern (°)				
NS.Gn	67.67 (2.73)	69.50 (3.68)	69.97 (4.12)	0.101†
SN.GoGn	30.42[a] (4.49)	32.80[a,b] (5.46)	34.53[b] (5.62)	0.026‡*
PoOr.GoMe	25.02[a] (5.37)	25.61[a] (4.68)	29.68[b] (5.32)	0.037‡*

[a,b]Values in the same row with different superscript letters are significantly different.
*Statistically significant at $P < 0.05$; †ANOVA; ‡Kruskal-Wallis test.

At the posttreatment stage, profile attractiveness was statistically greater in the 2E group than in the 4E group, and in the NE group it was similar to the others. At the long-term posttreatment stage, profile attractiveness was statistically greater in the 2E group than in the other groups (Tables VI and VII). In both stages, laypeople were more critical than orthodontists. There was interaction between treatment group and evaluator group only in the posttreatment stage (Table VI).

DISCUSSION

To select the sample, a minimum initial severity of full bilateral Class II malocclusion was established. This

Table III. Comparison among the 3 groups regarding the cephalometric variables at the long-term posttreatment stage (T2)

Variable	*NE group (n = 20)*	*2E group (n = 27)*	*4E group (n = 21)*	P *value*
Nasal measures (mm)				
Prn-Sn	16.65[b] (1.44)	15.33[a] (1.88)	16.57[a,b] (1.91)	0.023[‡*]
Prn-Pog′Ls	10.51 (5.00)	8.41 (4.23)	9.64 (5.96)	0.363[†]
Distance from lips to esthetics lines (mm)				
UL-Pog′Cm	−2.30 (2.48)	−1.44 (2.09)	−1.78 (2.87)	0.502[†]
LL-Pog′Cm	−1.79 (2.99)	−0.60 (2.47)	−0.99 (3.79)	0.426[†]
UL-Pog′Prn	−5.90 (2.92)	−4.87 (2.51)	−5.44 (3.40)	0.488[†]
LL-Pog′Prn	−3.75 (3.32)	−2.46 (2.82)	−3.23 (4.39)	0.454[†]
UL-Pog′Sn	2.44 (2.13)	3.00 (1.78)	2.88 (2.34)	0.635[†]
LL-Pog′Sn	0.78 (2.59)	1.79 (2.10)	1.85 (3.07)	0.322[†]
Soft tissue angular measurements (°)				
Cm.Sn.Ls	105.81 (9.37)	106.80 (10.39)	103.48 (11.34)	0.543[†]
Li.B′.Pog′	118.43 (17.05)	120.08 (14.61)	125.47 (12.79)	0.281[†]
Soft tissues thickness (mm)				
Prn-Nperp	20.10 (1.91)	19.47 (1.85)	19.73 (2.02)	0.548[†]
ANS-Sn	14.41 (2.33)	13.99 (3.49)	13.42 (7.12)	0.795[†]
A-A′	14.67 (2.40)	14.66 (2.93)	15.20 (2.19)	0.728[†]
UL-Ulp	17.48 (2.96)	18.58 (2.07)	18.42 (1.72)	0.233[†]
LL-Llp	12.51 (1.81)	13.33 (1.28)	12.97 (1.91)	0.248[†]
B-B′	11.69 (1.42)	12.04 (1.67)	11.98 (1.20)	0.709[†]
Pog-Pog′	12.50 (2.28)	13.03 (2.40)	13.84 (1.47)	0.130[†]
Convexity measures (°)				
n′.Sn.Pog′	126.13 (4.90)	126.77 (4.13)	128.02 (4.39)	0.385[†]
n′.Prn.Pog′	161.54 (7.31)	159.68 (5.74)	161.81 (5.98)	0.442[†]
n.A.P	4.77 (5.46)	6.89 (5.16)	5.74 (6.51)	0.448[†]
ULPog′.PoO	78.80 (12.43)	77.14 (9.19)	75.39 (13.45)	0.643[†]
Pog′Ls.NB	7.56 (6.89)	9.29 (4.75)	7.93 (6.06)	0.560[†]
Growth pattern (°)				
NS.Gn	67.16 (3.17)	68.96 (3.94)	69.46 (4.63)	0.153[†]
SN.GoGn	29.39 (4.35)	31.52 (5.88)	32.25 (7.13)	0.278[†]
PoOr.GoMe	23.85 (4.74)	23.91 (5.29)	27.11 (7.34)	0.119[†]

[a,b]Values in the same row with different superscript letters are significantly different.
*Statistically significant at $P < 0.05$; [†]ANOVA; [‡]Kruskal-Wallis test.

restriction made the sample representative of an "extreme" Class II malocclusion correction, because full cusp severity is often associated with severe overjets.[5,7,17]

Understanding patients' expectations is essential to offer appropriate treatments. If the orthodontist's perception does not match the patient's, the result may not be accepted, even with improved function. Patient opinion was represented in this study by the lay raters. Some studies have found a statistically significant difference in attractiveness as perceived by laypeople versus orthodontists.[18] Therefore, both types of evaluators were used and separately evaluated the groups.[14]

Androgynous silhouettes have been defended and used by previous authors for profile esthetics evaluation, because they eliminate other possible variables, such as skin, hair, and eyes.[19] Elimination of these influences allows the evaluators to focus on the profile, which had its position standardized on the horizontally oriented Frankfort plane.[20]

The survey conducted through a website enabled a randomized order of evaluations for each rater.[12] They were able to observe the images as many times as they wanted, revise their assessments if necessary, independently, and without time constraint.[21] A 10-point VAS allowed raters to easily express their subjective opinions.

The instruction for raters to observe all gallery images before starting the attractiveness judgment may explain the reproducibility in assessment and the reduced methodologic errors, because they had the chance to calibrate themselves according to the sample they were about to evaluate.[4]

The groups were matched regarding age in both periods evaluated (Table I). A different age distribution could interfere with profile attractiveness, because cephalometric differences occur in patients' faces with time.[22]

Several investigations on attractiveness described the occlusal relationship only as "good," with "dental alignment," or "without anomalies," or they did not mention

Table IV. Comparison among the 3 groups regarding the cephalometric changes during the long-term posttreatment period (T2 − T1)

Variable	*NE group (n = 20)*	*2E group (n = 27)*	*4E group (n = 21)*	P *value**
Nasal measures (mm)				
Prn-Sn	1.93 (1.58)	1.49 (1.31)	2.15 (1.46)	0.270
Prn-Pog′Ls	3.32 (1.85)	2.77 (2.43)	4.27 (3.87)	0.196
Distance from lips to esthetics lines (mm)				
UL-Pog′Cm	−1.27 (0.76)	−1.06 (1.24)	−1.65 (1.76)	0.306
LL-Pog′Cm	−1.39 (1.50)	−1.32 (1.58)	−2.17 (1.68)	0.150
UL-Pog′Prn	−1.79 (0.98)	−1.56 (1.50)	−2.29 (2.09)	0.291
LL-Pog′Prn	−1.85 (1.58)	−1.75 (1.89)	−2.85 (2.06)	0.101
UL-Pog′Sn	−0.56 (0.94)	−0.38 (1.09)	−0.85 (1.47)	0.395
LL-Pog′Sn	−0.80 (1.43)	−0.76 (1.30)	−1.30 (1.24)	0.335
Soft-tissue angular measurements (°)				
Cm.Sn.Ls	−2.41 (7.84)	−4.24 (7.06)	−5.17 (7.46)	0.483
Li.B′.Pog′	3.91 (16.31)	1.78 (13.87)	3.27 (11.64)	0.865
Soft tissues thickness (mm)				
Prn-Nperp	1.50 (1.25)	1.29 (1.61)	1.48 (1.69)	0.877
ANS-Sn	−0.54 (2.17)	−1.26 (2.13)	−1.16 (6.56)	0.818
A-A′	−0.36 (1.72)	−0.33 (1.88)	0.39 (1.72)	0.303
UL-Ulp	−0.79 (2.11)	−1.01 (2.29)	−1.13 (2.77)	0.897
LL-Llp	−0.12 (1.31)	−0.10 (1.86)	−0.10 (1.69)	0.999
B-B′	0.34 (1.05)	0.04 (1.41)	0.52 (1.30)	0.416
Pog-Pog′	0.80 (1.27)	0.95 (1.44)	1.61 (1.46)	0.138
Convexity measures (°)				
n′.Sn.Pog′	−2.34 (2.82)	−1.02 (2.84)	−2.05 (1.98)	0.190
n′.Prn.Pog′	1.40 (3.40)	2.34 (3.80)	2.15 (2.73)	0.624
n.A.P	0.25 (1.98)	0.20 (3.10)	−0.96 (3.59)	0.328
ULPog′.PoO	4.84 (6.16)	6.01 (6.78)	7.66 (5.60)	0.354
Pog′Ls.NB	−2.55 (2.14)	−2.96 (3.01)	−3.78 (3.44)	0.389
Growth pattern (°)				
NS.Gn	−0.52 (1.73)	−0.54 (1.50)	−0.50 (1.42)	0.996
SN.GoGn	−1.04 (2.63)	−1.28 (2.33)	−2.28 (3.46)	0.318
PoOr.GoMe	−1.17 (2.81)	−1.70 (2.98)	−2.57 (4.03)	0.389

*ANOVA.

Table V. Comparison of rater groups

Variable	*Lay people (n = 77)*	*Orthodontists (n = 77)*	P *value*
Male	26 (38.24%)	42 (61.76%)	0.009*†
Female	51 (59.30%)	35 (40.70%)	
Age (y)	34.31 ± 12.96	44.08 ± 11.20	0.000*‡

Values are presented as n (%) or mean ± SD.
*Statistically significant at $P < 0.05$; †Chi-square test; ‡Mann-Whitney test.

the occlusion.[23] Because the occlusion is closely related to facial attractiveness, the groups had to have good and similar occlusal finishing. This was assured by all 3 groups having similar OGS indexes (Table I). In addition, the overjet significantly influences facial attractiveness. The groups were also matched regarding the posttreatment and long-term posttreatment overjets.

Different forms of treatment, when evaluated in the long term, require comparability at the posttreatment stage. Out of 25 cephalometric variables, only 2 showed statistically significant differences between the groups (SN.GoGn and PoOr.GoMe; Table II). It is understandable that the 4E group had a more vertical pattern than the others, because extraction treatments are better accepted in cases with vertical growth pattern.[24-26] However, this difference among the groups did not influence the attractiveness scores, because the group with intermediate growth pattern (2E) received the better rating (Table VI). This demonstrates high cephalometric comparability among the 3 groups at the posttreatment stage.

At the long-term posttreatment stage, almost all cephalometric variables were similar. The exception was nasal projection, which was significantly smaller in the 2E group (Table III). The nasal projection after treatment had a slight, but not significant, difference between groups. In addition, there was a slightly greater, but also not significant, change in nasal projection in the other groups during the long-term posttreatment

Table VI. Influence of treatment protocol, type of rater profile, and their interaction in profile attractiveness at the posttreatment stage (2-way ANOVA)

Group/evaluator	*n*	*Mean*	*SD*	P *value*
Group				
NE	40[†]	4.76[a,b]	1.32	0.044*
2E	54[†]	5.35[b]	1.02	
4E	42[†]	4.53[a]	1.11	
Evaluator				
Laypeople	77	4.73	1.15	0.000*
Orthodontist	77	5.11	0.95	
Evaluator × group (interaction)				
NE				
Laypeople	20	4.69[a,b,c]	1.18	0.018*
Orthodontist	20	4.84[a,b,c]	1.50	
2E				
Laypeople	27	5.06[a,b]	1.00	
Orthodontist	27	5.63[c]	1.10	
4E				
Laypeople	21	4.34[a]	1.10	
Orthodontist	21	4.71[b,c]	1.15	

[a,b,c]Values with different superscript letters are significantly different.
*Statistically significant at $P < 0.05$; [†]The number of observations are doubled because the scores of both groups of raters are included.

Table VII. Influence of treatment protocol, type of rater profile, and their interaction in profile attractiveness at the long-term posttreatment stage (2-way ANOVA)

Group/evaluator	*n*	*Mean*	*SD*	P *value*
Group				
NE	40[†]	4.32[a]	1.12	0.017*
2E	54[†]	5.08[b]	0.96	
4E	42[†]	4.33[a]	1.06	
Evaluator				
Laypeople	77	4.35	1.43	0.000*
Orthodontist	77	4.99	1.00	
Evaluator × group (interaction)				
NE				
Laypeople	20	4.08	0.93	0.139
Orthodontist	20	4.56	1.33	
2E				
Laypeople	27	4.72	0.80	
Orthodontist	27	5.45	1.16	
4E				
Laypeople	21	4.07	0.93	
Orthodontist	21	4.60	1.22	

[a,b]Values with different superscript letters are significantly different.
*Statistically significant at $P < 0.05$; [†]The number of observations are doubled because the scores of both groups of raters are included.

period (Table IV). The sum of these factors may explain this difference at the long-term posttreatment stage.

No statistical difference were found between the changes at the long-term posttreatment period, demonstrating high cephalometric comparability between the groups (Table IV).

At the posttreatment stage, the 2E group was rated significantly more attractive than the 4E group, and both showed no statistically significant difference compared with the NE group (Table VI). This similarity of the 2E and NE groups is understandable because these treatment protocols showed similar final profiles, as also seen in an earlier study.[7]

Assuming that individual features, such as nasal projection, could influence profile attractiveness, average cephalograms of the 3 groups at the posttreatment stage were traced and superimposed, with Frankfort planes parallel, centered on the tip of the nose (pronasal point). Figure 2 shows that the 4E group presents with a more retrognathic profile, although the difference is not significant. In the same figure, the NE group seems similar to the other groups, except for a reduced facial height, which is consistent with the results obtained.

In Class II malocclusion, 2-premolar extractions are usually indicated when there is no crowding or cephalometric discrepancy in the mandibular arch, although 4-premolar extractions are indicated especially in the presence of anteroinferior dental crowding and/or cephalometric discrepancy.[24] Patients with full cusp Class II and greater crowding have significantly smaller apical bases than patients with mild crowding or without crowding.[27] In this way, it can be speculated that the difference in attractiveness between groups 2E and 4E is due to a likely smaller basal bone in the 4E compared with the 2E group.

In the present sample, the posttreatment difference in profile attractiveness appears to be more related to the same characteristics that guided the treatment protocol choice in complete Class II malocclusions.

At the long-term posttreatment stage, the NE and 4E groups had statistically similar facial esthetics (Table VII). Stephens et al also found statistical similarity in profile esthetics of patients treated without extraction compared with 4-premolar extractions 15 years after treatment.[2] The results corroborate with other research as well that did not observe significant differences between groups with 4-premolar extractions and without extractions in the amount of long-term development.[28]

The 2E group performed significantly better in relation to the other 2 groups (Table VII). This result diverges a bit from the posttreatment evaluation (Table VI), when there was no statistically significant difference between the NE and 2E groups.

Protruded lips are related to attractive profiles, so current trends indicate that fuller lips confer a young appearance.[29] Therefore, the common belief that

Fig 2. Posttreatment-stage superimposition on the Frankfort plane centered at the tip of the nose.

premolar extractions can lead to a "dished in" profile and a premature aging effect has prevented some orthodontists to use that treatment modality. In addition, with the development of devices and different techniques for molar distalization, nonextraction treatments generally prevail.[30] The tendency to avoid extractions is greater in retrognathic patients. Therefore, it is possible that the NE group had more retrognathic mandibles than the 2E group before treatment. At the long-term follow-up, the NE group continued to show nonsignificantly greater convexity than the 2E group, which might explain the difference[31] (Table III).

Furthermore, the relative size of the nose, as well as other structures that may cause an apparent lip retrusion, influences profile attractiveness. In this comparison, attractiveness may have been influenced by nasal projection (Fig 3).

Another superimposition on the Frankfort plane, centered on the pronasal point, was created with the average cephalograms of the 3 groups at the long-term posttreatment stage (Fig 4). It is possible to observe a shorter facial height in the NE group, and that in the NE and 4E groups both jaws and lips are more retrognathic compared with the 2E group. This confirms how the change in nasal projection plays an important role when comparing the jaws.

A single variable, such as a more retrognathic mandible, may not be statistically different (Table III). However, when an increased nasal projection is added, an important difference in attractiveness may appear (Figs 2, 3, and 4). It seems that when separately assessed the variables may not show statistical differences between the groups, but attractiveness judgment assesses all of the variables together, and the sum of the slight differences in the variables may result in a statistically different attractiveness.

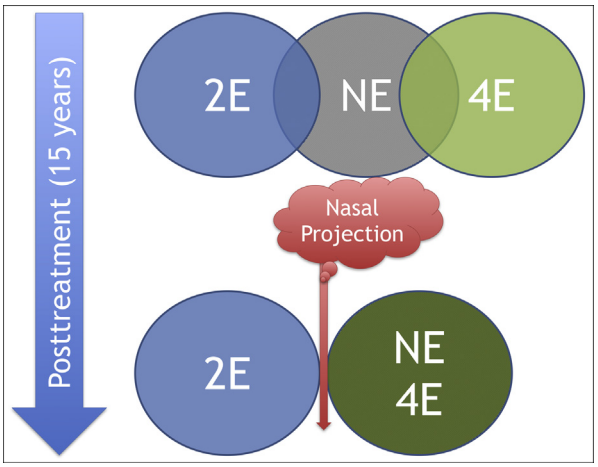

Fig 3. Supposed influence of nasal projection on profile attractiveness.

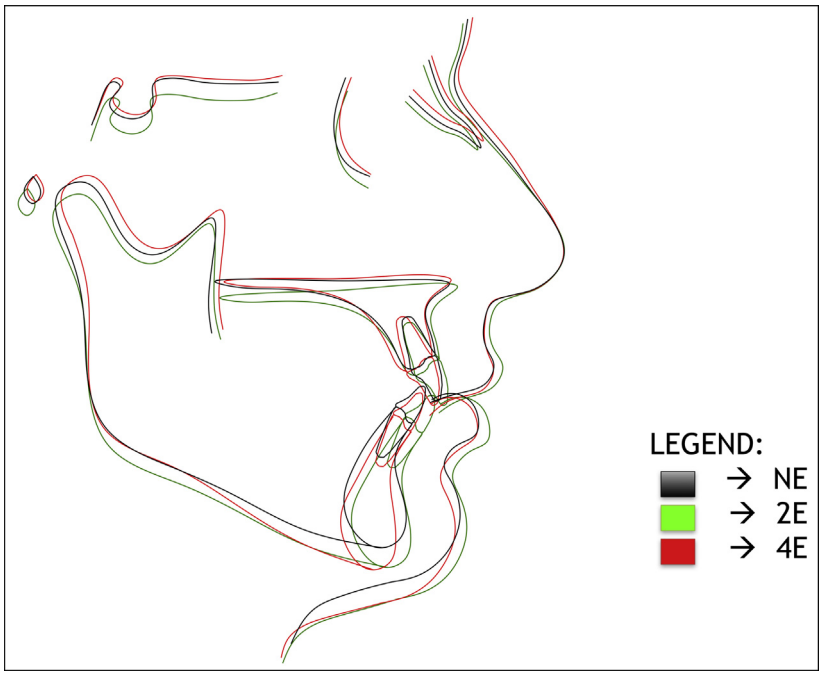

Fig 4. Long-term posttreatment–stage superimposition on the Frankfort plane centered at the tip of the nose.

CONCLUSIONS

With the present sample and methodology, the following can be concluded.

1. At the posttreatment stage, profile attractiveness was significantly greater in the 2-premolar extractions group compared with the 4-premolar extractions group, whereas the nonextraction group performed similarly to the other 2 groups.
2. At the long-term posttreatment stage, profile attractiveness was significantly greater in the 2-premolar extractions group.

ACKNOWLEDGMENT

This study received financial support from FAPESP (process number 2011/04603-6).

REFERENCES

1. Case SC. The question of extraction in orthodontia. Am J Orthod Dentofacial Orthop 1964;50:660-91.
2. Stephens CK, Boley JC, Behrents RG, Alexander RG, Buschang PH. Long-term profile changes in extraction and nonextraction patients. Am J Orthod Dentofacial Orthop 2005;128:450-7.
3. Strang RHW. Factors of influence in producing a stable result in the treatment of malocclusion. Am J Orthod Oral Surg 1946;32:313-32.
4. Janson G, Junqueira CH, Mendes LM, Garib DG. Influence of premolar extractions on long-term adult facial aesthetics and apparent age. Eur J Orthod 2016;38:272-80.
5. Janson G, Brambilla Ada C, Henriques JF, de Freitas MR, Neves LS. Class II treatment success rate in 2- and 4-premolar extraction protocols. Am J Orthod Dentofacial Orthop 2004;125:472-9.
6. Janson G, Barros SE, de Freitas MR, Henriques JF, Pinzan A. Class II treatment efficiency in maxillary premolar extraction and nonextraction protocols. Am J Orthod Dentofacial Orthop 2007;132:490-8.
7. Janson G, Fuziy A, de Freitas MR, Castanha Henriques JF, de Almeida RR. Soft-tissue treatment changes in Class II Division 1 malocclusion with and without extraction of maxillary premolars. Am J Orthod Dentofacial Orthop 2007;132:729.e721-8.
8. Janson G, Leon-Salazar V, Leon-Salazar R, Janson M, de Freitas MR. Long-term stability of Class II malocclusion treated with 2- and 4-premolar extraction protocols. Am J Orthod Dentofacial Orthop 2009;136:154.e1-10 [discussion 154-5].
9. Erdinc AE, Nanda RS, Dandajena TC. Profile changes of patients treated with and without premolar extractions. Am J Orthod Dentofacial Orthop 2007;132:324-31.
10. Casko JS, Vaden JL, Kokich VG, Damone J, James RD, Cangialosi TJ, et al., American Board of Orthodontics. Objective grading system for dental casts and panoramic radiographs. Am J Orthod Dentofacial Orthop 1998;114:589-99.
11. Navarro Rde L, Oltramari-Navarro PV, Fernandes TM, Oliveira GF, Conti AC, Almeida MR, et al. Comparison of manual, digital and lateral CBCT cephalometric analyses. J Appl Oral Sci 2013;21:167-76.
12. Janson G, Branco NC, Morais JF, Freitas MR. Smile attractiveness in patients with Class II division 1 subdivision malocclusions treated with different tooth extraction protocols. Eur J Orthod 2014;36:1-8.
13. Lim HJ, Ko KT, Hwang HS. Esthetic impact of premolar extraction and nonextraction treatments on Korean borderline patients. Am J Orthod Dentofacial Orthop 2008;133:524-31.
14. Torsello F, Graci M, Grande NM, Deli R. Relationships between facial features in the perception of profile attractiveness. Prog Orthod 2010;11:92-7.
15. Dahlberg G. Statistical methods for medical and biological students. New York: Interscience Publications; 1940.
16. Houston W. The analysis of errors in orthodontic measurements. Am J Orthod Dentofacial Orthop 1983;83:382-90.

17. Janson G, Nakamura A, Barros SE, Bombonatti R, Chiqueto K. Efficiency of Class I and Class II malocclusion treatment with four premolar extractions. J Appl Oral Sci 2014;22:522-7.
18. Abu Arqoub SH, Al-Khateeb SN. Perception of facial profile attractiveness of different antero-posterior and vertical proportions. Eur J Orthod 2011;33:103-11.
19. Coleman GG, Lindauer SJ, Tufekci E, Shroff B, Best AM. Influence of chin prominence on esthetic lip profile preferences. Am J Orthod Dentofacial Orthop 2007;132:36-42.
20. Macias Gago AB, Romero Maroto M, Crego A. The perception of facial aesthetics in a young Spanish population. Eur J Orthod 2012;34:335-9.
21. McNamara L, McNamara JA Jr, Ackerman MB, Baccetti T. Hard- and soft-tissue contributions to the esthetics of the posed smile in growing patients seeking orthodontic treatment. Am J Orthod Dentofacial Orthop 2008;133:491-9.
22. Bishara SE, Treder JE, Jakobsen JR. Facial and dental changes in adulthood. Am J Orthod Dentofacial Orthop 1994;106:175-86.
23. Gracco A, Cozzani M, D'Elia L, Manfrini M, Peverada C, Siciliani G. The smile buccal corridors: aesthetic value for dentists and laypersons. Prog Orthod 2006;7:56-65.
24. Bishara SE, Cummins DM, Jakobsen JR. The morphologic basis for the extraction decision in Class II, Division 1 malocclusions: a comparative study. Am J Orthod Dentofacial Orthop 1995;107:129-35.
25. Vaden JL. Alternative nonsurgical strategies to treat complex orthodontic problems. Semin Orthod 1996;2:90-113.
26. Tweed CH. The application of the principles of the edge-wise arch in the treatment of malocclusions: I. Angle Orthod 1941;11:5-11.
27. Janson G, Goizueta OE, Garib DG, Janson M. Relationship between maxillary and mandibular base lengths and dental crowding in patients with complete Class II malocclusions. Angle Orthod 2011;81:217-21.
28. Germec D, Taner TU. Effects of extraction and nonextraction therapy with air-rotor stripping on facial esthetics in postadolescent borderline patients. Am J Orthod Dentofacial Orthop 2008;133:539-49.
29. Berneburg M, Dietz K, Niederle C, Goz G. Changes in esthetic standards since 1940. Am J Orthod Dentofacial Orthop 2010;137:450.e451-9 [discussion 450-1].
30. Khan M, Fida M. Soft tissue profile response in extraction versus nonextraction orthodontic treatment. J Coll Physicians Surg Pak 2010;20:454-9.
31. Fortes HN, Guimaraes TC, Belo IM, da Matta EN. Photometric analysis of esthetically pleasant and unpleasant facial profile. Dental Press J Orthod 2014;19:66-75.