

## Outcomes of different Class II treatments

### Comparisons using the American Board of Orthodontics Model Grading System

## Klinische Ergebnisse nach unterschiedlichen Klasse-II-Therapien

### Vergleiche anhand des American Board of Orthodontics Model Grading System

Hatice Akinci Cansunar<sup>1</sup> · Tancan Uysal<sup>2</sup>

Received: 9 November 2013 / Accepted: 3 November 2015 / Published online: 20 April 2016  
© Springer-Verlag Berlin Heidelberg 2016

#### Abstract

**Objective** The aim of this study was to evaluate the clinical outcomes of three different Class II treatment modalities followed by fixed orthodontic therapy, using the American Board of Orthodontics Model Grading System (ABO-MGS).

**Materials and methods** As a retrospective study, files of patients treated at postgraduate orthodontic clinics in different cities in Turkey was randomly selected. From 1684 posttreatment records, 669 patients were divided into three groups: 269 patients treated with extraction of two upper premolars, 198 patients treated with cervical headgear, and 202 patients treated with functional appliances. All the cases were evaluated by one researcher using ABO-MGS. The  $\chi^2$ , Z test, and multivariate analysis of variance were used for statistical evaluation ( $p < 0.05$ ).

**Results** No significant differences were found among the groups in buccolingual inclination, overjet, occlusal relationship, and root angulation. However, there were significant differences in alignment, marginal ridge height, occlusal contact, interproximal contact measurements, and overall MGS average scores. The mean treatment time between the extraction and functional appliance groups was significantly different ( $p = 0.017$ ).

**Conclusion** According to total ABO-MGS scores, headgear treatment had better results than functional appliances. The headgear group had better tooth alignment than the extraction group. Headgear treatment resulted in better occlusal contacts than the functional appliances and had lower average scores for interproximal contact measurements. Functional appliances had the worst average scores for marginal ridge height. Finally, the functional appliance group had the longest treatment times.

**Keywords** Extraoral traction appliances · Tooth extraction · Orthodontic appliances · Functional appliances

#### Zusammenfassung

**Ziel** Unter Verwendung des ABO-MGS (American Board of Orthodontics Model Grading System) sollten die klinischen Ergebnisse von 3 unterschiedlichen Klasse-II-Therapiemodalitäten und anschließender kieferorthopädischer Behandlung mit festsitzenden Apparaturen evaluiert werden.

**Material und Methoden** Ein retrospektives Kollektiv behandelter Patienten wurde randomisiert aus den Archiven kieferorthopädischer Ambulanzen („postgraduate orthodontic clinics“) in verschiedenen türkischen Kliniken ausgewählt. Ausgehend von 1684 Patientenakten nach Behandlung wurden 669 Patienten in 3 Gruppen aufgeteilt: 269 Patienten mit Extraktion zweier oberer Prämolaren, 198 mit zervikalem Headgear und 202 Patienten, die mit funktionellen Apparaturen behandelt worden waren. Sämtliche Fälle wurden von einem Untersucher anhand des ABO-MGS evaluiert. Zur statistischen Evaluierung dienten der  $\chi^2$ -, der Z- und der multivariate ANOVA(“analysis of variance”)-Test ( $p < 0,05$ ).

**Ergebnisse** Keine signifikanten Unterschiede zeigten sich zwischen den Gruppen hinsichtlich bukkolingualer

---

Hatice Akinci Cansunar: Assistant Professor.

---

Tancan Uysal: Professor and Chair. and Prof. Dr.

---

✉ Tancan Uysal  
tancan.uysal@ikc.edu.tr

<sup>1</sup> Department of Orthodontics, Faculty of Dentistry, Inonu University, Malatya, Turkey

<sup>2</sup> Department of Orthodontics, Faculty of Dentistry, Izmir Katip Celebi University, Izmir, Turkey

Inklination/Kippung, Overjet, okklusaler Beziehung und Wurzelangulation, wohl dagegen hinsichtlich Alignment, Randleistenhöhe, okklusalem Kontakt, interproximalen Kontaktmessungen und durchschnittlichen Gesamt-MGS-Scores. Die mittlere Behandlungszeit zwischen der Gruppe mit Extraktion und der mit funktioneller Apparatur war signifikant unterschiedlich ( $p = 0,017$ ).

**Schlussfolgerung** Den ABO-MGS-Gesamtscores entsprechend waren die Ergebnisse nach Headgear-Therapie besser als nach funktioneller Behandlung. Das Alignment war in der Headgear-Gruppe besser als in der Extraktionsgruppe. Die Behandlung mit einem Headgear führte zu besseren okklusalen Kontakten als die funktionelle Behandlung und wies geringere mittlere Scores im Bereich interproximaler Kontaktmessungen auf. Die funktionellen Apparaturen hatten die schlechtesten Scores für die Randleistenhöhe. Schließlich waren die Behandlungszeiten in der Gruppe mit funktioneller Apparatur am längsten.

**Schlüsselwörter** Extraorale Traktionsapparaturen · Zahnextraktion · Kieferorthopädische Apparaturen · Funktionelle Apparaturen

## Introduction

The correction of Class II malocclusions has always been an issue of concern for orthodontists [16]. If maxillary and mandibular arches are malpositioned because of dental problems, skeletal problems, or a combination of both, Class II malocclusion could occur [9]. Class II malocclusion includes various skeletal and dentoalveolar components generally [6].

Patients' age, aspects related to the malocclusion, and financial situation affect the treatment plan for Class II malocclusion [19, 21]. Numerous methods are available to treat Class II malocclusions [13]; however, if a patient is expected to continue to grow, a nonextraction method with extraoral headgear or a removable functional appliance is usually administrated [2, 5, 12, 23]. The most commonly used appliance for extraoral anchorage is headgear [11]. Extraoral appliances are used to correct the dentoalveolar relationship between the mandible and maxilla [11]. The aim, however, is to provide space for retraction of the anterior segment; thus, two upper premolars could be extracted [3, 4, 22]. Two upper premolar extraction protocols require less patient compliance than nonextraction treatment with extraoral headgear or a removable functional appliance [13].

McNamara [15] investigated 277 children with Class II malocclusion and discovered that mandibular skeletal retrusion was most common, while maxillary skeletal protrusion was not common. Therefore, adjustment and

redirection of mandibular growth is the essential goal of most Class II treatment procedures. Although it is apparent that growth adjustment for Class II therapy is very effective in certain individuals, little is known about the mechanisms involved in whether treatment goals are achieved. Therefore, the best timing of Class II therapy and its associated effect on the numerous risks and benefits of therapy is of clinical significance [14].

In 1997, directors of the American Board of Orthodontics (ABO) developed an Objective Grading System (OGS) to objectively quantify tooth positions for scoring posttreatment case records. This system comprises eight measures that are evaluated by using a numeric measure: alignment, marginal ridge height, buccolingual inclination, occlusal relationship, occlusal contact, overjet, interproximal contact, and root angulation. OGS has a specific gauge (ABO measuring gauge) to standardize the measurements by the examiners. This system was implemented in the evaluation of the final casts and panoramic radiographs of each case. The ABO-OGS attempts to assess the outcome of orthodontic treatment. The language was updated and changed from the Objective Grading System to the Model Grading System (MGS) in 2007. Thus, we used the MGS in the present study [9].

There are several studies in the literature assessing different methods for treating Class II malocclusion with cephalometric evaluation or other occlusal indices. No study to date, however, has compared treatment results and duration of two upper premolar extraction, headgear or functional appliances followed by fixed appliance therapy protocols for treating Class II malocclusion with the ABO-MGS.

The aim of this study was to evaluate the clinical outcomes of three Class II treatment modalities (two upper premolar extraction, headgear, and functional orthopedic treatment) followed by fixed orthodontic therapy, using the ABO-MGS. The null hypothesis to be tested states that there is no significant difference in orthodontic treatment outcomes and duration of two upper premolar extraction, headgear, or functional appliance treatments followed by fixed appliance protocols in the treatment of Class II malocclusion.

## Materials and methods

As a retrospective study, files of patients who met the inclusion criteria below were randomly selected from the archives of nine postgraduate orthodontic clinics in different cities in Turkey. In the current study, only cases treated by orthodontists were selected in order to standardize treatment outcomes. The orthodontic competence of the specialists was generally similar.

Advertisement

During the selection of cases, the following criteria were evaluated:

### Inclusion criteria

- Class II patients treated with two upper premolar extraction, headgear, or functional appliances,
- completion of treatment with a fixed orthodontic appliance,
- fixed appliances consisting only of wires and brackets (e.g., no quad helix, rapid palatal expander),
- patients began and completed treatment in the same postgraduate clinic,
- patients' treatment was begun and completed by the same operator, and
- the cases included a final panoramic radiograph.

Cases were excluded if

- Class II treatment was completed with only headgear or functional appliances without fixed orthodontic therapy,
- cases were treated by orthodontic teaching staff,
- cases were treated with four premolar extraction,
- treatment was completed for personal reasons,
- only digital dental casts were available,
- dental plaster casts were broken,
- incomplete/missing records,
- negative chart entries due to the lack of cooperation or poor oral hygiene,
- more than three chart entries signifying broken appliances or brackets, and
- more than three chart entries of missed appointments.

Data collection procedures were retrospectively planned and defined. Pretreatment malocclusion severity was measured using the ABO discrepancy index. Both authors investigated the correlation between pretreatment malocclusion severity and total OGS (MGS) score and found that there was no correlation between pretreatment malocclusion severity and total OGS (MGS) [1]. Thus, the initial malocclusion severity logically did not influence the end result of the orthodontic treatment.

Of the 1684 posttreatment records, 669 patients (347 females and 322 males, average age: 14.3 years at start of treatment) meeting inclusion criteria were divided into three groups: group 1 comprised 269 patients (124 females and 145 males) treated with two upper premolar extraction; group 2 comprised 198 patients (103 females and 95 males) treated with cervical headgear; group 3 comprised 202 patients (120 females and 82 males) treated with a functional appliance (Twin block or fixed functional appliances, Forsus and Herbst). The fixed appliance phase of the orthodontic therapy was combined in all cases (especially

for the headgear group). All cases had pretreatment and posttreatment orthodontic records, including panoramic and cephalometric radiographs, as well as dental casts. All cases were treated using traditional Roth prescription with 0.018-inch bracket slots in all university graduate orthodontic clinics.

The ABO-MGS for scoring dental casts and panoramic radiographs includes eight measures: alignment, marginal ridges, buccolingual inclination, occlusal relationships, occlusal contacts, overjet, interproximal contacts, and root angulation [7]. The ABO measuring gauge was used to score casts. A score of 0 indicates ideal alignment and occlusion, while scores of 1 or 2 indicate deviations from normal [7]. ABO-MGS scores in each of the eight categories and total case scores measured in the cast/radiograph evaluation form were recorded. Treatment time was calculated using the dates of initial application and removal of fixed appliances.

Furthermore, posttreatment MGS scores were used to classify treatment as passing, undetermined, or failing, based on the ABO's instruction that cases with a score of less than 20 commonly passed, and cases with scores more than 30 were generally unsuccessful. Scores of 20–30 were classified as undetermined scores [7].

The principal investigator was initially trained in the ABO-MGS using the ABO Calibration Kit from March 2011 and a tutorial using the ABO gauge. Only one investigator (H.A.C.) evaluated all cases. Radiographs and study casts were both scored by the examiner who was unaware of the group allocation.

### Statistical analysis

A post hoc statistical power analysis indicated that a multivariate analysis of variance (MANOVA) design with 1 factor with 3 groups with an average of 198 ( $n_1 = 246$ ,  $n_2 = 176$ ,  $n_3 = 172$ ) subjects each, for a total of 594 subjects, and 8 response variables achieved 99.9 % power to test factor A if a Wilks' Lambda Approximate  $F$  Test is used with a 5 % significance level.

To assess intraexaminer reliability, a subsample of 20 patients was randomly selected from the main sample. The measurements were repeated 8 weeks after the first measurements. A paired sample  $t$  test was applied to the first and second measurements and the differences between measurements were evaluated.

Treatment type proportions were compared at each quality level using the  $Z$  test with the Bonferroni-adjusted significance levels once a statistically significant relationship was found between the treatment quality and treatment types in question. The aforementioned treatment types were also compared in terms of treatment duration and the

eight MGS component scores using a MANOVA; having obtained a statistically significant multivariate test, the follow-up univariate tests and Bonferroni pair-wise comparisons were examined. Box’s *M* test was applied, and it was observed variance–covariance matrices of the outcome variables were equal across groups (Box’s *M* = 4.013, *F* = 0.666, *p* = 0.677) as well as the Levene’s univariate tests for homogeneity of group variances which also yielded nonsignificant results for the treatment time (*p* = 0.050) and the MGS (*p* = 0.555). Having homogeneous variances across groups, on the other hand, the outcome variables violated the normality assumption based on the Shapiro–Wilk’s test. However, we believe that, under many conditions, violating the multivariate assumption, especially when variances are equal, does not necessarily invalidate the results. Departures from multivariate normality generally have only very slight effects on type I error rates of the four MANOVA statistics, but Roy’s greatest characteristic root may sometimes be an exception. IBM SPSS Statistics version 20 was used for the analyses. When the *p* value was less than 0.05, the statistical test was considered significant.

**Results**

The paired sample *t* test results for the intra-examiner reliability indicated that the first and second measurements and the differences between measurements were insignificant (*p* = 0.625). The total mean treatment times for the three groups are shown in Table 1. A significant difference was found in the mean treatment time between the two upper premolar extraction group (mean 27.60 ± 11.4 months) and the functional appliance group (mean 30.38 ± 11.2 months; *p* = 0.017). Treatment time for the headgear group (mean 28.69 ± 12.6 months) was not significantly different from the two upper premolar extraction and functional appliance groups.

Mean MGS scores for the eight measured variables and descriptive statistics for each component are given in

Table 2. A significant difference was found in alignment between the extraction and headgear groups (*p* = 0.027). The headgear group (mean MGS score = 2.22) had better tooth alignments than the extraction group (mean MGS score = 2.63).

Regarding marginal ridge height measurements, a significant difference was found between the headgear and functional appliance groups (*p* = 0.021). The functional appliance group (mean MGS score = 3.81) had a higher average score for marginal ridge height measurements.

No significant difference was found among three groups in buccolingual inclination (*p* = 0.183), overjet (*p* = 0.696), occlusal relationship (*p* = 0.185), and root angulation (*p* = 0.092).

A significant difference was found in occlusal contact between the headgear group and functional appliance group (*p* = 0.034). The headgear group (mean MGS score = 2.73) had better occlusal contact than the functional appliance group (mean MGS score = 3.65).

Interproximal contact measurements showed significant differences between the two upper premolar extraction and headgear groups (*p* = 0.003). Similarly, significant differences were found in interproximal contacts between the two upper premolar extraction group and functional appliance group (*p* = 0.041). The headgear group (mean MGS score = 2.73) had a lower average score for interproximal contact measurements.

Based on the significant differences in the variables and in overall MGS average scores, the null hypothesis of this study was rejected. When comparing the overall MGS average scores of the three groups, significant differences were found in alignment measurements between the headgear group and functional appliance group (*p* = 0.026; Table 3). The headgear group (mean MGS score = 16.80) had better ABO-MGS scores than the functional appliance group (mean MGS score = 19.05). However, cases classified as passing were not significantly different among two upper premolar extraction (60 %), headgear (66 %), and functional appliance groups (60 %; Fig. 1).

**Tab. 1** Mean treatment time (months) for the two upper premolar extraction, headgear, and functional appliances groups  
**Tab. 1** Durchschnittliche Behandlungszeit (Monate) für die Gruppen mit 2 Prämolarenextraktionen, Headgear und funktionellen Apparaturen

Groups	N	Mean (month)	SD	Min.	Max.	F values	Sig.	Pairwise comparisons		
								2-upper-premolar extraction	Headgear	Functional Appliance
2-upper-premolar extraction	269	27.60	11.45	26.19	29.01	4.13	0.017*		0.194	0.017*
Headgear	198	28.69	12.66	27.04	30.34			0.194		1.000
Functional appliance	202	30.38	11.24	28.75	32.00			0.017*	1.000	
Total	669	28.76	11.80	26.99	28.82					

N sample size, SD standard deviation, Min minimum, max maximum, Sig. significance

\* *p* < 0.05, partial-η<sup>2</sup> = 0.011

**Tab. 2** Descriptive values and multiple comparisons  
**Tab. 2** Deskriptive Werte und mehrere Vergleiche

ABO-OGS Parameter	Group	N	Mean	Std. Deviation	Minimum	Maximum	F values	Sig.	Pairwise comparisons		
									2-upper- premolar extraction	Headgear	Functional Appliance
Alignment	2-upper- premolar extraction	269	2.63	1.79	0	9	4.440	0.012*		0.027*	0.052
	Headgear	198	2.22	1.64	0	8			0.027*		1.000
	Functional	202	2.26	1.43	0	8			0.052	1.000	
Marginal ridge height	2-upper- premolar extraction	269	3.32	2.34	0	11	3.989	0.019*		1.000	0.097
	Headgear	198	3.14	2.38	0	10			1.000		0.021*
	Functional	202	3.81	2.27	0	10			0.097	0.021*	
Buccolingual inclination	2-upper- premolar extraction	269	1.72	1.77	0	12	1.706	0.183		0.922	0.206
	Headgear	198	1.57	1.72	0	8			0.922		1.000
	Functional	202	1.45	1.49	0	6			0.206	1.000	
Overjet	2-upper- premolar extraction	269	4.72	3.94	0	16	0.363	0.696		1.000	1.000
	Headgear	198	4.44	3.19	0	18			1.000		1.000
	Functional	202	4.52	3.33	0	16			1.000	1.000	
Occlusal contact	2-upper- premolar extraction	269	3.43	3.44	0	17		0.034*		0.121	1.000
	Headgear	198	2.73	3.17	0	15	3.403		0.121		0.043*
	Functional	202	3.65	3.62	0	15			1.000	0.043*	
Occlusal relationship	2-upper- premolar extraction	269	1.84	2.72	0	13	1.692	0.185		1.000	0.203
	Headgear	198	2.12	2.90	0	13			1.000		1.000
	Functional	202	2.40	3.36	0	20			0.203	1.000	
Interproximal contact	2-upper- premolar extraction	269	0.37	0.86	0	4	6.347	0.002**		0.003**	0.041*
	Headgear	198	0.14	0.56	0	4			0.003**		1.000
	Functional	202	0.20	0.63	0	4			0.041*	1.000	
Root angulation	2-upper- premolar extraction	269	0.19	0.52	0	2	2.394	0.092		0.087	1.000
	Headgear	198	0.07	0.30	0	2			0.087		0.728
	Functional	202	0.14	0.69	0	8			1.000	0.728	

The mean difference is significant at the 0.05 level

N sample size, Sig. significance

\*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ , Multivariate Tests:  $F(18, 1166) = 2.935$ ,  $p = 0.000$ , partial- $\eta^2 = 0.043$

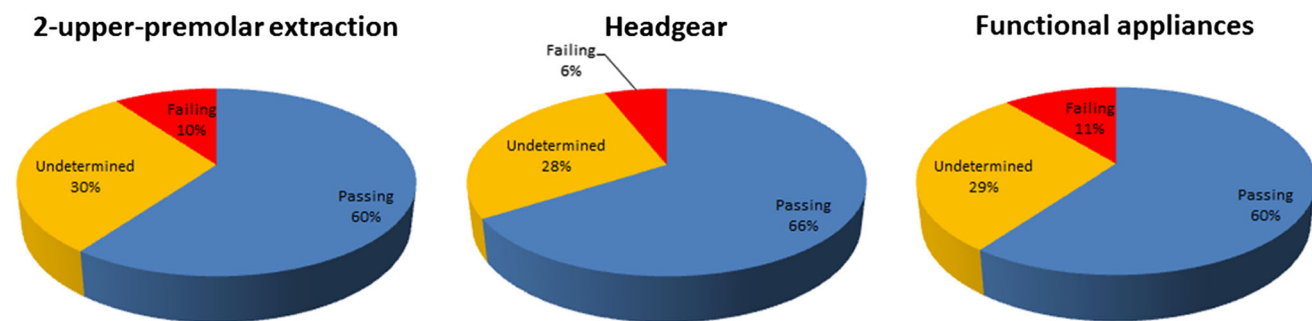
**Tab. 3** Overall average Objective Grading System (OGS) scores for the two upper premolar extraction, headgear, and functional appliances groups

**Tab. 3** Durchschnittliche OGS (Objective Grading System)-Scores der Gruppen mit 2 Prämolarenextraktionen, Headgear und funktionellen Apparaturen

Groups	N	Mean	SD	Min.	Max.	F values	Sig.	Pairwise comparisons		
								2-upper-premolar extraction	Headgear	Functional appliance
2-upper-premolar extraction	269	18.34	8.65	17.31	19.36	3.647	0.026*	0.165	1.000	
Headgear	198	16.80	8.54	15.60	17.99			0.165	0.026*	
Functional appliance	202	19.05	8.41	17.87	20.24			1.000	0.026*	

N sample size, SD standard deviation, Min minimum, Max maximum, Sig. significance

\*  $p < 0.05$ , partial- $\eta^2 = 0.011$



**Fig. 1** Percentage of cases classified as passing, undetermined, and failing based on the posttreatment Model Grading System (MGS) score for the three different treatment groups

**Abb. 1** Prozentsätze von als erfolgreich, unbestimmt bzw. insuffizient eingestuftem Fällen in den 3 Behandlungsgruppen, basierend auf dem MGS(Model Grading System)-Score nach abgeschlossener Behandlung

**Discussion**

There are several studies in the literature assessing different methods to treat Class II malocclusion with cephalometric evaluation or other occlusal indices. Occlusal indices are helpful for clinicians in diagnosis, research design, decision-making, evaluating orthodontic treatment need, and clinical outcomes [7, 17, 18].

The Peer Assessment Rating (PAR) index was developed to assess malocclusion. Comprehensive clinical assessment includes various factors, e.g., facial and dental esthetics, arch form, vertical control, root resorption, periodontal health, and treatment efficiency.

Validity and reliability of the ABO-MGS were confirmed and consequently used in the evaluation of orthodontic records [17]. The ABO-MGS provides a method to objectively assess outcome and success of orthodontic treatment [7]. Onyeaso and Begole [18] found that the standards involved in the ABO-MGS are more strict than those used in the Peer Assessment Rating (PAR) or the index of complexity outcome and need (ICON) for assessing the orthodontic treatment outcome. Thus, we

chose the ABO-MGS to evaluate different treatment modalities in Class II cases.

Comparison of treatment results and duration of two upper premolar extraction, headgear, or functional appliances followed by fixed appliance therapy protocols with the ABO-MGS had not been previously reported. Therefore, the present study examined dental casts and panoramic radiographs to evaluate orthodontic treatment results and duration of two upper premolar extraction, headgear, or functional appliance treatments when treating Class II malocclusions with the ABO-MGS.

Okunami et al. [17] assessed differences between digital and plaster dental casts to score the ABO-MGS. However, they reported that the recent digital program is inadequate for scoring all parameters as required by the ABO-MGS. Thus, only plaster models were used and digital models were excluded in the current study.

A significant difference was found in the mean treatment times between the two upper premolar extraction (mean  $27.60 \pm 11.4$  months) and functional appliance (mean  $30.38 \pm 11.2$  months) groups. Regarding reasons that may affect the length of orthodontic treatment, Fink and

Smith [10] found a statistically significant relationship in 4 of the 18 variables examined (pretreatment ANB angle, extraction of premolars, pretreatment mandibular plane angle, and the number of broken appointments). Almeida-Pedrin et al. [8] compared the duration of Class II treatment with the cervical headgear, pendulum appliance, and extraction of two maxillary premolars followed by fixed appliance therapy. Extraction of two maxillary premolars followed by fixed appliance had the shortest treatment duration. While it is attempted to produce skeletal and dental effect with functional treatment, only a dental effect is achieved with two upper premolar extraction treatment. Long treatment duration is not a surprise for functional treatment.

A significant difference was found in alignment between the two upper premolar extraction and headgear groups. Xu et al. [24] investigated extraction versus nonextraction orthodontic treatment and reported no significant differences for tooth alignment, midline symmetry, overbite, overjet, or posterior occlusion between the groups. However, in the present study, we found that the headgear group had better tooth alignment than the extraction group. Headgear treatment is a nonextraction treatment modality; thus, it may have affected these outcomes.

The headgear and functional appliance groups had significant differences regarding the marginal ridge height and occlusal contact measurements. The headgear group had better marginal ridge height measurements and occlusal contact than the functional appliance group. However, the functional appliances group had long treatment duration.

Interproximal contact measurements were significantly different between the two upper premolar extraction and headgear groups. Similarly, a significant difference was found in interproximal contacts between the two upper premolar extraction and functional appliance groups. Premolar extraction was associated with more crowding, more serious buccal segment occlusion, greater overjet, and a larger midline deviation. Because it is a nonextraction alternative, if there is less crowding or less serious buccal segment occlusion, headgear treatment can be preferred. In the present study, it was found that the headgear group had a lower average score and better values for interproximal contact measurements. Thus, pretreatment characteristics might affect treatment outcomes.

Based on the significant differences in overall MGS average scores, the null hypothesis of this study was rejected. When comparing the overall MGS average scores of the three groups, significant differences were found in alignment measurements between the headgear and functional appliance groups. Pinzan-Vercelino et al. [20] compared the outcomes and duration of Class II malocclusion treatment with two maxillary premolar extractions and the pendulum appliance protocol with study models and initial cephalograms. Using the PAR index used to evaluate occlusal

outcomes, they stated that the posttreatment occlusal position is similar between the groups. However, because the PAR index measures only one outcome of treatment, i.e., straight teeth, and requires both pre- and posttreatment casts to generate a valid score, the PAR might not capture all the fine details of dental alignment, while MGS evaluates posttreatment models, capturing all fine details.

Almeida-Pedrin et al. [8] compared the cephalometric effects, dental-arch changes, and efficacy of Class II treatment with the cervical headgear, pendulum appliance, or extraction of two maxillary premolars followed by fixed appliance therapy. They stated that the effects of treatment with cervical headgear or pendulum appliance and extraction of two maxillary premolars followed by fixed appliances were similar from both cephalometric and occlusal results. In this study, the headgear group finished better than functional appliance group according to ABO-MGS. However, cases classified as passing were not significantly different [two upper premolar extraction (60 %), headgear (66 %), and functional appliance groups (60 %)].

There are several limitations of this study, including the lack of gender differentiation, inclusion of removable and fixed functional appliance treatments into a single group, and—to obtain large sample sizes—selection of cases from the archives of nine postgraduate university orthodontic clinics in different cities. If the cases had been selected from only one archive, the resulting sample size would not have reached the power of the current analyses. An important consideration is that the duration and quality of treatment parallel the practitioner's level of orthodontic experience; thus, in the current research we used cases that were treated only by experienced residents in orthodontics. The orthodontic competence of the residents was generally similar (in the 3rd or 4th year of orthodontic postgraduate education). All included cases were treated with traditional Roth prescription with 0.018-inch bracket slots.

In Class II malocclusion, treatment duration and the occlusal outcomes are associated with patient compliance to achieve the correct molar relationship or to continue the molar relationship during anterior tooth retraction. In spite of the necessity of patient compliance in the three protocols evaluated in present study, patient compliance should be higher or stricter in the headgear and functional appliance groups. Although the headgear group did not have shorter treatment duration, the cases were considered well finished according to ABO-MGS. Thus, headgear distalization is preferred to correct Class II malocclusion.

## Conclusion

- The outcome for the headgear group was better than that for the functional appliances group according to



ABO-MGS scores. However, cases classified as passing were not significantly different among the three groups.

- The longest treatment time was found for the functional appliance group.
- The headgear group had better tooth alignments than the two upper premolar extraction group.
- The functional appliance group had a higher average score for marginal ridge height measurements.
- The headgear group had better occlusal contact than the functional appliance group. The headgear group had a lower average score for interproximal contact measurements.

**Acknowledgments** The authors of this study thank Ataturk, Baskent, Cumhuriyet, Cukurova, Dicle, Erciyes, Inonu, Selcuk, Suleyman Demirel Universities' Faculties of Dentistry, Departments of Orthodontics for permitting use their archives. The authors would like to thank Dr. Bulent Ozkan for his advice and guidance on the statistical analysis.

#### Compliance with ethical guidelines

**Conflict of interest** H.A. Cansunar and T. Uysal state that there are no conflicts of interest. The accompanying manuscript does not include studies on humans or animals.

#### References

1. Akinci Cansunar H, Uysal T (2014) Relationship between pre-treatment case complexity and orthodontic clinical outcomes determined by the American Board of Orthodontics criteria. *Angle Orthod* 84:974–979
2. Arvystas MG (1985) Nonextraction treatment of Class II, division 1 malocclusion. *Am J Orthod* 88:380–395
3. Baumrind S, Korn EL, Boyd RL, Maxwell R (1996) The decision to extract: part I. Interclinician agreement. *Am J Orthod Dentofac Orthop* 109:297–299
4. Baumrind S, Korn EL, Boyd RL, Maxwell R (1996) The decision to extract: part II. Analysis of clinicians' stated reasons for extraction. *Am J Orthod Dentofac Orthop* 109:393–402
5. Bishara SE, Ziaja RR (1989) Functional appliances: a review. *Am J Orthod Dentofac Orthop* 95:250–258
6. Burkhardt D, McNamara JA, Baccetti T (2003) Maxillary molar distalization or mandibular enhancement: a cephalometric comparison of comprehensive orthodontic treatment including the pendulum and the Herbst appliances. *Am J Orthod Dentofac Orthop* 123:108–116
7. Casco JS, Vaden JL, Kokich VG, Damone J, James RD et al (1998) Objective grading system for dental casts and panoramic radiographs. *Am J Orthod Dentofac Orthop* 114:589–599
8. De Almeida-Pedrin RR, Henriques JF, de Almeida RR, de Almeida MR, McNamara JA (2009) Effects of the pendulum appliance, cervical headgear, and 2 premolar extractions followed by fixed appliances in patients with Class II malocclusion. *Am J Orthod Dentofac Orthop* 136:833–842
9. De Oliveira JN Jr, de Almeida RR, de Almeida MR, de Oliveira JN (2007) Dentoskeletal changes induced by Jasper jumper and cervical headgear appliances followed by fixed orthodontic treatment. *Am J Orthod Dentofac Orthop* 132:54–62
10. Fink DF, Smith RJ (1992) The duration of orthodontic treatment. *Am J Orthod Dentofac Orthop* 102:45–51
11. Firouz M, Zernik J, Nanda R (1992) Dental and orthopedic effects of high pull headgear in treatment of Class II division 1 malocclusion. *Am J Orthod Dentofac Orthop* 102:197–205
12. Graber TM (1969) Maxillary second molar extraction in Class II malocclusion. *Am J Orthod* 56:331–353
13. Janson G, Graciano JT, Henriques JF, Freitas MR, Pinzan A, Pinzan-Vercelino CR (2006) Occlusal and cephalometric Class II division 1 malocclusion severity in patients treated with and without extraction of 2 maxillary premolars. *Am J Orthod Dentofac Orthop* 129:759–767
14. King GJ, Keeling SD, Hocevar RA, Wheeler TT (1990) The timing of treatment for Class II malocclusions in children: a review of the literature. *Angle Orthod* 60:87–97
15. McNamara JA (1981) Components of Class II malocclusion in children 8–10 years of age. *Angle Orthod* 51:177–202
16. Mossaz CF, Byloff FK, Kiliaridis S (2007) Cervical headgear vs pendulum appliance for the treatment of moderate skeletal Class II malocclusion. *Am J Orthod Dentofac Orthop* 132:616–623
17. Okunami TR, Kusnoto B, BeGole E, Evans CA, Sadowsky C, Fadavi S (2007) Assessing the American Board of Orthodontics objective grading system: digital vs plaster dental casts. *Am J Orthod Dentofac Orthop* 131:51–56
18. Onyeaso CO, Begole AE (2007) Relationship between index of complexity, outcome and need, dental aesthetic index, peer-assessment rating index, and American Board of Orthodontics objective grading system. *Am J Orthod Dentofac Orthop* 131:248–252
19. Petrone J, Fishell J, Berk NW, Kapur R, Sciote J, Weyant RJ (2003) Relationship of malocclusion severity and treatment fee to consumer's expectation of treatment outcome. *Am J Orthod Dentofac Orthop* 124(1):41–45
20. Pinzan-Vercelino CRM, Janson G, Pinzan A, de Almeida RR, de Freitas MR, de Freitas KMS (2009) Comparative efficiency of Class II malocclusion treatment with the pendulum appliance or two maxillary premolar extractions and edgewise appliances. *Eur J Orthod* 31:333–340
21. Proffit WR, Tulloch JF (2002) Preadolescent Class II problems: treat now or wait? *Am J Orthod Dentofac Orthop* 121(6):560–562
22. Proffit WR (1994) Forty-year review of extraction frequencies at a university orthodontic clinic. *Angle Orthod* 64:407–414
23. Vargervik K, Harvold EP (1985) Response to activator treatment in Class II malocclusions. *Am J Orthod* 88:242–251
24. Xu TM, Liu Y, Yang MZ, Huang W (2006) Comparison of extraction versus nonextraction orthodontic treatment outcomes for borderline Chinese patients. *Am J Orthod Dentofac Orthop* 129:672–677