

Comparison of Treatment Outcome Assessment for Class I Malocclusion Patients: Peer Assessment Rating versus American Board of Orthodontics-Objective Grading System

Mihee Hong¹, Yoon-Ah Kook², Seung-Hak Baek³, Myeng-Ki Kim⁴

¹Department of Orthodontics, The Catholic University of Korea, Uijeongbu St. Mary's Hospital, Uijeongbu,

²Department of Orthodontics, The Catholic University of Korea, Seoul St. Mary's Hospital, Seoul,

³Department of Orthodontics, Dental Research Institute, School of Dentistry, Seoul National University, Seoul,

⁴Department of Dental Services Management and Informatics, School of Dentistry, Seoul National University, Seoul, Korea

Purpose: The purpose of this retrospective study is to investigate the degree of coincidence between the peer assessment rating (PAR) index and American Board of Orthodontics objective grading system (ABO-OGS) in the assessment of orthodontic treatment outcomes of Class I malocclusion cases.

Materials and Methods: The sample consisted of 26 Class I patients. The PAR index was used for evaluation of pre- (T0) and posttreatment (T1) casts, and the ABO-OGS for assessment of T1 casts. If there was a reduction in PAR scores from T0 to T1 of more than 30%, the label 'PAR+' was given to the case, and if not, it was labeled 'PAR-'. If the ABO-OGS was less than 27, the label 'OGS+' was given to the case and if not, it was labeled 'OGS-'. 'A PAR-only qualified group' (PAR+), 'ABO-OGS-only qualified group' (OGS+), 'both indices qualified group' (PAR+/OGS+), and 'both indices disqualified group' (PAR-/OGS-) were compared with a Wilcoxon rank-sum test, sensitivity/specificity test and Spearman's correlation test.

Result: PAR scores for T0, T1, and percentage reduction were 21.1, 6.4, and 65.9%, respectively, and 35.4 for ABO-OGS. The distribution of the 'PAR+/OGS+', 'PAR+', and 'PAR-/OGS-' group was 19.3%, 76.9%, and 3.8%, respectively. The T0-PAR, T1-PAR and PAR point reductions for the 'PAR+' group were significantly higher than those of 'PAR+/OGS+' groups (23.1 vs. 15.6; 6.7 vs. 4.6; and 16.5 vs. 11.0; all $P < 0.05$). However, the PAR-percentage reduction and treatment duration between the two groups were not statistically different (70.0% vs. 67.0%, $P = 0.4325$; 24.1 months vs. 25.0 months, $P = 0.4057$). The T1-ABO-OGS score for 'PAR+' group was significantly higher than that of the 'PAR+/OGS+' groups (38.2 vs. 24.0, $P < 0.001$).

Conclusion: Since the fraction of the 'PAR+/OGS+' group was less than 20% and there was no significant correlation between PAR-percentage reduction and T1-ABO-OGS, development of a new index system for the accurate evaluation of treatment outcome is needed.

Key Words: American Board of Orthodontics objective grading system; Angle Class I malocclusion; Health care quality indicators; Orthodontics; Outcome assessment; Peer assessment rating index

Corresponding Author: **Myeng-Ki Kim**

Department of Dental Services Management and Informatics, School of Dentistry, Seoul National University, 101 Daehak-ro, Jongno-gu, Seoul 110-749, Korea

TEL : +82-2-740-8791, FAX : +82-2-743-7633, E-mail : meeree@snu.ac.kr

Received for publication April 27, 2014; Returned after revision May 26, 2014; Accepted for publication June 3, 2014

Copyright © 2014 by Korean Academy of Dental Science

© This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

The subjective, experience-based clinician's judgment for orthodontic outcome has been the gold standard in most clinical settings. However, practitioner variation was highlighted in the largest international study undertaken to assess orthodontic treatment needs and outcomes¹⁾. Therefore, indices were developed to objectively measure treatment outcomes and to represent the professional standard²⁾.

Among the various outcomes³⁾ such as occlusal, skeletal, soft tissue and functional, the occlusal outcome is the basic unit. There are two representative occlusal indices: the peer assessment rating (PAR) index and the American Board of Orthodontics objective grading system (ABO-OGS). The PAR index was developed to provide an objective assessment of treatment outcomes and has been used extensively as a method of audit in Europe⁴⁾. On the other hand, the ABO-OGS helps set standards of excellence for orthodontists in America and is used in the phase III candidate case report examination to evaluate the candidates' best work⁵⁾.

Onyeaso and Begole⁶⁾ mentioned that the ABO-OGS required more-stringent standards than the PAR for assessing the outcome of orthodontic treatment. Fox and Chapple⁷⁾ showed that different occlusal indices gave different failure rates, despite the same patients.

Since PAR and ABO-OGS were introduced to the Korean orthodontic society, several outcome studies have been performed. Kim et al.⁸⁾ found that treatment duration was correlated with extraction and the PAR index. Cheong and Park⁹⁾ in an evaluation of the study casts of 25 patients with permanent dentition, using the PAR index, concluded that the treatment reduced the PAR index by 85.05%. Son et al.¹⁰⁾ quantitatively determined the contributing factors of relapse tendency after orthodontic treatment using the ABO-OGS.

Although the PAR index focuses on improvement by treatment, the ABO-OGS evaluates the completeness of the results. To our knowledge, there has been no published study showing that the ABO-OGS and PAR were used to evaluate orthodontic treatment outcomes for the same subjects in Korea. Professional agreements on outcome quality need to be searched for continuously to contribute to our professional standards. How are these indices related to each other in assessing the same treatment results? How do they agree on the best or worst outcome?

The purpose of this retrospective study is to investigate the degree of coincidence between the PAR index and ABO-OGS in the assessment of orthodontic treatment outcomes of Class I malocclusion patients. The null hypothesis of this study was that there was no significant difference between the PAR index and ABO-OGS in the assessment of orthodontic treatment outcomes of Class I malocclusion patients.

Materials and Methods

1. Samples

The samples consisted of 26 Angle's classification Class I patients (eight males and 18 females; mean pretreatment age, 26 yr 7 mo±10 yr 11 mo; mean treatment duration, 24.6±10.0 months) treated from March 2010 to August 2013 at Seoul St. Mary's Hospital, The Catholic University of Korea. This retrospective study was reviewed and approved by the Institutional Review Board of The Catholic University of Korea (KC13RISI0655).

The inclusion criteria are as follows: Class I malocclusion and permanent dentition at an initial stage and cases where both pre- and posttreatment dental casts were made. In order to improve the homogeneity of the samples, the cases that had mixed dentition, were treated with orthognathic surgery, or had partial orthodontic treatment for prosthesis were excluded.

Table 1. Demographic data of the samples

Variable	Number (%)
Gender	Male 8 (30.8)
	Female 18 (69.2)
Extractions	Yes 12 (46.2)
	No 14 (53.8)
Age at diagnosis (yr)	<20 8 (30.8)
	≥20 18 (69.2)

The ratio of gender was approximately 7 : 3 female to male (Table 1). The percentage of non-extraction treatment was higher than extraction treatment (53.8% vs. 46.2%) (Table 1). Adult patients (defined as ≥20 years old) were 69.2% (Table 1).

2. Outcome Assessment by PAR Index

Both pre- (T0) and posttreatment (T1) casts were measured with the PAR index. The seven components of the PAR Index were upper labial segment alignment, lower labial segment alignment, right buccal occlusion, left buccal occlusion, overjet, overbite, and centerline to evaluate changes in the PAR score, PAR-point and PAR-percentage reductions were used. The PAR-point reduction was the difference between T0 and T1 scores. The PAR-percentage reduction was calculated as follows: % PAR=[(T0-PAR score-T1-PAR score)×100]/T0-PAR score.

According to Richmond et al.¹¹⁾, a case was judged as “improved” if a 30% reduction in PAR score was achieved. If there was more than a 30% reduction in PAR scores from T0 to T1, the label ‘PAR+’ was given to the case and if not, it was labeled PAR-.

3. Outcome Assessment by ABO-OGS

Only T1 casts were measured for the ABO-OGS score. Seven components, except for root angulations, which can be scored from a panoramic radiograph, were measured to obtain the total ABO-OGS score. The components are as follows: alignment/rotations, marginal ridges, buccolingual inclination, overjet, occlusal contacts, occlusal

relationships, and interproximal contacts.

If the ABO-OGS was less than 27, the label ‘OGS+’ was given to the case and if not, it was labeled ‘OGS-’. The acceptance level of the cast/radiographic evaluation for the ABO was defined as 27 points. Although this study didn’t include the radiographic criterion of ‘root angulations’ for comparison with the PAR index, it followed the acceptance level. Therefore the level of standard is more lenient than that of the original OGS evaluation by the difference of in points for root angulations: no points for parallel roots, one point for nonparallel roots, and two points for a contacting adjacent root¹²⁾.

4. Verification of Intra-operator Reproducibility

Period measured both T0 and T1 casts for all the components of PAR and OGS. All variables from those cases were re-measured after two months by the same examiner. Additionally, for several cases where the point difference between two measurements was more than seven, without referencing the previous data, a third measurement was fulfilled by the same operator. In these cases, clinical photos and bites were used as references. Intra-examiner reliability was assessed by intra-class correlation (ICC).

Three sets of 10 random numbers were generated in an Excel program (Microsoft, Redmond, WA, USA) and applied to the three sets of databases, T0-PAR, T1-PAR and OGS, respectively. An ICC for each 10-case set was calculated for each data set (0.9444, 0.7291, and 0.9723). An excellent agreement between the first and second measurements was demonstrated by a correlation coefficient of 0.94 for T0-PAR and 0.97 for the ABO-OGS. However, T1-PAR showed a middle level of correlation between the two measurements.

5. Statistical Analysis

A paired t-test was used to test the difference between T0- and T1-PAR scores of each paired

Table 2. Comparison of PAR index and ABO-OGS

Sample (n=26)	PAR index				ABO-OGS	Treatment duration (mo)
	T0	T1	T0~T1	Reduction (%)	T1	
Mean±SD	21.1±6.7	6.4±13.0	14.7±7.1***	65.9±20.4	35.4±9.6	24.6±6.1

PAR: peer assessment rating, ABO-OGS: American Board of Orthodontics objective grading system, SD: standard deviation, T0: pretreatment, T1: posttreatment.

A paired t-test was performed.

***P<0.001.

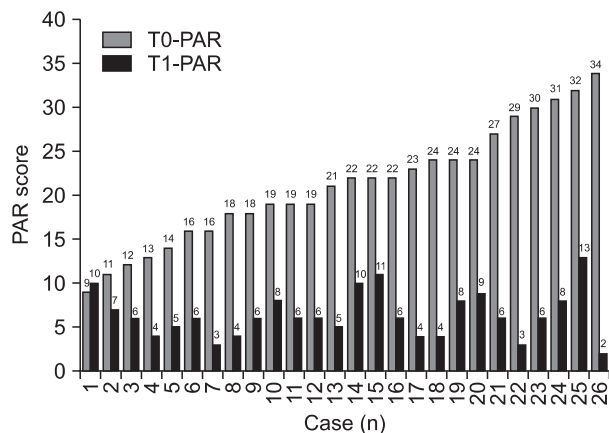


Fig. 1. Distribution of the values of T0-PAR and T1-PAR scores. The values on the top of the bars indicate the PAR score for each case. PAR: peer assessment rating, T0: pretreatment, T1: posttreatment.

case by treatment (SAS Institute, Cary, NC, USA). Spearman’s rank correlation test was applied for assessing the relationships among T0- and T1-PAR scores, the ABO-OGS and treatment duration.

Fisher’s exact test was chosen as an independence test for the four different acceptance level groups; the PAR-only qualified (PAR+) group, the ABO-OGS-only qualified (OGS+) group, the both indices qualified (PAR+/OGS+) group, and the both indices disqualified (PAR-/OGS-) group. A sensitivity/specificity test was used to understand the relationship between PAR and ABO-OGS through these four groups. A Wilcoxon rank-sum test was used to compare overall index scores and each component of the indices between the ‘PAR+’ group and ‘PAR+/OGS+’ group.

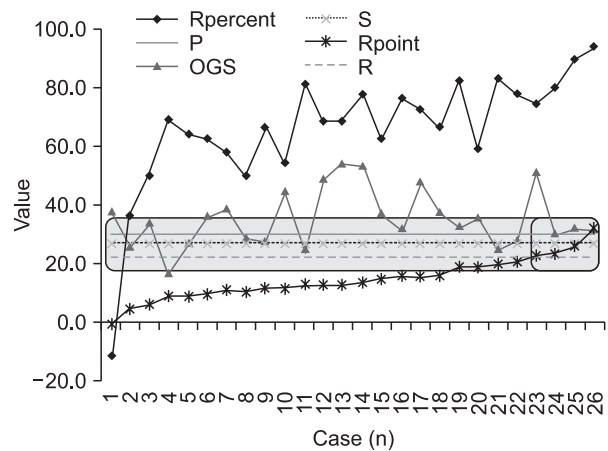


Fig. 2. Distribution of the values of the assessment indices. In the small square box at the left side of the rectangular area, cases 23 to 26 represent “greatly improved cases,” but their OGS scores are above S (=27), therefore, these cases belong to the only PAR+ group. PAR: peer assessment rating, ABO-OGS: American Board of Orthodontics objective grading system, Rpercent: PAR-percent reduction, P: 30% reduction cut off line; Rpoint above this line is PAR+ and Rpoint below this line is PAR-, OGS: ABO-OGS score, S: ABO-OGS cut off=27 line; OGS score above 27 line means OGS- and OGS below this line is OGS+, Rpoint: PAR-point reduction, R: PAR-point reduction=22 line; Rpoint above 22 line means the “greatly improved “cases” of the PAR index.

Result

The mean PAR scores for T0, T1, and PAR-percentage reduction were 21.1, 6.4, and 65.9%, respectively, and the ABO-OGS for T1 was 35.4 (Table 2). Although the fraction of patients with more than 10 points on the T0-PAR score was 96.1%, that with more than 10 points on a T1-PAR score was 11.5% (Fig. 1). Usually, fewer than 10 points means the occlusion of the cast would close to the ideal occlusion⁹. Therefore, 84% of the sample that had more than 10 points at T0 were

Table 3. Correlation among T0-PAR, T1-PAR, T1-ABO-OGS, and PAR-percentage reduction

	T0-PAR	T1-PAR	T1-ABO-OGS	PAR-percentage reduction	Treatment duration (mo)
T0-PAR	1.0	0.0348 (P=0.8657)	0.1236 (P=0.5473)	0.5088 (P=0.0079*)	-0.4129 (P=0.0361*)
T1-PAR		1.0	0.3872 (P=0.0506)	-0.7945 (P<0.001***)	0.0731 (P=0.7226)
T1-ABO-OGS			1.0	-0.1914 (P=0.3488)	-0.1179 (P=0.5662)
PAR-percentage reduction				1.0	-0.3182 (P=0.1132)
Treatment duration (mo)					1.0

T0: pretreatment, T1: posttreatment, PAR: peer assessment rating, ABO-OGS: American Board of Orthodontics objective grading system.

A Spearman's rank correlation test was performed.

*P<0.05, ***P<0.001.

Table 4. Distribution by acceptance level in PAR percentage improvement and ABO-OGS score

ABO-OGS	PAR		P-value
	≥30% reduction (PAR+)	<30% reduction (PAR-)	
≤27 (OGS+)	19.3% (n=5, PAR+/OGS+)	0% (n=0, OGS+)	0.8077
>27 (OGS-)	76.9% (n=20, PAR+)	3.8% (n=1, PAR-/OGS-)	

PAR: peer assessment rating, ABO-OGS: American Board of Orthodontics objective grading system, PAR+: PAR-only qualified group, PAR-: PAR disqualified group, OGS+: ABO-OGS-only qualified group, OGS-: ABO-OGS disqualified group, PAR+/OGS+: both indices qualified group, PAR-/OGS-: both indices disqualified group.

A Fisher's exact test was performed.

improved close to ideal occlusion at T1. When point reductions were more than 22 points, the case was classified as 'greatly improved'. Four cases received more than 27 points on T1-ABO-OGS, despite a reduction of more than 22 PAR-points (Fig. 2). Therefore, although these four cases were 'greatly improved' and classified as high-quality cases by PAR score, these cases did not belong to the ABO-OGS-qualified cases (Fig. 2).

A Spearman's rank correlation test was fulfilled for the PAR scores, ABO-OGS scores, and PAR-percentage reduction. The T1-PAR score was negatively and middle-highly related with the amount of PAR-percentage reduction (-0.79, P<0.001) (Table 3). This finding suggested that Class I malocclusion was more effectively treated by orthodontic treatment when the value of the T0-PAR index was larger. However, there was no significant

correlation between PAR-percentage reduction and T1-ABO-OGS (Table 3).

Only five cases that showed an improvement in PAR score of more than 30% and fewer than 27 ABO-OGS points, which means the 'PAR+/OGS+' group (19.3%). The number of cases that showed an improvement in PAR score of more than 30% and more than 27 ABO-OGS points, which means the 'PAR+' group, was 20 (76.9%) (Table 4). However, there was no case that showed an improvement in PAR score of less than 30% and fewer than 27 ABO-OGS points, which means the 'OGS+' group (0%). And there was only one case that showed an improvement in PAR score of less than 30% and more than 27 ABO-OGS points, which means the 'PAR-/OGS-' group (3.8%) (Table 4). Since Fisher's exact test showed that the table probability was 0.8077 (Table 4), the distribution of acceptance

Table 5. Sensitivity/specificity of PAR index and ABO-OGS

	Evaluation tools (2×2 table)		Terminology	PAR	OGS
	OGS+	OGS-			
PAR+	5 (TP)	20 (FP)	Sensitivity	5/(5+0)=1.0000	5/(20+5)=0.2000
PAR-	0 (FN)	1 (TN)	Specificity	1/(20+1)=0.0476	1/(1+0)=1.0000
			Diagnostic accuracy	(5+1)/26=0.2307	(5+1)/26=0.2307

PAR: peer assessment rating, ABO-OGS: American Board of Orthodontics objective grading system, OGS+: ABO-OGS only qualified group, OGS-: ABO-OGS disqualified group, PAR+: PAR only qualified group, PAR-: PAR disqualified group, TP: true positive, FP: false positive, FN: false negative, TN: true negative.

Table 6. Comparison of values of PAR index between 'PAR+' group and 'PAR+/OGS+' group

Measurement	PAR+ group (n=20)		PAR+/OGS+ group (n=5)		Comparison between two groups, P-value	
	T0-PAR	T1-PAR	T0-PAR	T1-PAR	T0-PAR	T1-PAR
Alignment						
Upper labial	8.7±3.4	2.0±1.5	5.0±2.0	0.8±1.4	0.0099**	0.0288*
Lower labial	8.6±3.0	2.3±1.5	7.8±2.8	1.6±0.9	0.3788	0.1733
Lateral (right and left) occlusion	1.6±1.7	1.3±0.8	0.2±0.4	1.4±1.1	0.0091**	0.4425
Overjet	2.7±1.3	0.4±0.6	1.6±1.5	0.2±1.0	0.0815	0.3323
Overbite	1.2±1.1	0.7±0.7	0.6±0.5	0.6±0.9	0.1409	0.3967
Centerline	0.4±0.8	0.2±0.5	0.4±0.5	0	0.3652	0.1997
Total PAR score	23.1±5.8	6.7±2.7	15.6±5.0	4.6±1.5	0.0123*	0.0448*
PAR-percent reduction	70.0±12.0		67.0±19.0		0.4325	
PAR-point reduction	16.5±6.2		11.0±5.9		0.0412*	
Treatment duration (mo)	24.1±4.8		25.0±10.3		0.4057	

PAR: peer assessment rating, OGS: objective grading system, PAR+: PAR-only qualified group, PAR+/OGS+: both indices qualified group, T0: pretreatment, T1: posttreatment.

Values are presented as mean±standard deviation.

A Wilcoxon rank-sum test was performed.

*P<0.05, **P<0.01.

groups by cut-off level from the ABO-OGS and PAR index was independent of each other. This means that ABO scoring doesn't affect the PAR index score, and vice versa.

Furthermore, if we assume that the PAR index would be the gold standard, the sensitivity of ABO-OGS is 0.2000 and the specificity of the ABO-OGS is 1.0000 (Table 5). The false positive rate is 0.0000 and diagnostic accuracy, in which the sum of true positives and true negatives is divided by the total, is 0.2307 (Table 5). If we assume that the ABO-OGS would be the gold standard, the sensitivity of PAR is 1.0000 and the specificity of the PAR is 0.0476. The false positive rate is 0.9524 and diagnostic accuracy

is 0.2307, too (Table 5). This means that acceptance by the PAR index is more lenient than that of ABO in evaluating occlusion. Diagnostic accuracy indicates a low level of concurrence between the two evaluation indices.

According to these data, 96.2% of Angle Class I malocclusion patients (n=25; 'PAR+' group and 'PAR+/OGS+' group) showed a reduction in PAR scores of more than 30%. T0-PAR and T1-PAR scores and PAR-point reduction for the 'PAR+' group were significantly higher than those of the 'PAR+/OGS+' groups (23.1 vs. 15.6; 6.7 vs. 4.6; and 16.5 vs. 11.0; all P<0.05) (Table 6). Treatment duration and PAR-percentage reduction between

Table 7. Comparison of ABO-OGS between 'PAR+' group and 'PAR+/OGS+' group

Measurement	PAR+ group (n=20)	PAR+/OGS+ group (n=5)	P-value
Alignment/rotations	5.7±2.3	4.6±1.5	0.1149
Marginal ridge	5.1±2.7	3.0±2.2	0.0750
Buccolingual inclination	8.1±4.4	4.6±3.6	0.0658
Overjet	6.1±2.9	3.0±4.0	0.0279*
Occlusal contacts	8.2±5.5	4.2±3.6	0.0710
Occlusal relationships	4.9±3.8	4.6±2.3	0.4182
Interproximal contacts	0.1±0.3	0	0.2578
Total ABO-OGS score	38.2±8.7	24.0±4.0	0.0004***

ABO-OGS: American Board of Orthodontics objective grading system, PAR: peer assessment rating, PAR+: PAR-only qualified group, PAR+/OGS+: both indices qualified group.

Values are presented as mean±standard deviation.

A Wilcoxon rank-sum test was performed.

*P<0.05, ***P<0.001.

the two groups were not statistically different (24.1 months vs. 25.0 months, P=0.4057; 70.0% vs. 67.0%, P=0.4325) (Table 6). The T1-ABO-OGS score for the 'PAR+' group was significantly higher than that of the PAR+/OGS+ group (38.2 vs. 24.0, P<0.01) (Table 7).

Discussion

It became apparent that PAR and ABO-OGS are different approaches to outcome assessment. The assessment of this sample appears to have a reduction in PAR index of approximately 70%, which means "great improvement in occlusion", but the level of average end occlusion by OGS was lower when compared to the ABO's test standard. As Deguchi et al.³⁾ stated that there was no statistically significant correlation between posttreatment PAR scores and ABO-OGS scores for the same cases, this study showed that there was no statistical correlation between T0-PAR and T1-ABO-OGS, between T1-PAR and T1-ABO-OGS, or even between PAR-percent reduction and T1-ABO-OGS.

1. Understanding Apples and Oranges

Both indices were developed by orthodontic expert groups in their countries so as to reflect a professional consensus or standard^{5,11,13)}. In this

study, criteria-weighting for the PAR index by professional groups, which needs a consensus process among Korean orthodontic society members in the future, was not applied. There are two differences in applying these indices for orthodontic treatment outcomes: considering the initial malocclusion status and the measurement unit.

Basically, PAR was designed to quantify the treatment effect, where the initial malocclusion severity is considered to reimburse the treatment fee based on a clinician's input. On the other hand, to become an ABO diplomate, passing a clinical examination is the last gateway. Candidates should submit the best cases from their practices to be qualified by the ABO. The initial malocclusion status of submitted cases ranges from simple to complicated based on the requirements; however, all the cases need to pass the same cutoff level, which was 27 or fewer points in the OGS cast-radiographic evaluation.

Additionally, ABO-OGS measures individual teeth and adds points based on the deviation distance (mm), while PAR analyzes segment units, such as the upper anterior segment and left lateral segment, and adds points based on definitions, for example one point for a "one-quarter to one-half lower incisor width" deviation of midline. The

inter-arch coordination evaluation criteria from each tooth, such as the “overjet” and the “occlusal relationship” are more meticulous in the ABO-OGS. This seemed to be the reason why there was no significant correlation between T1PAR and T1-ABO-OGS (Table 3). In addition, although the samples in the present study consisted of Class I malocclusion cases that have relatively homogeneous characteristics, the ICC values for inter-operator reliability showed a difference between T1 PAR score and OGS (0.7 vs. 0.9). The PAR index measured the amount of malocclusion by segment. Compared to T0 casts, T1 casts needed more subtle approaches, like individual tooth measurement in OGS, to assess the degree of fine finishing in treatment outcome.

The ABO-OGS might be a better index for assessing the quality of finished cases, because the OGS evaluates finished casts more precisely in all three planes (first, second, and third order), while the PAR index might be superior for evaluating therapeutic improvement³⁾. In addition, even though it is important to quantify change and the outcome of a clinical intervention to determine how effective the intervention process was¹⁾, it is not appropriate for OGS to be used for assessing pretreatment malocclusion. Since the measurement range of the grading system for the ABO-OGS was a difference or deviation from 0.5 mm to 2 mm of each tooth for each criterion, a value of measurement exceeding more than 2mm has the same score with a 2 mm difference or deviation¹⁴⁾. Therefore, this system is not designed to assess treatment need or quantify the amount of initial malocclusion.

2. Improvement Quantity vs. Standard

We set the two cut-off points as a reduction of more than 30% in PAR score and fewer than 27 points in the ABO-OGS. According to these two cut-off points, one case was not qualified from both indices. Twenty out of 25 cases that showed

a reduction in the PAR index of more than 30% had more than 28 points in the ABO-OGS (PAR+ group) (Table 4). The four cases that were “greatly improved” by more than 22 reduction points from the T0-PAR score belonged to the PAR+ group. However, the ABO-OGS was not able to appreciate the huge change by treatment as a higher outcome.

In the present study, both the PAR+ and PAR+/OGS+ groups were qualified as “improved” by the PAR index; however, some of them were qualified by the ABO-OGS, but others were not (mean OGS scores; 38.2 vs. 24.0, $P<0.001$) (Table 7). Based on the T0-PAR scores, the PAR+ group might be a worse malocclusion group than the PAR+/OGS+ group (mean T0-PAR; 23.1 vs. 15.6, $P<0.05$) (Table 6). The mean T0-PAR scores of both groups were significantly different (23.1 vs. 15.6, $P<0.05$) (Table 6). Among the subcomponents of PAR, upper labial segment alignment and combined lateral buccal occlusion showed significant differences (8.7 vs. 5.0; 1.6 vs. 0.2, all $P<0.05$) (Table 6).

The treatment duration between the two groups was not different (Table 6). Also, there was no statistical difference in PAR-percent reduction between the PAR+/OGS+ and PAR+ groups (Table 6). However, the mean PAR-point reduction of the PAR+ group was substantially larger than that of ‘PAR+/OGS+’ (16.5 vs. 11.0, $P<0.05$) (Table 6). Therefore, T1-PAR scores were higher in the PAR+ group than in the PAR+/OGS+ group (6.7 vs. 4.6, $P<0.05$) (Table 6). Among the criteria, only the “upper labial segment alignment” showed a statistical difference (2.0 vs. 0.8, $P<0.05$) (Table 6). The ABO-OGS scores were significantly different between the PAR+ and PAR+/OGS+ groups, too (38.2 vs. 24.0, $P<0.001$) (Table 7). When comparing the PAR+ and PAR+/OGS+ groups in the ABO-OGS, no specific components, except for overjet, explained the overall statistical difference between the two groups (6.1 vs. 3.0, $P<0.05$) (Table 7). Therefore, during the same duration, the level of treatment outcome for cases in a worse pretreatment

condition turned out to be the lower than those in a better initial condition, even though the absolute quantity of improvement was larger.

This study exhibited the differences in assessment of orthodontic treatment outcome between PAR and ABO-OGS despite the homogeneous Class I malocclusion group. Therefore, further studies are needed to develop a new system for accurate evaluation of pretreatment status and treatment outcome using combination of the advantages of both indices.

Conclusion

The null hypothesis of this study was rejected.

Since the fraction of the PAR+/OGS+ group was less than 20% and there was no significant correlation between PAR-percentage reduction and T1-ABO-OGS, development of a new index system for accurate evaluation of pretreatment status and treatment outcome is needed.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgement

We thank the postgraduate students and faculty members of St. Mary's Hospital, The Catholic University, Seoul, Korea.

References

- Huang GJ, Richmond S, Vig KW. Evidence-based Orthodontics. Chichester: Wiley-Blackwell; 2011.
- Turbill EA, Richmond S, Andrews M. A preliminary comparison of the DPB's grading of completed orthodontic cases with the PAR index. *Br J Orthod.* 1994; 21: 279-85.
- Deguchi T, Honjo T, Fukunaga T, Miyawaki S, Roberts WE, Takano-Yamamoto T. Clinical assessment of orthodontic outcomes with the peer assessment rating, discrepancy index, objective grading system, and comprehensive clinical assessment. *Am J Orthod Dentofacial Orthop.* 2005; 127: 434-43.
- Roberts CT, Richmond S. The design and analysis of reliability studies for the use of epidemiological and audit indices in orthodontics. *Br J Orthod.* 1997; 24: 139-47.
- Casko JS, Vaden JL, Kokich VG, Damone J, James RD, Cangialosi TJ, Riolo ML, Owens SE Jr, Bills ED. Objective grading system for dental casts and panoramic radiographs. American Board of Orthodontics. *Am J Orthod Dentofacial Orthop.* 1998; 114: 589-99.
- Onyeaso CO, Begole EA. 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 Dentofacial Orthop.* 2007; 131: 248-52.
- Fox NA, Chapple JR. Measuring failure of orthodontic treatment: a comparison of outcome indicators. *J Orthod.* 2004; 31: 319-22.
- Kim HH, Lee KH, Kim JC. The treatment change of PAR (peer assessment rating) index and cephalometric measurements in Class I malocclusion patients. *Korean J Orthod.* 1999; 29: 277-84.
- Cheong SB, Park YG. Evaluation of orthodontic treatment outcomes and post-treatment changes by the PAR index. *Korean J Orthod.* 2001; 31: 393-401.
- Son WS, Cha KS, Chung DH, Kim TW. Quantitative evaluation and affecting factors of post-treatment relapse tendency. *Korean J Orthod.* 2011; 41: 154-63.
- Richmond S, Shaw WC, Roberts CT, Andrews M. The PAR index (peer assessment rating): methods to determine outcome of orthodontic treatment in terms of improvement and standards. *Eur J Orthod.* 1992; 14: 180-7.

12. Case preparation [Internet]. St. Louis (MO): The American Board of Orthodontics [cited 2014 Apr 25]. Available from: <http://www.americanboardortho.com/professionals/clinicalexam/casereportpresentation/preparation/measurementinstruments.aspx>
13. Richmond S, Shaw WC, O'Brien KD, Buchanan IB, Jones R, Stephens CD, Roberts CT, Andrews M. The development of the PAR index (peer assessment rating): reliability and validity. *Eur J Orthod.* 1992; 14: 125-39.
14. Grading system for dental casts and panoramic radiographs [Internet]. St. Louis (MO): The American Board of Orthodontics [cited 2014 Apr 25]. Available from: <http://www.americanboardortho.com/professionals/downloads/Grading%20System%20Casts-Radiographs.pdf>