

Orthodontic Treatment—Improvement and Standards Using the Peer Assessment Rating Index

Chukwudi Ochi Onyeaso^a; Ellen A. BeGole^b

ABSTRACT

This article is aimed at assessing the orthodontic treatment outcome in an accredited graduate orthodontic clinic in the United States, using the original Peer Assessment Rating (PAR) Index. A sample of 100 pre- and posttreatment study models was randomly selected from the clinic model storage. One investigator assessed all the components of the PAR Index. Descriptive statistics and Scheffe test were used to analyze the data. The mean weighted PAR point reduction was 22.11, and mean percentage reduction was 86.20. “Greatly improved” conditions were seen in 50% of patients while 47% accounted for the “improved category.” Only 3% belonged to the “worse or no different” group. Statistically significant differences were observed between the three improvement groups for weighted pre- and posttreatment PAR scores, reduction, and percentage reduction. The results indicated an excellent improvement of the patients, with patients being treated to a high standard and a great proportion of patients with a clear need for treatment. (*Angle Orthod* 2006;76:260–264.)

KEY WORDS: Orthodontic; Treatment; Outcome

INTRODUCTION

In orthodontics, it is important to objectively assess whether a worthwhile improvement has been achieved in terms of overall alignment and occlusion for an individual patient and the greater proportion of a practitioner’s caseload.¹ Several indices have been developed specifically to assess the success of orthodontic treatment by comparing pre- and posttreatment records to register the outcome of orthodontic care.^{2–4} Summers occlusal index,⁵ a treatment index, has also been used to assess the outcome of treatment.^{6,7}

In 1987, the Peer Assessment Rating (PAR) Index was developed over a series of six meetings with a group of 10 experienced orthodontists (British Orthodontic Standards Working Party).⁸ The PAR Index provides a single summary score for all the occlusal

anomalies which may be found in a malocclusion. The score provides an estimate of how much a case deviates from normal alignment and occlusion. The difference in scores between the pre- and posttreatment patients reflects the degree of improvement and, therefore, the success of the treatment.

The PAR Index has been used to evaluate treatment standard among general practitioners and orthodontic specialists.^{9,10} The PAR Index has its limitations, principally because of the high weight assigned to overjet, and difficulties still arise from the application of one weighting system to all malocclusions because occlusal features vary in importance in different classes of malocclusion.¹¹ However, it remains a valid and reliable index in the assessment of orthodontic treatment outcome. Treatment outcome can be assessed using the PAR Index in two ways: using the numerical reduction in the weighted PAR score or using the percentage reduction in the weighted PAR score.

Unlike in Europe, there is paucity of information on the outcome of orthodontic treatment in the United States assessed using the PAR Index. The report of Holman et al¹² on extraction vs nonextraction orthodontic treatment using the PAR Index applied a US weighting system which was developed specifically for white United States patients.

Therefore, this article is aimed at assessing the orthodontic treatment outcome in an accredited graduate

^a Faculty Lecturer, Department of Preventive Dentistry, University of Ibadan, Ibadan, Nigeria.

^b Professor of Biostatistics in Orthodontics, Department of Orthodontics, College of Dentistry, University of Illinois at Chicago, Chicago, Ill.

Corresponding author: Chukwudi Ochi Onyeaso BDS, FWACS, Department of Preventive Dentistry, University of Ibadan, Ibadan, Oyo State 234, Nigeria (e-mail: coonyeaso@yahoo.com)

Accepted: April 2005. Submitted: February 2005.

© 2006 by The EH Angle Education and Research Foundation, Inc.

orthodontic clinic in the United States, using the original PAR Index.

MATERIALS AND METHODS

A sample of 100 patients was randomly selected from the model store of an accredited graduate orthodontic clinic in the United States. No consideration was given to race or sex in the selection of patients for the sample. The duration of treatment was calculated by subtracting the age at the onset of treatment as recorded on the pretreatment models from the age at the end of treatment as indicated on the posttreatment models.

The pretreatment and posttreatment PAR scores were recorded according to Richmond et al¹ All the components of the PAR Index were assessed. Treatment outcome can be assessed using the PAR Index in two ways: using the numerical reduction in the weighted PAR score or using the percentage reduction in the weighted PAR score.

Several conventions have been derived from the panel of 74 examiners (members of the British Orthodontic Society) to categorize improvement:

1. "Worse or no different," patients who show less than a 30% reduction in weighted PAR score.
2. "Improved," patients showing greater than or equal to 30% reduction in weighted PAR score.
3. "Greatly improved," generally a reduction of 22 weighted PAR points (greatly improved reflects the severity of the initial malocclusion).

In assessing the standard of orthodontic treatment, a high standard of care is achieved when the proportion of patients treated by an individual practitioner falling into the worse or no different category is negligible and the mean percentage reduction in weighted PAR score is high (eg, greater than 70%). If a high proportion of patients also fall into the category "greatly improved" (eg, greater than 40%), then this suggests that the practitioner is treating a substantial proportion of patients with significant malocclusion to high standard of care.

Intraexaminer reliability

The errors of the method for the recording of the PAR Index were evaluated from double recordings of 20 randomly selected patients from the original sample. The pretreatment and posttreatment study casts were evaluated a second time; 40 double recordings were performed. The random or accidental error for the weighted and unweighted PAR Index scores was evaluated with the formula:

$$S_i = \sqrt{\sum \frac{d_i}{2n}}$$

TABLE 1. Mean Unweighted and Weighted Pretreatment PAR Scores^a

PAR Components	Mean Pretreatment PAR Scores (Total Sample)	
	Unweighted	Weighted
Upper right segment	1.08	0.00
Upper anterior	5.43	5.43
Upper left segment	1.21	0.00
Lower right segment	1.20	0.14
Lower anterior segment	3.84	3.63
Lower left segment	1.03	1.57
Right buccal occlusion	0.88	1.30
Overjet	1.45	1.31
Overbite	1.18	2.38
Centerline	0.46	1.84
Left buccal occlusion	0.78	0.83
Total	18.55 ± 9.34	23.83 ± 11.55

^a PAR indicates Peer Assessment Rating.

Where *d_i* is the difference between the double determinations and *n* is the number of double determination.¹³ For further measure of the reliability of the measurements, the intraclass correlation coefficient was calculated using the Spearman rank-order correlation coefficient.¹⁴

Statistical analysis

Besides the use of descriptive statistics in the analysis of the whole data, Scheffe test was applied to ascertain statistically significant differences in the three different categories of improvement. All statistical analyses were carried out using the Statistical Package for Social Sciences for Windows software (SPSS Inc, Chicago, Ill).

RESULTS

From the 40 double recordings, the *S_i* value was 0.30 for unweighted total PAR Index and 0.67 for the weighted total PAR Index. There was no systematic bias between the two sets of PAR scores for the double recordings.

Intraexaminer correlation coefficient results revealed that pretreatment (a) vs pretreatment (b) showed excellent correlation (*r* = .99; *P* < .0001) and posttreatment (a) vs posttreatment (b) indicated very good correlation (*r* = .83; *P* < .0001).

Pretreatment and posttreatment unweighted and weighted PAR scores are shown in Tables 1 and 2. The upper anterior segment had the highest mean for both unweighted and weighted pretreatment PAR scores. In the posttreatment PAR score, the right buccal occlusion had the highest mean for unweighted PAR score, whereas overjet had the highest for weighted.

TABLE 2. Mean Unweighted and Weighted Posttreatment PAR Scores^a

PAR Components	Mean Posttreatment PAR Scores (Total Sample)	
	Unweighted	Weighted
Upper right segment	0.00	0.01
Upper anterior segment	0.14	0.13
Upper left segment	0.01	0.00
Lower right segment	0.07	0.01
Lower anterior segment	0.14	0.13
Lower left segment	0.06	0.02
Right buccal occlusion	0.16	0.16
Overjet	0.10	0.60
Overbite	0.08	0.19
Centerline	0.09	0.36
Left buccal occlusion	0.04	0.09

^a PAR indicates Peer Assessment Rating.

TABLE 3. PAR Index Descriptives for the Whole Sample^a

Data for Whole Sample	Mean	SD
Start PAR score		
Weighted	23.83	11.55
Unweighted	18.55	9.34
Finish PAR score		
Unweighted	0.96	1.80
Weighted	1.72	3.79
PAR reduction		
Weighted	22.11	11.72
Unweighted	17.59	9.59
Percentage Reduction		
Weighted	86.20	27.44
Unweighted	89.01	24.31

^a PAR indicates Peer Assessment Rating.

Table 3 shows the PAR Index descriptors for the whole sample. The mean start weighted PAR score was 23.83 ± 11.55 (SD) and the finish PAR score was 1.72 ± 3.79 (SD). The weighted PAR point reduction was 22.11 ± 11.72 (SD), whereas the mean percentage reduction was 86.20 ± 27.44 (SD).

The numbers in the right-hand column in Table 4 indicate the number corrected out of the number of patients in whom improvement was indicated or possible, as shown on the first column on the left-hand side. In parentheses are the corresponding percentages. The upper anterior segment malocclusions had the highest percentage improvement (94.7%) followed by overjet (92.8%), as shown in Table 4.

Table 5 shows the relationship of pre- and posttreatment PAR scores, reduction, and percentage reduction in weighted PAR scores for the three categories of improvement. The greatly improved category had 50 patients (50%), and 47 (47%) were in the improved category, while only 3 (3%) were found to belong to the worse or no different category (Figure 1).

TABLE 4. Improvement of Individual PAR Components^a

PAR Component	Cases in Which Improvement was Possible	Number (%) Improved
Upper right segment	—	—
Upper anterior segment	94	89 (94.7)
Upper left segment	—	—
Lower right segment	2	1 (50)
Lower anterior segment	—	—
Lower left segment	2	1 (50)
Right buccal occlusion	35	27 (77.1)
Overjet	69	64 (92.8)
Overbite	55	48 (87.3)
Centerline	42	33 (78.6)
Left buccal occlusion	32	28 (87.5)

^a PAR indicates Peer Assessment Rating.

TABLE 5. The Relationship of Pre- and Posttreatment, Reduction, and Percentage Reduction in Weighted PAR Scores for the Three Classifications of Improvement^a

	Pretreatment	Posttreatment	Reduction	% Reduction
Worse or no difference (3 cases)				
Mean	10.3	10.0	0.33	1.13
Minimum	0	0	0	0
Maximum	29	28	1	3.4
Improved (47 cases)				
Mean	32.66	1.28	31.38	96.40
Minimum	23.00	0	22.0	71.43
Maximum	49.00	12.00	49.0	100.0
Greatly improved (50 cases)				
Mean	33.26	1.36	31.90	90.61
Minimum	23	0	22	0
Maximum	51	12.00	51	100.0

^a Using Scheffe test with significance level at .05, statistically significant differences were found between all groups for all variables.

A multiple comparison test was used, which indicated statistically significant differences (Scheffe $< .05$) between the three groups for all the variables recorded (pretreatment, posttreatment, reduction, and percentage reduction in weighted PAR scores).

Patients having less than 22 pretreatment weighted PAR scores were 46 in number (46%). The average duration of treatment was 26 months. Finish weighted PAR scores between 0 and 5 points accounted for 88%.

DISCUSSION

For a practitioner to demonstrate consistently high standards, the proportion of the caseload in the worse or no different category should be negligible, and the mean percentage reduction should be as high as possible.

A high standard of treatment may be judged ac-

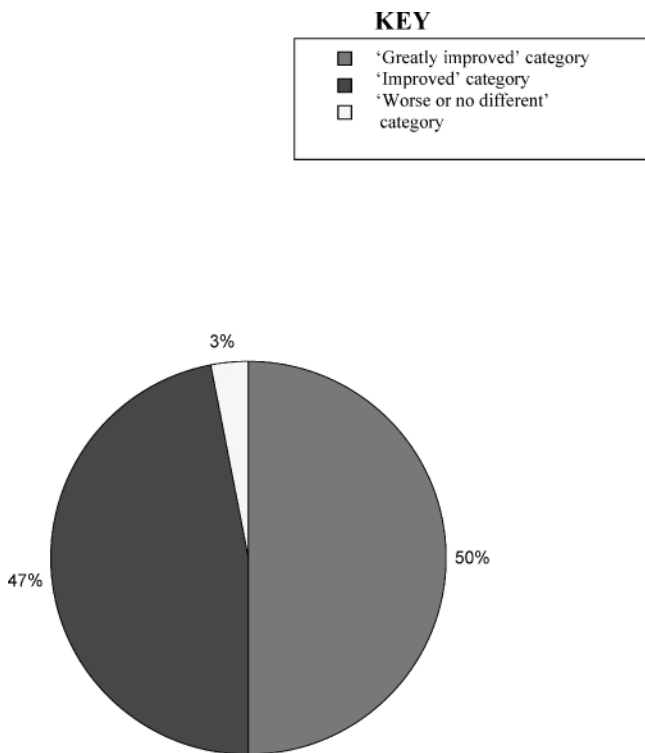


FIGURE 1. Pie chart illustration of the different categories of treatment outcome in the sample.

According to the mean percentage reduction in weighted PAR score for an individual practitioner's caseload, eg, greater than 70%.¹ The present US study has revealed a mean percentage reduction in weighted PAR score of over 86%. This parameter was reported to be about 47% in the General Dental Service Orthodontics in England and Wales.¹⁵ In a pilot study to assess the standard of orthodontic treatment in Norway, the mean percentage reduction in weighted PAR score for all the patients was 78%, with four of the patients in the worse or no different category. In a related previous study in the United States,¹² the mean percentage reduction in weighted PAR score was 79.41% for the extraction group and 77.63% for the nonextraction group.

For a practitioner to produce high standards and treat those patients who have perhaps a greater need for treatment, not only must the mean percentage reduction for the caseload be high (eg, greater than 70%), but the percentage of patients having been greatly improved should also be high (eg, greater than 40%).¹ In the present study, the greatly improved category was 50%. Firestone et al¹⁶ reported, for the greatly improved category, an average of 27% for patients treated by postgraduate and undergraduate students in Bern, Switzerland. In the same Swiss report,¹⁶ it was 28% in 1983 and 40% in 1993 for the postgraduate students. The same Swiss report indicated 12%

of patients in worse or no different group, as against 3% in the present US study. Kerr et al¹⁷ reported 89.3% of patients as either improved or greatly improved with 16 (10.7%) classified as worse or no different, in assessing the effectiveness of removable orthodontic appliances.

When considering the change in PAR score for a group of subjects, it must be borne in mind that the amount of improvement possible under the nomogram system is influenced by the initial PAR score.¹⁷ If this is less than 22, it is impossible for the patients to become greatly improved. In the UK (Glasgow) study by Kerr et al,¹⁷ this was found to apply to one third (49) of the patients in their study, as against 46% in the present US study. This high percentage of patients with initial PAR score of less than 22 in the present study contributed in lowering the percentage of patients belonging to the greatly improved category.

Although this study revealed many patients with posttreatment PAR score indicating ideal occlusions, most of them could not be classified as greatly improved. According to Kerr et al,¹⁷ if all the patients in a given study had initial PAR scores of less than 22, none could be greatly improved, and one might, therefore, infer that the treatment regimen used was in some way defective. Consequently, it would seem reasonable to identify the proportion of patients in this category to allow a more objective assessment of the outcome.

Longer dual arch fixed appliances were reported as having the tendency to reduce PAR scores to lower levels.¹⁵ The present study assessed patients treated with dual arch fixed appliances (preadjusted edgewise technique) because it is the practice at the center. The average duration of treatment for the present study sample was 26 months. The random sample of 1010 patients treated with removable or fixed appliances (or both) in the General Dental Services of England and Wales showed that one out of every five patients would be categorized as worse or no different, with an overall mean percentage reduction in weighted PAR score of 55%.¹ When the standard of treatment was assessed in relation to the appliance used, the use of upper and lower fixed appliances produced the best standard of treatment (71.4% reduction in weighted PAR score). In addition, the patient was less likely to be worse after treatment than other patients treated with removable or functional appliances. Single arch removable and fixed appliances showed a mean percentage reduction of around 50%, with a substantial proportion of patients not benefiting from orthodontic treatment.

The present study showed that the mean pretreatment weighted PAR score was 23.83 and the post-treatment weighted PAR score was 1.72. Firestone et

al¹⁶ reported 25 in 1983 and 24.8 in 1993 for pretreatment PAR scores, and 8.3 in 1983 and 5.3 in 1993, for posttreatment PAR scores in Switzerland. Turbill et al¹⁵ reported pretreatment weighted PAR scores of 26.94, 26.74, and 24.74 for patients treated with removable appliances, respectively. Their corresponding posttreatment weighted PAR scores were 12.79, 15.19, and 11.40, respectively. The study of Holman et al¹² showed pretreatment PAR score of 30.01 for extraction group as against 25.21 for the nonextraction group in a related study that used the American weighting system.

CONCLUSIONS

- Improved and greatly improved categories of improvement constituted about 97% of the patients, with only 3% in the worse or no different group.
- The mean weighted PAR reduction was 22.11, whereas the mean weighted percentage reduction was 86.20, indicating excellent improvement of the patients according to the PAR Index.
- This accredited graduate orthodontic clinic in the United States is treating a great proportion of patients with a clear need for treatment to a high standard. This could be a reflection of the standard of treatment in the United States.

REFERENCES

1. 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;14:180–187.
2. Eismann D. A method of evaluating efficiency of orthodontic treatment. *Trans Eur Orthod Soc.* 1974;223–232.
3. Eismann D. Reliable assessment of morphological changes results from orthodontic treatment. *Eur J Orthod.* 1980;2:19–25.
4. Gottlieb EL. Grading your orthodontic treatment results. *J Clin Orthod.* 1975;9:143–154.
5. Summers CJ. The occlusal index: a system for identifying and scoring occlusal disorders. *Am J Orthod.* 1971;59:552–566.
6. Pickering EA, Vig P. The occlusal index used to assess orthodontic treatment. *Br J Orthod.* 1975;2:47–51.
7. Elderton RJ, Clark JD. Orthodontic treatment in general dental service assessed by the occlusal index. *Br J Orthod.* 1983;10:178–186.
8. Richmond S, Shaw WC, O'Brien KD, Buchanan IB, Stephens CD, Roberts CT, Andrews M. The development of the PAR Index (Peer Assessment Rating): reliability and validity. *Eur J Orthod.* 1992;14:125–139.
9. Richmond S, Shaw WC, Stephens CD, Webb WG, Roberts CT, Andrews M. Orthodontics in the General Dental Services of England and Wales: a critical assessment of standards. *Br Dent J.* 1993;174:315–329.
10. O'Brien KD, Shaw WC, Roberts CT. The use of occlusal indices in assessing the provision of orthodontic treatment by the hospital orthodontic Service of England and Wales. *Br J Orthod.* 1993;20:25–35.
11. Hamdan AM, Rock WP. An appraisal of the Peer Assessment Rating (PAR) Index and a suggested new weighting system. *Eur J Orthod.* 1999;21:181–192.
12. Holman JK, Hans MG, Nelson S, Powers MP. An assessment of extraction versus non-extraction orthodontic treatment using the Peer Assessment Rating (PAR) Index. *Angle Orthod.* 1998;68:527–534.
13. 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;24:139–147.
14. Fisher RA. *Statistical Methods for Research Workers.* New York, NY: Hafner Publishing Co; 1973, p 10–15.
15. Turbill EA, Richmond S, Wright JC. A closer look at General Dental Service orthodontics in England and Wales 1: factors influencing effectiveness. *Br Dent J.* 1999;187:211–216.
16. Firestone AR, Hasler RU, Ingervall B. Treatment results in dental school orthodontic patients in 1983 and 1993. *Angle Orthod.* 1999;69:19–26.
17. Kerr WJS, Buchanan IB, McColl JH. Use of the PAR Index in assessing the effectiveness of removable orthodontic appliances. *Br J Orthod.* 1993;20:351–357.