The effects of premolar-extraction: A long-term comparison of outcomes in "clear-cut" extraction and nonextraction Class II patients

Suwannee Luppanapornlarp, DDS, MSD; Lysle E. Johnston, Jr., DDS, PhD

Over the past decade or so, the day-to-day practice of orthodontics has undergone a series of profound, but basically evolutionary, changes. The specialty's current attitude toward premolar extraction, however, represents an abrupt, puzzling discontinuity. From a high of perhaps 60% to 80% 20 years ago[1], the extraction rate has declined to 30%, or perhaps even lower.[2] In a recent survey, O'Connor[2] found that over half of more than 800 responding orthodontists had reduced their extraction rate during the previous 5 years; only about 4% reported an increase. Because O'Connor's figures come from personal estimates, rather than from a review of records, individual clinicians may have over- or underestimated their true extraction rate by as much as 15-20%[3]; whatever the absolute rate, however, a downward trend seems clear. On the face of it, a reduction in the rate of premolar extraction would seem desirable. Unfortunately, the refereed literature provides little in the way of a rationale for this sudden change and suggests no reliable (that is, probably stable) alternative to extraction as a routine answer to the common problem of protrusion

Abstract

Discriminant analysis was used to assess the anatomical basis of the extraction/nonextraction decision in 238 former Saint Louis University Class II edgewise patients. The resulting discriminant scores (based on six measures of protrusion and crowding) were used to divide this parent sample into three prognostic subgroups: clear-cut extraction, clear-cut nonextraction, and a borderline stratum containing both extraction and nonextraction patients. The "clear-cut" patients—those at the tails of the distribution—were then contacted and asked to return for follow-up records (cephalograms, models, clinical examination); in the end, 62 (33 extraction and 29 nonextraction) were recalled. The average post-treatment interval was about 15 years.

Premolar extraction produced a significantly greater reduction in hard-and soft-tissue protrusion. During the post-treatment period, however, both groups underwent essentially the same change: decreased profile convexity and a pattern of dental change/relapse that was correlated with antero-posterior mandibular displacement. Because of their greater initial crowding and protrusion, the various effects summed to make the extraction patients significantly more protrusive at recall. Both treatments produced mesial mandibular displacement, extraction significantly more than nonextraction; however, at recall the two groups did not differ with respect to the signs and symptoms of dysfunction.

The present findings, therefore, fail to support the common, influential belief that premolar extraction frequently causes "dished in" profiles, "distalized" mandibles, and, ultimately, craniomandibular dysfunction.

Key Words

Premolar extraction • Nonextraction • Profile • Craniomandibular disorders • Growth • Relapse

Submitted: January 1993 **Revised and accepted for publication:** June 1993 Angle Orthod 63:257-272

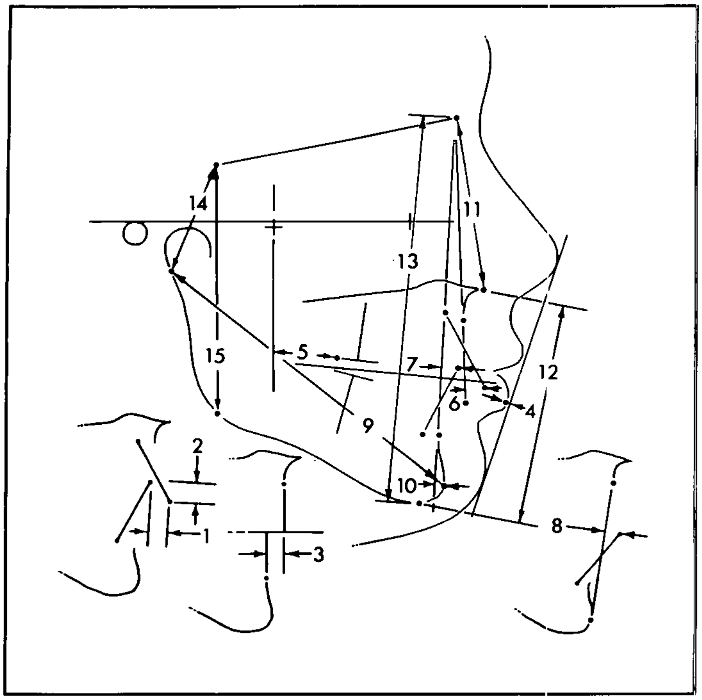

Figure 1A

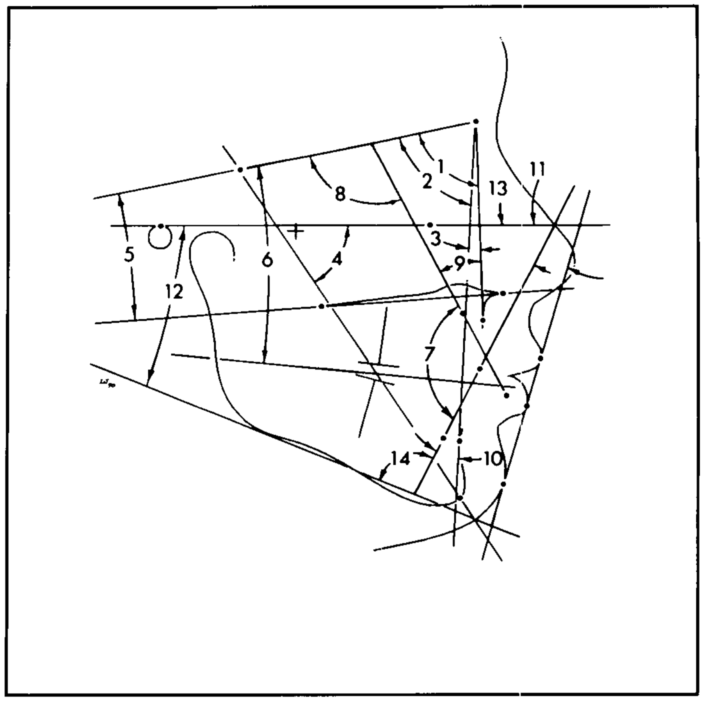

Figure 1B

Figure 1A-B Descriptive cephalometric analysis. A) linear measures; B) angular measures. The numbers correspond to the enumeration in the first column of Tables II and IV.

and crowding.

In response to this dilemma, we have used discriminant analysis to identify and recall matched samples of "borderline" extraction and nonextraction patients (i.e., those who could have been treated either way).[4-6] Although on average the two groups had presented with almost identical malocclusions and facial form, after treatment they showed a number of marked differences, most of which were still apparent nearly 15 years later. These differential treatment effects, however, had little if any discernible effect on stability, aesthetics (as judged by the patients, themselves), or the well-being of the musculature and temporomandibular joints. Unfortunately, despite this and other, more important evidence that craniomandibular disorders (CMD) are not caused by occlusion, by malocclusion, or by conventional orthodontic therapy, many contemporary clinicians now go to great lengths—expansion, bite-jumping, air-rotor "stripping"—to circumvent the need for premolar extraction. Given this spirit of therapeutic adventurism on the one hand and a lack of support and encouragement from the refereed literature on the other, we would argue that it is the "clear-cut" extraction patient who is now most in danger of clinical misadventure at the hands of those who would avoid extraction at any cost. We propose, therefore, to assess directly the extent to which these patients are penalized vis-à-vis those who, because of a lack of crowding and protrusion, seem ideally suited to nonextraction treatment, a strategy that has thus far escaped criticism. The results of this comparison should give some measure of the importance and urgency of the specialty's search for alternatives to premolar extraction.

Materials and methods

As part of an ongoing investigation of Class II treatment alternatives, discriminant analysis was used to define empirically the anatomical basis of the extraction/nonextraction decision in a large pool of former patients. The present report deals with long-term outcomes in the two extreme groups, the patients who apparently were susceptible to only one or the other of the two treatment strategies.

In preparation for this study, attempts were made to contact each of 2500 Saint Louis University Class II edgewise (.022") patients who completed treatment between 1969 and 1980. Because the goal of this study was to define homogeneous sub-groups and because the method was clearly going to require access to the records of several thousand ex-patients, the parent sample was composed entirely of Caucasians.* In the end, 238 of these former patients agreed to return for follow-up records should they meet the inclusion criteria that would be inferred from their pre-treatment records.

Discriminant analysis (subprogram DISCRIM-

*We are, however, currently repeating this line of research on an equivalent pool of African-American patients recruited from a number of university clinics and private practices.

INANT, SPSS-X, release 3.1, SPSS, Inc., Chicago, IL) was applied to 89 measurements obtained from each potential subject's pre-treatment cephalograms, study models, and examination form. These various independent variables were chosen to provide an abstraction of practically any morphological characteristic on which the initial treatment decisions could have been based. In addition to age and sex, there were 19 linear and 16 angular cephalometric measures (most of which are illustrated in Figure 1), as well as midline deviation, depth of curve of Spee, upper and lower inter-molar and inter-canine width, tooth-size (incisors, canines, and premolars), available space (measured in two segments per side), discrepancy, arch-depth, and lower anterior irregularity. A transparent digitizer (Scriptel RDT-1212, Scriptel Corp., Columbus, OH) and a commercial digitization program (Dentofacial Planner, version 4.22A, Dentofacial Software, Toronto, Canada) were used to generate the individual measurements from the cephalograms and photocopies of the occlusal surfaces of the pre-treatment models. To improve the reliability of the analysis, the more technically demanding landmarks were traced on a 0.003" frosted-acetate overlay prior to digitization.

The resulting discriminant function (fueled by upper and lower arch-length discrepancy, U1-SN, L1-NB, lower irregularity, and upper Z-angle) then was used to assign standardized discriminant scores to each of the 238 potential subjects (Figure 2A). Based on these scores, the former-patients were divided into three prognostic subgroups: extreme extraction (discriminant scores from 0.6 to 4.2), borderline extraction and nonextraction (scores near zero), and extreme nonextraction (scores from -1.1 to -4.8). The borderline patients were recalled first; these results have been reported elsewhere.[4-6] The present report deals with the 142 patients at the tails of the discriminant-score distribution.

Although all of the extreme patients had previously agreed to participate, 36 could not be re-contacted, 17 reversed their original decision, 13 had defective records, 7 were pregnant, 6 failed their appointment, and 1 had been retreated. In the end, 62 patients, 33 extraction (15 male, 18 female) and 29 nonextraction (11 male, 18 female), signed an informed consent document approved by the Saint Louis University Institutional Review Board and returned for follow-up records and a detailed clinical examination. At the time of recall, three extraction and two nonextraction subjects decided not to allow lateral cephalograms to be taken, thereby reducing the sample for the cephalometric portion of the study from 62 to 57. Each subject was paid $25 to defray the cost of participation. The average posttreatment interval was 15.3 years (range, 10.8-22.5 years). The participating extraction patients had discriminant scores that ranged from 0.7 to 4.2; the nonextraction, from -1.2 to -2.4 (Fig. 2B).

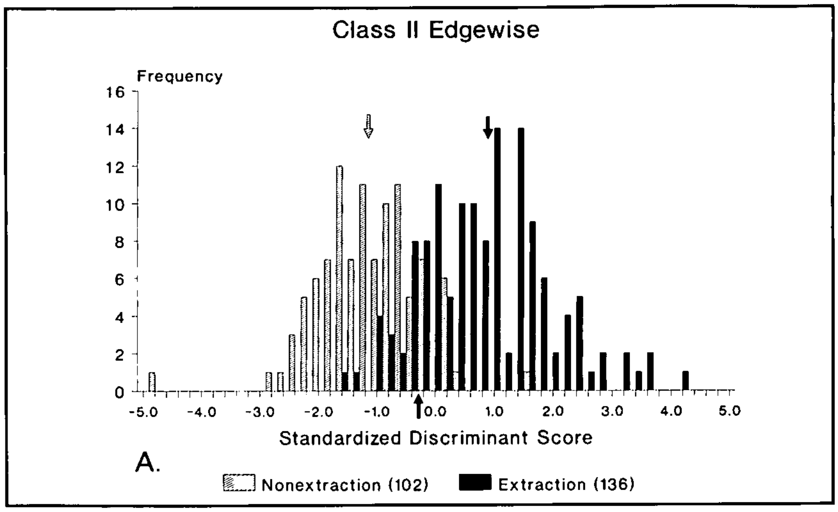

Figure 2A

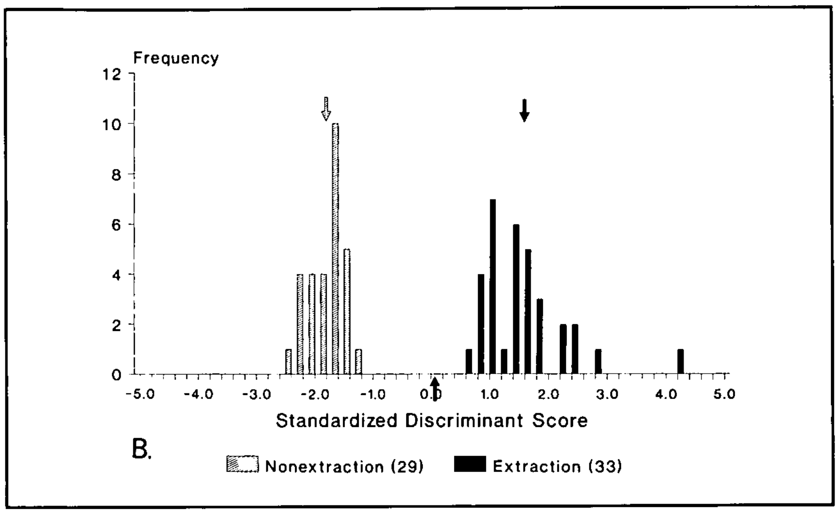

Figure 2B

Figure 2A-B Discriminant scores. A) Parent sample of 238 Class II subjects; B) "Extreme" subsamples. Note that the recalled subjects came from the tails of the distribution where there is little overlap (i.e., little empirical evidence of clinical uncertainty). Arrows denote mean discriminant scores (centroids).

Cephalometric analysis

Each subject's posttreatment and recall lateral cephalograms were subjected to the descriptive analysis that earlier had been used to generate the discriminant function. The main goal of this analysis was to characterize the skeletal, dental, and soft-tissue changes that had occurred, both during treatment and in the 10-20 years thereafter. In addition, detailed regional superimposition was employed to quantify the source of the anteroposterior correction/relapse of the molar relationship and overjet. Although this so-called "pitchfork analysis" (Figure 3) has been described elsewhere,[5,7-9] it is still relatively new and thus

Figure 3
"Pitchfork" analysis: the antero-posterior skeletal and dental components of the molar and overjet change (treatment corrections or post-treatment change). The growth and/or displacement of the maxilla and mandible are measured relative to cranial base (SE registration; MFOP orientation). The movements of the maxillary and mandibular molars and incisors are measured relative to their respective basal bone (regional superimposition). All measurements are executed parallel to MFOP and are given signs appropriate to the nature of their contribution to the change in molar relationship and incisor overjet. As a result, the algebraic sum of the various skeletal and dental changes equals the treatment change in the molar relationship and incisal overjet.

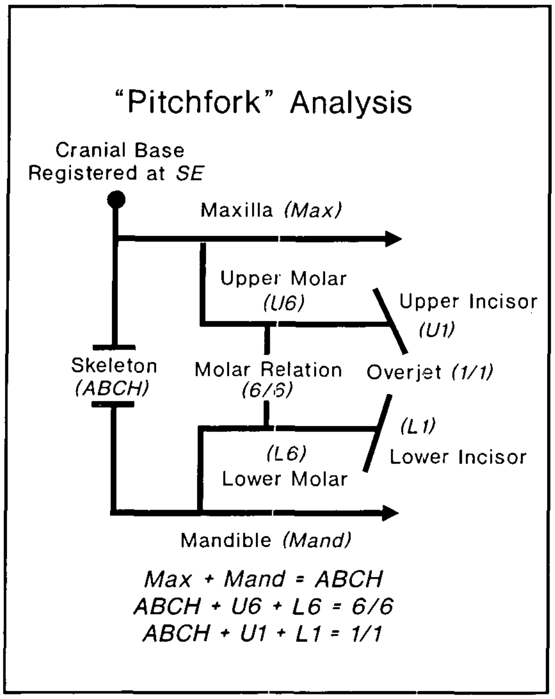

Figure 3

requires a somewhat more detailed description.

The dental and skeletal measurements of the pitchfork analysis were executed parallel to the mean functional occlusal plane (MFOP; initial/recall average), and each was given a sign appropriate to its impact on the molar or overjet correction: positive, if it improved the relationship (as would be the case, say, with forward growth of the mandible or distal movement of the maxillary molars and incisors); negative, if it made them worse (e.g., forward growth of the maxilla or mesial movement of the maxillary dentition). The various bony architectural features and landmarks employed in this part of the analysis were transferred by regional superimposition throughout each subject's three-film series from the film on which each could best be seen. Based on methods of cranial base, maxillary, and mandibular superimposition inferred from Björk's implant studies[10-12], we measured:

1) Mandibular symphyseal movement relative to maxilla (ABCH). This estimate of change in the relationship between maxillary and mandibular basal bone was measured as a shift in D-point as seen from the vantage point of a "best-fit" maxillary superimposition.

2) The displacement (as a result of growth, orthopedic changes, or functional shifts) of maxillary (Max) and mandibular (Mand) basal bone relative to cranial base. Max was measured from a maxillary superimposition as the apparent displacement of SE point, the point at which the averaged greater wings of the sphenoid cross planum sphenoidale; Mand was calculated by the algebraic subtraction of Max from ABCH.

3) The movement of the first molars (measured at the mesial contact point and at a point midway between the apices) and central incisors (at the incisal edge) relative to basal bone. Molar displacement was divided into two components: bodily (the movement of the apical point) and tipping (contact point movement minus apical displacement). Upper tooth movement was measured from a best-fit maxillary superimposition; lower tooth movement, from a superimposition employing an MFOP orientation and a D-point registration.

The resulting analysis is internally consistent: the various measurements of skeletal and dental change must add up to both the molar and overjet corrections. For example, the algebraic sum of ABCH (i.e., Max plus Mand) and molar movement relative to basal bone (U6 plus L6) is equal to the magnitude of the molar correction; the sum of the growth of the jaws and the movement of the incisors is equal to the overjet correction. Because of the technical difficulty of this analysis, the tracing, superimposition, and measurement were done by hand (with the aid of digital calipers); no digitization was employed.

Finally, to test the influential (albeit ungrammatical) argument that premolar extraction commonly "distalizes" the mandible, Björk's technique of mandibular regional superimposition (mandibular canal, unerupted molars, symphyseal architecture) was used to carry C-point, the approximate center of the pre-treatment condyle, through to the posttreatment and recall tracings.[8] Cranial base superimposition was then used to measure (parallel to MFOP) the apparent displacement of this point during the treatment and posttreatment periods. The use of regional superimposition to transfer an arbitrary pre-treatment condylar point to the other tracings in the series mimics the behavior of a metallic implant. Moreover, this technique does not require that we be able to visualize the precise condylar outline (which, in any event, would remodel over time and thus could not support an analysis of bodily displacement).

Model Analysis

The various measurements of arch length, arch width, discrepancy, and irregularity[13] that were used in the discriminant analysis were also obtained from digitized photocopies of the occlusal surfaces of the posttreatment and recall models.

Clinical Examination

Our clinical examination used the 62 items of the so-called "Craniomandibular Index" (CMI) of

Fricton and Schiffman.[14-15] These measures (Table I) provide a systematic evaluation of both the presumed signs and symptoms of joint dysfunction (mandibular movement, 16 items; temporomandibular joint noise, 4 items; temporomandibular joint capsular palpation, 6 items) and muscle pain (extra-oral palpation, 18 items; intra-oral palpation, 6 items; neck muscle palpation, 12 items). The clinical examinations were performed by the senior author after two calibration sessions: the first by Dr. John Rugh at the University of Texas Health Science Center at San Antonio and the second a month later by Dr. Rugh and Dr. James Fricton in St. Louis.

Statistical Analysis

With the aid of a table of random numbers, 10 three-film series (five extraction and five nonextraction) were selected, retraced, redigitized, and remeasured. Dahlberg's formula,

$$SD_E = \sqrt{\Sigma D^2/2N}$$

where D is the difference between double determinations, was then used to calculate the error standard deviation (SD_E) for each of the variables in the basic discriminant analysis. Given the replication of 10 series, N=30 for the basic descriptive measures, and N=20 for the measures of change. As may be seen in Tables II, III, and VI, the various linear descriptive measurements generally have error standard deviations that are less than 1 mm; the angular measures, less than 2 degrees.

Between-groups differences for initial, posttreatment, and recall data, means and for average treatment, posttreatment and net change were analyzed by means of completely randomized t-tests. The Mann-Whitney U statistic, corrected for ties, was used to test for significant ($P<.05$) between-groups differences in the number of positive scores on the 26 joint-dysfunction (D) and 36 muscle-palpation (P) items, as well as for both categories, combined.[16] Finally, Pearson product-moment coefficients (r) were calculated to assess the strength of the relationship between mandibular growth and/or displacement relative to cranial base (Mand) and the various estimates of molar and incisor movement during and after treatment.

Results

The two samples recalled here were at the extremes of a distribution of discriminant scores that were based on six measures of crowding and protrusion. As a result, there were a number of significant pretreatment differences (Figure 4A), many of which still were seen at recall (Figure 4B). As may be inferred from the start/finish/recall

Table I
Craniomandibular Index: Menu of items surveyed at recall (after Fricton & Schiffman, 1987[15])

Scale	Items	N
Dysfunction (D)—26 points		
Mandibular Movement (MM)	Opening	
	< 40 mm	1
	Passive stretch < 42 mm	1
	Restricted	1
	Painful	1
	"Jerky"	1
	S-deviation > 2 mm	1
	Lateral deviation > 2 mm	1
	Locks open	1
	Locks closed	1
	Protrusion	
	Pain	1
	Limitation < 7 mm	1
	Laterotrusion (R/L)	
	Pain	2
	Limitation < 7 mm	2
	Rigidity	1
TMJ Noise (TN)	Click (R/L)	2
	Crepitus/popping (R/L)	2
TMJ Capsule Palpation (TP)	Lateral (R/L)	2
	Superior (R/L)	2
	Posterior (R/L)	2
Palpation (P)—36 points		
Extra-oral Palpation (EP)	Temporalis (R/L)	
	Anterior	2
	Deep	2
	Middle	2
	Masseter (R/L)	
	Anterior	2
	Deep	2
	Inferior	2
	Posterior Digastric (R/L)	2
	Medial Pterygoid (R/L)	2
	Vertex (R/L)	2
Intra-oral Palpation (IP)	Pterygoid (R/L)	
	Lateral	2
	Medial	2
	Temporalis Insertion (R/L)	2
Neck Muscle Palpation (NP)	Sternocleidomastoid (R/L)	
	Superior	2
	Middle	2
	Inferior	2
	Trapezius (R/L)	
	Upper	2
	Insertion	2
	Splenius Capitus	2

Table II
Start, Finish, and Recall: Descriptive and inferential statistics for representative cephalometric and demographic variables

Measure	SD_E (N=30)	Start			Finish			Recall		
		Extr.	Non.	t	Extr.	Non.	t	Extr.	Non.	t
Linear (mm)										
1. Overjet	0.7	7.4	7.5	ns	2.9	2.8	ns	4.2	4.3	ns
2. Overbite	0.4	3.5	4.2	ns	1.6	1.5	ns	3.5	3.5	ns
3. Wits A/B	0.8	2.8	1.3	*	0.9	0.2	ns	2.6	1.5	ns
4. L Lip-E-Plane	0.5	1.7	-3.0	**	-1.2	-3.5	**	-3.9	-6.3	**
5. U6-PTV	0.6	18.6	19.9	ns	21.0	19.1	*	24.9	22.9	*
6. U1-NA	0.6	4.8	4.1	ns	2.0	2.7	ns	2.1	3.2	ns
7. L1-NB	0.4	6.2	2.8	**	5.5	4.5	ns	4.7	3.3	*
8. L1-APog	0.4	1.5	-0.9	**	1.5	1.3	ns	0.2	-0.0	ns
9. Ar-Gn	0.5	103.3	106.4	ns	107.7	109.2	ns	113.7	114.3	ns
10. Pog-NB	0.3	2.0	2.6	ns	3.0	3.2	ns	3.9	3.7	ns
11. N-ANS	0.4	51.5	53.4	*	53.9	54.7	ns	56.1	56.8	ns
12. ANS-Me	0.5	68.7	62.7	**	71.4	66.1	**	74.2	68.7	**
13. N-Me	0.4	118.0	114.0	*	123.4	119.1	*	128.6	123.9	*
14. S-Ar	0.4	34.6	34.5	ns	36.1	35.6	ns	38.2	37.3	ns
15. S-Go	0.2	72.4	73.5	ns	76.4	77.2	ns	82.9	82.7	ns
Molar rel	0.3	-1.0	-0.9	ns	2.2	2.0	ns	1.6	1.4	ns
PNS-A	0.8	48.8	49.5	ns	48.7	49.5	ns	51.5	51.1	ns
Angular (degrees)										
1. SNA	0.3	80.8	80.5	ns	79.0	79.1	ns	79.9	79.2	ns
2. SNB	0.2	74.9	76.3	*	74.8	76.1	ns	75.6	76.4	ns
3. ANB	0.2	6.0	4.0	**	4.2	3.1	*	4.3	2.8	*
4. Y-axis	0.3	58.8	54.7	**	59.1	55.9	**	59.2	56.2	*
5. Pal-SN	0.5	7.2	8.9	*	8.2	9.6	ns	8.0	9.4	ns
6. FOP-SN	1.5	20.0	18.2	ns	20.6	18.7	ns	17.6	15.8	ns
7. 1/1	1.9	123.1	130.0	**	126.9	127.0	ns	133.1	135.1	ns
8. U1-SN	1.5	104.3	104.1	ns	99.7	101.7	ns	97.6	99.2	ns
9. U1-NA	1.4	23.5	23.8	ns	20.7	22.6	ns	7.8	20.0	ns
10. L1-NB	1.4	27.5	21.8	**	28.2	27.4	ns	24.9	22.1	ns
11. Z-Angle (U) .	1.9	70.4	77.1	**	75.2	79.2	**	83.2	85.5	*
12. FMA	0.4	24.7	18.9	**	24.5	19.6	**	22.8	18.3	*
13. FMIA	1.4	58.1	66.3	**	57.5	60.4	ns	61.0	65.7	*
14. IMPA	1.6	97.2	94.9	ns	98.1	100.0	ns	96.2	96.0	ns
Demographic (years)										
Age	. .	12.9	13.1	ns	14.8	14.8	ns	30.2	30.1	ns

*P<.05; **P<.01

Table III
Dental Arch Measures: Means and t-scores for between-treatment differences

Measure	SD_E (N=30)	Start Extr.	Start Nonex.	Start t	Finish Extr.	Finish Nonex.	Finish t	Recall Extr.	Recall Nonex	Recall t
Intercanine width										
Maxillary	0.3	32.4	31.0	**	33.6	31.9	**	33.3	31.4	**
Mandibular	1.3	25.1	24.3	ns	26.5	25.2	**	25.1	23.8	**
Intermolar width										
Maxillary	0.5	47.1	47.2	ns	46.4	48.9	**	46.5	49.2	**
Mandibular	0.3	42.3	43.2	ns	39.8	43.0	**	40.3	43.5	**
Arch Length										
Maxillary	0.6	71.2	70.9	ns	62.9	70.0	**	60.3	67.8	**
Mandibular	0.4	61.3	61.6	ns	52.7	61.0	**	50.2	57.7	**
Discrepancy										
Upper	1.2	-5.8	-0.9	**	-1.4	-1.9	ns	-3.6	-4.1	ns
Lower	1.0	-7.3	-0.8	**	-2.8	-2.9	ns	-6.5	-5.5	ns
Irreg. Index	0.7	7.2	2.9	**	0.6	0.6	ns	3.2	3.7	ns
Curve of Spee	0.1	1.9	1.9	ns	0.4	0.5	ns	1.0	0.9	ns

*P<.05; **P<.01

descriptive statistics summarized in Tables II and III, the extraction patients presented with more protrusive dentitions and convex profiles, greater discrepancy and irregularity, and increased lower anterior face height. In general, the skeletal differences tended to persist, whereas many of the dental differences were eliminated by treatment.

In terms of change (Tables IV-VI), the premolar-extraction treatments produced significantly more incisor and lip retraction (about 2 mm), irregularity and discrepancy reduction (on average, 5 mm), and molar anchorage loss (2 mm). Because both treatments showed essentially the same pattern of posttreatment change (which would include "relapse"), the final net between-groups differences were similar to those seen at the end of treatment. In passing, it should be noted (Table VI) that in both groups the condyles came forward during treatment; however, this mesial displacement was significantly more pronounced in the premolar-extraction patients. Finally, as summarized in Table VII, many of the incisor and molar movements both during and after treatment bore a significant linear correlation to Mand, a measure of the anteroposterior movement of mandibular basal bone relative to cranial base.

Despite a variety of statistically significant skeleto-dental differences at all three time-points, there were no significant differences in the prevalence of the putative signs and symptoms of dysfunction catalogued in the CMI or its two component parts, the Palpation and Dysfunction indices (Table VIII). The extraction patients aver-

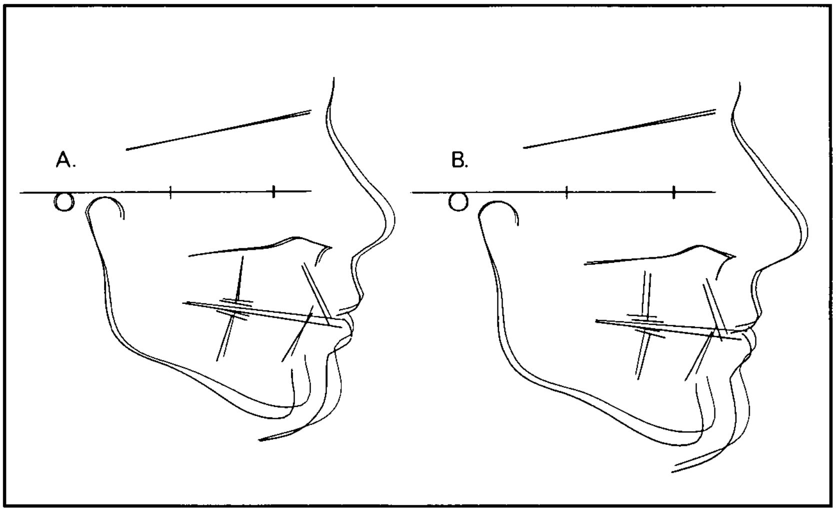

Figure 4A-B

Figure 4A-B
Recalled "extreme" extraction and nonextraction Class II patients (averaged tracings superimposed on FH and registered at PTV: A) Initial, and B) Recall. Thick lines, extraction (N=30); thin lines, nonextraction (N=27). Note that, on average, the two groups were different prior to treatment and were still different 10 to 20 years later.

Table IV
Treatment, posttreatment, and net change: Descriptive and inferential statistics for representative cephalometric and demographic variables

Measure	Treatment			Posttreatment			Net		
	Extr.	Non.	t	Extr.	Non.	t	Extr.	Non.	t
Linear (mm)									
1. Overjet	-4.5	-4.7	ns	1.3	1.5	ns	-3.2	-3.1	ns
2. Overbite	-1.9	-2.7	ns	1.9	2.0	ns	-0.1	-0.7	ns
3. Wits A/B	-1.9	-1.1	ns	1.7	1.3	ns	-0.2	0.3	ns
4. L Lip-E-Plane	-3.0	-0.6	**	-2.7	-2.8	ns	-5.7	-3.3	**
5. U6-PTV	2.4	-0.8	**	3.9	3.8	ns	6.3	3.0	**
6. U1-NA	-2.8	-1.4	ns	0.1	0.5	ns	-2.7	-0.9	**
7. L1-NB	-0.7	1.7	**	-0.8	-1.2	ns	-1.5	0.5	**
8. L1-APog	0.0	2.1	**	-1.3	-1.3	ns	-1.3	0.9	**
9. Ar-Gn	4.5	2.7	ns	5.9	5.2	ns	10.4	7.9	ns
10. Pog-NB	1.0	0.6	ns	0.9	0.5	ns	2.0	1.	**
11. N-ANS	2.4	1.6	ns	2.2	2.1	ns	4.6	3.7	ns
12. ANS-Me	2.7	3.4	ns	2.8	2.6	ns	5.5	6.0	ns
13. N-Me	5.4	5.0	ns	5.2	4.8	ns	10.5	9.8	ns
14. S-Ar	1.5	1.1	ns	2.1	1.7	ns	3.6	2.8	ns
15. S-Go	4.0	3.6	ns	6.5	5.6	ns	10.5	9.2	ns
Molar rel	3.2	2.9	ns	-0.6	-0.6	ns	2.6	2.3	ns
PNS-A	-0.1	0.0	ns	2.8	1.6	*	2.8	1.7	*
Angular (degrees)									
1. SNA	-1.9	-1.4	ns	0.9	0.2	ns	-0.9	-1.2	ns
2. SNB	-0.1	-0.3	ns	0.9	0.4	ns	0.8	0.2	ns
3. ANB	-1.8	-1.0	*	0.1	-0.3	ns	-1.7	-1.3	ns
4. Y-axis	0.3	1.2	*	0.0	0.3	ns	0.3	1.5	ns
5. Pal-SN	1.0	0.7	ns	-0.2	-0.2	ns	0.8	0.5	ns
6. FOP-SN	0.6	0.4	ns	-3.0	-2.9	ns	-2.3	-2.4	ns
7. 1/1	3.8	-3.0	*	6.1	8.1	ns	10.0	5.1	ns
8. U1-SN	-4.6	-2.4	ns	-2.0	-2.4	ns	-6.6	-4.9	ns
9. U1-NA	-2.7	-1.2	ns	-3.0	-2.6	ns	-5.7	-3.8	ns
10. L1-NB	0.7	5.6	**	-3.3	-5.2	ns	-2.6	0.3	ns
11. Z-Angle (U)	4.8	2.1	**	3.1	4.1	ns	7.9	6.2	ns
12. FMA	-0.2	0.7	ns	-1.7	-1.3	ns	-1.9	-0.6	ns
13. FMIA	-0.7	-5.8	**	3.5	5.3	ns	2.9	-0.6	ns
14. IMPA	0.9	5.1	*	-1.9	-4.0	ns	-1.0	1.1	ns
Demographic (years)									
Age	1.9	1.7	ns	15.4	15.3	ns	17.3	17.0	ns

*P<.05; **P<.01

Table V
Change in dental-arch dimensions: Means and t-scores for between-treatment differences

Measure	Treatment			Posttreatment			Net		
	Extr.	Nonex.	t	Extr.	Nonex.	t	Extr.	Nonex.	t
Intercanine width									
Maxillary	1.2	0.9	ns	-0.3	-0.5	ns	0.9	0.4	ns
Mandibular	1.4	0.8	ns	-1.5	-1.4	ns	-0.0	-0.6	ns
Intermolar width									
Maxillary	-0.7	1.7	**	0.1	0.3	ns	-0.6	2.0	**
Mandibular	-2.5	-0.2	**	0.5	0.5	ns	-1.9	0.3	**
Arch Length									
Maxillary	-8.3	-0.9	**	-2.6	-2.2	ns	-10.9	-3.1	**
Mandibular	-8.6	-0.6	**	-2.5	-3.3	ns	-11.1	-4.0	**
Discrepancy									
Upper	4.5	-1.0	**	2.3	-2.2	ns	2.2	-3.2	**
Lower	4.6	-2.1	**	-3.7	-2.6	ns	0.9	-4.7	**
Irreg. Index	-6.6	-2.3	**	2.6	3.1	ns	-4.0	0.8	**
Curve of Spee	-1.5	-1.5	ns	0.6	0.4	ns	-0.9	-1.0	ns

*P<.05; **P<.01

aged 4.67 signs/symptoms of dysfunction; the nonextraction, 4.66.

Discussion

In contrast to earlier papers in which discriminant analysis was used to minimize between-treatment differences (i.e., susceptibility bias), it was used in the present study to maximize these differences. By identifying and eliminating patients who could have been treated either way, we were left with patients who seem to have been ideally suited to the treatments they received. If the critics of premolar extraction are correct, the present study—because of its design, its sample size, and its relatively small technical error—should produce results that are consistent with their various claims and conjectures. Moreover, such a confirmation would, at long last, provide evidence that the search for alternatives to premolar extraction is at least of potential benefit to the patient. The data, however, are not encouraging.

As might be expected, premolar-extraction had a greater impact (by 2 to 3 mm) on the profile. Indeed, the nonextraction treatments generally produced less of a change in denture position and profile convexity than occurred in the posttreatment period, presumably at the hands of the normal pattern of facial growth. It should not be inferred, however, that the extraction profiles were too "flat" on recall. Instead, it was the nonextraction patients who tended to have concave faces, whereas the extraction patients more often had what nonextraction advocates might call "nice,

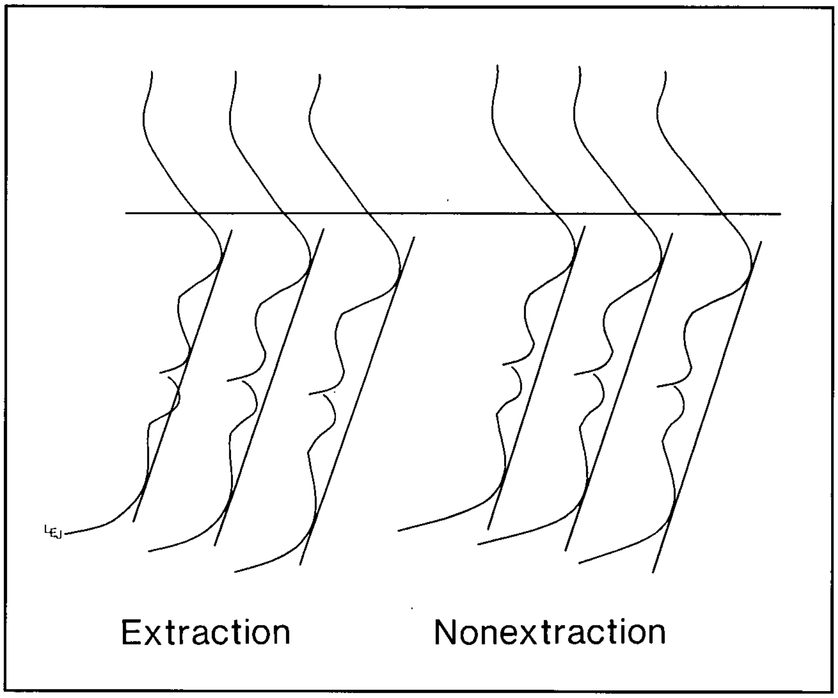

Figure 5

Figure 5
Mean start, finish, and recall facial polygons for the extreme extraction and nonextraction samples. At recall, it was the nonextraction subjects who tended to have the "flatter" profiles.

Table VI
Dental and skeletal components of the molar and overjetcorrection: Means and t-scores for between-treatment differences

Measure	SD_E (N=20)	Treatment Extr.	Treatment Non.	Treatment t	Posttreatment Extr.	Posttreatment Non.	Posttreatment t	Net Ext.	Net Non.	Net t
Skeleton—Jaw displacement relative to cranial base										
MAX	0.6	-1.1	-0.9	ns	-1.9	-1.2	ns	-3.0	-2.1	ns
MAND	1.1	3.8	2.8	ns	3.2	2.4	ns	7.0	5.2	ns
Net Skel. (ABCH)	0.8	2.7	2.0	ns	1.3	1.2	ns	4.1	3.1	ns
C-Point	0.2	1.8	1.0	*	0.5	0.6	ns	2.3	1.6	ns
Dentition—Molar movement/relapse relative to basal bone										
U6										
Bodily	0.6	-2.0	-0.3	**	-0.5	-0.6	ns	-2.5	-0.9	**
Tipping	0.5	-0.1	1.0	**	-2.0	-2.2	ns	-2.1	-1.2	ns
L6										
Bodily	0.8	4.2	2.3	**	0.3	0.2	ns	4.5	2.5	**
Tipping	0.9	-1.5	-1.8	ns	0.5	1.0	ns	-1.0	-0.8	ns
Net 6	..	0.7	1.2	ns	-1.8	-1.7	ns	-1.1	-0.5	ns
Dentition—Incisor movement/relapse relative to basal bone										
U1	0.7	3.6	1.8	*	-1.4	-1.0	ns	2.2	0.8	ns
L1	0.8	-2.0	0.8	**	-1.6	-1.7	ns	-3.6	-0.9	**
Net 1	..	1.6	2.6	ns	-3.0	-2.8	ns	-1.5	-0.2	ns
Total Correction—ABCH plus Net 6; ABCH plus Net 1										
6/6	0.4	3.4	3.1	ns	-0.5	-0.5	ns	2.9	2.7	ns
Overjet	1.0	4.3	4.6	ns	-1.7	-1.6	ns	2.6	3.0	ns

*P<.05 **P<.01

full, pleasing profiles" (Fig. 5). Given the initial crowding and protrusion of the extraction patients (and the relative spacing and retrusion of the nonextraction patients), this result is unremarkable; it is mentioned at the outset merely to counter the popular notion that "dished-in profiles" are a unique, obligatory by-product of first-premolar extraction therapy.

The striking feature of the present results is that they match, almost line-for-line, those obtained in an earlier extraction/nonextraction comparison[5] in borderline prognostic sub-groups. The two strategies produced about the same treatment differential in each study, and each had about the same effect within both the extreme and borderline strata. Further, the posttreatment change was independent, not only of the type of patient (i.e., discriminant score), but also the type of treatment. All patients tended to undergo about the same change. Although this result may seem surprising, it really is not: if future change were a simple function of individual facial form, cephalometric prediction would have yielded long ago to the various univariate and multivariate statistical techniques that have been brought to bear. Instead, much of the final orthodontic outcome can be seen as a dentoalveolar compensation for the usual pattern of posttreatment jaw growth and/or displacement (Table VII) combined with the distinctive, relatively stereotyped impact of the treatment chosen by the clinician.

As summarized in Table VII for both borderline and extreme subjects, many of the dental components of the molar and overjet correction and posttreatment change bear a significant correlation to the antero-posterior growth and/or displacement of the mandible (Mand). In a previous paper[5], tooth movement was correlated significantly with ABCH, a skeletal variable derived from the same maxillary regional superimposition that was used to estimate the movement of the upper molars and incisors. Because the resulting shared superimposition errors would serve to inflate the correlation coefficients, we have chosen instead to use Mand, a more technically independent measure of facial skeletal growth. In any event, significant linear correlation does not imply causation: the growth of the jaws might influence the movement of the teeth or, with equal plausibility, the movement of the teeth might influence the displacement and even the growth of the jaws. The literature offers support for both points of view.

In young patients, the mandible generally outgrows the maxilla. If premolars are extracted, the upper buccal segments almost always move mesially. In non-growing adults, however, there is no maxillo-mandibular differential, and the upper buccal segments can be held in place.[4,17] This interesting difference is perhaps a reflection of the presence or absence of dentoalveolar compensation for the excess growth of the mandible. Conversely, the present report and others[8,18-21] argue that orthodontic treatment (especially extraction treatment) commonly would increase the distance between centric occlusion and centric relation, presumably due to a forward displacement of the mandible, rather than distal drift of the glenoid fossa. We would suggest, therefore, that it might be better to assume provisionally that the jaws and the teeth influence one another during treatment (i.e., when dentoalveolar compensation is constrained by brackets and archwires), but that jaw-growth (and/or displacement) is the primary event once the appliances have been removed. Given this interpretation, some common forms of posttreatment relapse (Table VII) can be seen as dentoalveolar compensations for the pattern of posttreatment jaw growth.[5,22-23] Indeed, a good portion of the final result is a by-product of growth and the relatively stereotyped impact of the clinician's choice of treatment. On the face of it, this interpretation implies that some sort of treatment prediction should be possible. Unfortunately, things are not quite so simple. As noted recently by Sinclair:[24]

> "It becomes clear that our ability to predict future growth in a clinically meaningful manner must remain low. Growth prediction is not a simple matter. Adding a "mean increment" of growth derived from a similar group of patients to an individual's measurements is unlikely to produce an accurate individual growth prediction."

By definition, an "accurate individual growth prediction" can only come from a method that can assign a different increment to two patients of the same age and sex. Tracing error, sensitivity to initial conditions, and variations in clinical proficiency and patient cooperation all generate random "noise" that would tend to obscure the often subtle individual effects we would like to predict. An inability to account for individual variation—to know who will grow more than average and who will grow less—does not, however, mean that the future is totally unknown. There is change, and much of it—the mean increment—tends to affect everyone. It is of sufficient magnitude to be a factor in various kinds of long-term planning.

For example, when parents buy clothes for their children, they commonly leave a little room for growth, even though they may not be able to predict the exact increment. They know roughly how the average child grows (some rather than none; larger rather than smaller; inches rather than microns or yards) and they know at what ages these changes are likely to occur. These same parents would perhaps be surprised to know that it is not at all unusual for their children's orthodontic treatment to be planned without nearly as much concern for the probable effects of growth. This attitude is the end result of spending the entire cephalometric era, the last 50-60 years or so, locked into extreme, theoretically barren positions. As a result, some now argue that the face does not change (Brodie's "pattern concept"), whereas others say that it does, but in a way that is totally unpredictable. Either way, treatment planning reduces to a largely static process in which growth plays no part. Clearly, this is a potentially dangerous over-simplification.

First of all, Brodie's so-called "pattern concept" was wrong: the pattern does change and in so doing tends to affect many aspects of treatment, including even the details of tooth movement. For example, we have long known (say, from Björk[25] or Lande[26]) that most growing subjects—even those with Class II malocclusions–will get bigger and in the process will undergo a reduction in the mandibular and occlusal plane angles, a flattening of the facial profile (much of it due to the continued growth of the nose and a forward rotation of the chin), and a reduction in arch space.

Table VII
Coefficients of linear correlation (r) between mandibular displacement relative to cranial base (Mand) and the components of the change in molar relation and overjet

Component	Treatment Change		Posttreatment Change	
	Borderline	Extreme	Borderline	Extreme
U6				
Bodily	0.04	-0.10	-0.10	-0.38**
Tipping	-0.34**	-0.25	-0.57**	-0.37**
L6				
Bodily	-0.17	0.32*	-0.27*	-0.12
Tipping	-0.31*	-0.61**	-0.38**	-0.37**
Molars (6/6)	0.25*	0.58**	0.33**	0.52**
U1	-0.24	0.14	-0.66**	-0.47**
L1	-0.28*	-0.61**	-0.74**	-0.47*
Overjet (1/1)	0.06	0.25	-0.09	0.33*

*P<.05 **P<.01

The failure of growth prediction, therefore, stems not from being unable to predict change, but rather from a proven inability to account for individual departures from the mean increment of change. Given the magnitude of this increment, even if true prediction were to prove impossible (as the junior author has long argued[27-28]), it would still be of considerable utility to know the most likely pattern of growth and the probable impact of a given type of treatment. In combination with a characterization of the individual pre-treatment face, the resulting "projection" would do much to prevent the various aesthetic "horror stories" (overly flattened profiles, for example) that are so often exploited by participants in the current extraction/nonextraction debate.

Given our results, we would argue that three sets of "expected" (i.e., average) skeletal and dental effects would suffice for Class II treatment planning: one for premolar-extraction treatment, one for nonextraction treatment, and one for the effect of posttreatment "growth." Rough estimates of some of these average increments (taken to the

Table VIII
Joint and muscle evaluation: Positive responses for dysfunction and muscle palpation

	Extraction (N=33)				Nonextraction (N=29)		
Sex	Dysfunct.	Palpation	Total	Sex	Dysfunct.	Palpation	Total
M	2	0	2	M	3	1	4
F	2	6	8	F	3	3	6
F	2	0	2	F	7	6	13
M	5	4	9	M	3	0	3
F	4	0	4	F	3	2	5
M	0	4	4	M	1	0	1
M	3	0	3	F	6	3	9
M	0	0	0	M	0	0	0
F	6	2	8	M	3	0	3
M	6	0	6	F	1	0	1
M	1	1	2	F	2	1	3
M	2	2	4	M	2	5	7
F	2	1	3	M	0	0	0
M	6	1	7	F	1	5	6
F	1	0	1	M	6	2	8
F	5	3	8	F	4	0	4
F	6	2	8	F	7	3	10
F	5	0	5	M	0	0	0
M	6	0	6	M	0	0	0
F	2	0	2	F	0	0	0
F	2	0	2	F	8	2	10
M	7	0	7	F	0	5	5
F	0	0	0	M	0	0	0
F	3	4	7	F	7	2	9
M	2	0	2	F	0	2	2
M	4	0	4	F	7	1	8
F	1	2	3	F	3	1	4
F	9	2	11	F	5	2	7
F	4	2	6	F	5	2	7
M	4	7	11	..	..	..	..
M	1	1	2	..	..	..	..
F	0	0	0	..	..	..	..
F	6	1	7	..	..	..	..
Total	109	45	154		87	48	135
Mean	3.30	1.36	4.67		3.00	1.66	4.66

nearest 0.25 mm or so) are summarized in Table IX. These data, in conjunction with earlier investigations of Class II treatment[7,29-35] and cephalometric prediction[27-28], argue that the average millimetric impact of both the normal pattern of facial growth and the various common treatment options constitute a series of relatively constant building blocks from which a rational, individualized treatment plan can be constructed. For example, in terms of effect on denture position, mean increments of change (plus = protrusion; minus = retrusion) can be estimated for a variety of common treatment options: extraoral traction, -1 mm; fully preangulated brackets ("straight-wire"), +1 mm; upper second-molar extraction, -1.5 mm; first-premolar extraction, -3 mm; second-premolar extraction, -2 mm; posttreatment growth, -3 mm; sagittal appliance treatment, +2 to +3 mm; and so forth. In addition to its effect on antero-posterior position, it would also seem prudent to consider a treatment's expected long-term impact on the alignment of the dental arches (i.e., its ability to deal with problems that commonly constitute the chief complaint).

The present investigation and many others before it (e.g., Little, Wallen, and Riedel[36], Glen, Sinclair and Alexander[37], and Paquette, Beattie, and Johnston[5]) document a long-term posttreatment arch-length reduction of 2-3 mm that is apparently independent of treatment strategy or type of Class II. Thus, although the nonextraction patients started out with little discrepancy and irregularity, it was the extraction patients who were better able to accommodate the posttreatment change that tends to affect all groups more or less equally. As a result, the extraction patients improved, whereas the nonextraction patients worsened; the between-treatment differences were statistically significant and large enough to allow the originally crowded and protrusive extraction patients to achieve parity with their much less severely affected nonextraction cohorts. Moreover, the extraction patients were treated without any mean collapse of the upper intercanine width, and, more to the point of this discussion, the nonextraction patients were treated without resorting to routine expansion (Table III). Given this conservative approach to treatment, our data may underestimate the arch-length reduction that would be seen following more aggressive nonextraction therapy.

Our data also offer little support for the component parts of the venerable assertion that extraction produces distal condylar displacement, and that this displacement, in turn, produces TMD. First, as judged by regional superimposition, both treatments produced or permitted forward mandibular displacement, premolar-extraction significantly more so than nonextraction. Given that space-closure commonly pulls the upper buccal segments mesially (especially in growing subjects[7,17]), this difference should come as no surprise. Nonextraction treatment, in contrast, commonly "distalizes" the maxillary buccal segments (Table VI; Paquette et al.[5]). Perhaps as a result, four of the five patients whose mandibles appeared to undergo posterior displacement were in the nonextraction sub-group. Clearly, these results are difficult to explain within the framework of the various posterior-displacement hypotheses.

Ultimately, the issue of condylar position will probably prove to be a "red herring" whose surprising longevity is a function of dentistry's historical infatuation with mechanistic and biologically parochial conjecture. In the real world, however, the choice of treatment strategies—extraction or nonextraction—appears to be of little long-term functional significance: both sets of patients presented with about the same number of signs and symptoms of dysfunction. This is the same result that was obtained when statistically-defined borderline treatment samples were recalled (extraction and nonextraction orthodontics[6]; orthodontics and orthognathic surgery[38]). Although each of these studies found a low prevalence of signs and symptoms and no significant between-treatment differences, the present patients had a slightly higher average, about 4 signs/symptoms per subject. This variation among workers and studies underscores the difficulty of comparing CMD data, even when pains have been taken to calibrate the investigators.[39-41] In any event, all of our within-study comparisons argue that the choice of treatment options is of little functional significance. Moreover, given that the various orthodontic and surgical treatments in these investigations probably had markedly different mean impacts on condylar position, the lack of statistically significant functional differences implies that a small iatrogenic change in condylar position is also of no particular import.

On balance, the present results do not support the notion that one treatment strategy or the other is universally superior, either morphologically or functionally. Rather, they argue that each tends to produce a characteristic repertoire of additive effects and that each, therefore, would be an appropriate answer to a different type of malocclusion. Moreover, the fact that premolar-extraction therapy levied no obvious functional or aesthetic penalties implies that there is no pressing biological need to search for alternatives to extraction. Accordingly, it would seem appropriate at this juncture to call on the leaders of this search to provide comparable long-term data on the performance of flaring and expansion (by any other name) as a routine alternative to extraction. As we wait, it might be well to consider the decades of literature and experience and logic that are arrayed against contemporary efforts to reinvent what in the past proved to be a square wheel.

Table IX
Average increments of change*: Treatment, posttreatment, and net (borderline[5] and extreme data combined)

Measure	Treatment		Post-Treatment ("growth")	Net	
	Extr.	Nonextr.		Extr.	Nonextr.
Linear (mm)					
U1	+4.5	+1.5	-1.0	+3.0	+0.5
L1	-2.0	+1.0	-1.75	-3.75	-1.0
U6	-2.25	+0.75	-2.75	-5.0	-2.0
L6	+3.0	+0.25	+0.75	+3.75	+1.25
L Lip-E	-3.0	-0.5	-2.75	-5.5	-3.5
Angular (°)					
Z-angle	+6.0	+2.5	+3.5	+9.0	+6.0

*U1, U6, + = distal; L1, L6, + = mesial.

Conclusions

The average treatment and posttreatment changes seen in statistically defined, "clear-cut" premolar-extraction and nonextraction Class II edgewise patients were similar to those seen earlier in a study of "borderline" patients. Specifically:

1) Premolar extraction reduces soft- and hard-tissue convexity by 2-3 mm, whereas nonextraction therapy has little effect;

2) In general, posttreatment changes (including an additional convexity reduction) are about the same in both groups;

3) When growth is finished, clear-cut nonextraction patients tend to have "flatter" profiles than do premolar-extraction patients who present with ponderable crowding and spacing;

4) Pre- and posttreatment tooth movements tend to be related to the pattern of jaw-growth; some forms of relapse, therefore, may be a dentoalveolar compensation for residual posttreatment growth;

5) In nonextraction treatment, the upper buccal segments are commonly "distalized," whereas they tend to come forward if premolars have been extracted;

6) There is a tendency for the mandible to be displaced mesially during treatment (extraction more so than nonextraction);

7) In terms of the signs and symptoms commonly associated with CMD, there is apparently no significant long-term difference between extraction and nonextraction therapy.

9) The ability to account for individual variation from the mean future increment, although a formal requirement for true prediction, is not an absolute clinical necessity: because growth and treatment (either extraction or nonextraction) appear to have about the same average effect on a wide range of prognostic sub-groups, clinically useful predictions could be based on the initial face and estimates of the mean increments of change due to growth and treatment.

Acknowledgments

Supported by N.I.D.R. grant DE08716 and by donations from the Orthodontic Education and Research Foundation (St. Louis). This help is greatly appreciated. In addition, the authors wish to thank Drs. John Rugh and James Fricton for introducing us to the "Craniomandibular Index" and Drs. Carroll-Ann Trotman, Fedon A. Livieratos, and James A. McNamara, Jr., for their constructive comments during the preparation of this manuscript.

Based in part on a thesis submitted by the senior author in partial fulfillment of the degree of Master of Science in dentistry, Saint Louis University. It was the recipient of an Award of Special Merit from the American Association of Orthodontists (1993).

Author Address

Lysle E. Johnston, Jr.
Department of Orthodontics and Pediatric Dentistry
School of Dentistry
The University of Michigan
Ann Arbor, MI 48109-1078

S. Luppanapornlarp is an instructor in the Department of Comprehensive Dentistry and a clinical instructor in the Department of Orthodontics, Faculty of Dentistry, Mahidol University, Bangkok, Thailand

L.E. Johnston, Jr., is the Robert W. Browne Professor of Dentistry and Chairman, Department of Orthodontics and Pediatric Dentistry, The University of Michigan, Ann Arbor, MI

References

1. Peck S, Peck H. Frequency of tooth extraction in orthodontic treatment. Am J Orthod 1979;76:491-6.
2. O'Connor BMP. Contemporary trends in orthodontic practice: A national survey. Am J Orthod Dentofac Orthop 1993;103:163-70.
3. Weintraub JA, Vig PS, Brown C, Kowalski CJ. The prevalence of orthodontic extractions. Am J Orthod Dentofac Orthop 1990;96:462-6.
4. Johnston LE Jr, Paquette DE, Beattie JR, Cassidy DW Jr, McCray, JF, and Killiany, D.M.: The reduction of susceptibility bias in retrospective comparisons of alternative treatment srategies. In: Dryland Vig K, Vig PS (eds.) Clinical Research as the Basis of Clinical Practice. Craniofacial Growth Series, Volume 25, Ann Arbor: Center for Human Growth and Development, The University of Michigan, 1991:155-77.
5. Paquette DE, Beattie JR, Johnston LE Jr. A long-term comparison of non-extraction and bicuspid-extraction edgewise therapy in "borderline" Class II patients. Am J Orthod Dentofac Orthop 1992;102:1-14.
6. Beattie, JR, Paquette, DE, Johnston LE Jr. The functional impact of extraction and nonextraction treatments: A long-term comparison in patients with "borderline," equally-susceptible Class II malocclusions. in press, Am J Orthod Dentofac Orthop 1993.
7. Johnston LE Jr. A comparative analysis of Class II treatments. In: McNamara JA Jr., Carlson DS, Vig PS, Ribbens KA (eds.) Science and clinical judgment in orthodontics. Craniofacial Growth Series, Volume 19, Ann Arbor: Center for Human Growth and Development, The University of Michigan, 1986:103-48.
8. Luecke PE III, Johnston LE Jr. The effect of maxillary first premolar extraction and incisor retraction on mandibular position: Testing the central dogma of "functional orthodontics." Am J Orthod Dentofac Orthop 1992;101:4-12.
9. Hashim HA, Godfrey, K. The reproducibility of Johnston's cephalometric superimposition method. Aust Orthod J 1990;11:227-31.
10. Björk A. Prediction of mandibular growth rotation. Am J Orthod 1969;55:585-99.
11. Björk A, Skieller V. Normal and abnormal growth of the mandible. A synthesis of longitudinal cephalometric implant studies over a period of 25 years. Eur J Orthod 1983;5:1-46.
12. Björk A, Skieller V. Postnatal growth and development of the maxillary complex, In: McNamara JA Jr (ed.) Factors affecting the growth of the midface. Craniofacial Growth Series, Volume 6, Ann Arbor: Center for Human Growth and Development, The University of Michigan, 1986:61-99.
13. Little RM. The irregularity index: A quantitative score of mandibular anterior anterior alignment. Am J Orthod 1975;68:554-63.
14. Fricton JR, Schiffman EL. Reliability of a craniomandibular index. J Dent Res 1986;65:1359-64.
15. Fricton JR, Schiffman EL. The craniomandibular index: Validity. J Prosthet Dent 1987;58:222-8.
16. Siegel S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956:116-27.
17. Harris EF, Dyer GS, Vaden JL. Age effects on orthodontic treatment: Skeletodental assessments from the Johnston analysis. Am J Orthod Dentofac Orthop 1991;100:531-6.
18. Eggleston WB Jr, Eckelberry JW. An electromyographic and functional evaluation of treated orthodontic cases. Ann Arbor: Department of Orthodontics, University of Michigan, 1961.
19. Mellion JT. The relationship between orthodontic relapse and the magnitude of centric slides: A clinical and cephalometric study during the first six months of retention. Master's Thesis, St. Louis: Department of Orthodontics, Saint Louis University Medical Center, 1980.
20. Hall KC. The spatial relationship between centric occlusion and centric relation during contemporary Class II therapy. Master's Thesis, Saint Louis: Department of Orthodontics, Saint Louis University Medical Center, 1979.
21. Johnston LE Jr, EICO Orthodontic Study Group of Ohio. A gnathological assessment of centric slides in post-retention othodontic patients. J Prosthet Dent 1988;60:712-5.
22. Behrents RG, Harris EF, Vaden JL, Williams RA, Kemp DH. Relapse of orthodontic treatment results: Growth as an etiologic factor. Bul Charles H Tweed Int Found 1989;17:65-80.
23. Nanda RS, Nanda SK. Considerations of dentofacial growth in long-term retention and stability: Is active retention needed? Am J Orthod Dentofac Orthop 1992;101:297-302.
24. Sinclair PM. The clinical application of orthopedic forces: Current capabilities and limitations. In: Bone biodynamics in orthodontic and orthopedic treatment. Carlson DS, Goldstein SA (eds.), Craniofacial Growth Series, Volume 27, Ann Arbor: Center for Human Growth and Development, c1992:351-88.
25. Björk A. The face in profile. Svensk Tandläk-Tidskr 1948;40(5B):1-180.
26. Lande MJ. Growth of the bony facial profile. Angle Orthod 1952;22:78-90.
27. Johnston LE. A statistical evaluation of cephalometric prediction. Angle Orthod 1968;38:284-304.
28. Greenberg LZ, Johnston LE Jr. Computerized prediction: The accuracy of a contemporary long-range forecast. Am J Orthod 1975;67:243-52.
29. Webb MA. The quantitative effects of maxillary second molar extraction in the correction of Class II, Division 1, malocclusion. Department of Orthodontics, Saint Louis University Medical Center, 1981.
30. Morrell MV. The quantitative effects of maxillary first-bicuspid extraction and maxillary first- and mandibular second-bicuspid extraction in the treatment of Class II, Division 1, malocclusions. Master's Thesis, St. Louis: Department of Orthodontics, Saint Louis University Medical Center, 1984.
31. Brown RK. A quantitative analysis of the relative contributions of maxillary and mandibular molar movement and growth to both the extraction and non-extraction correction of Class II, Division 1 malocclusions with the Begg technique. Master's Thesis, St. Louis: Department of Orthodontics, Saint Louis University Medical Center, 1985.

32. Kesling CK. A cephalometric evaluation of maxillary first molar post-treatment stability as a function of direction of movement during treatment. Master's Thesis, St. Louis: Department of Orthodontics, Saint Louis University Medical Center, 1985.
33. Harper DL. The effects of pre-treatment morphology, treatment change, and post-treatment growth on relapse. Master's Thesis, St. Louis: Department of Orthodontics, Saint Louis University Medical Center, 1986.
34. Kessel JA. A comparative analysis of Class II treatment with the straight wire appliance. Master's Thesis, St. Louis: Department of Orthodontics, Saint Louis University Medical Center, 1987.
35. Habern MD. A cephalometric appraisal of Class II sagittal appliance therapy. Certificate Paper, St. Louis: Department of Orthodontics, Saint Louis University Medical Center, 1990.
36. Little RM, Wallen TR, Riedel RA. Stability and relapse of mandibular anterior alignment: First premolar extraction cases treated by traditional edgewise orthodontics. Am J Orthod 1981;80:349-65.
37. Glenn G, Sinclair PM, Alexander RG. Non-extraction orthodontic therapy: Posttreatment dental and skeletal stability. Am J Orthod Dentofac Orthop 1987;92:37-41.
38. Cassidy DW Jr, Herbosa EG, Rotskoff, KS, Johnston LE Jr. A comparison of Surgery and Orthodontics in "borderline" Class II, division 1 adults. In press, Am J Orthod Dentofac Orthop,1993.
39. Dworkin SF, LeResche S, Von Korff M. Studying the natural history of TMD: Epidemiologic perspectives on physical and psychological findings. In: Dryland Vig K, Vig PS, eds. Clinical Research as the Basis of Clinical Practice. Craniofacial Growth Series, Volume 25, Ann Arbor: Center for Human Growth and Development, The University of Michigan, 1991:39-60.
40. Dworkin SF, LeResche L, DeRouen,T. Assessing clinical signs of temporomandibular disorders: Reliability of clinical examiners. J. Prosthet Dent 1990;63:574-579.
41. Mohlin B, Pilley JR, Shaw WC. A survey of craniomandibular disorders in 1000 12-year-olds. Study design and baseline data in a follow-up study. Eur J Orthod 1991;13:111-123.