Clin Oral Invest (2001) 5:57–62 © Springer-Verlag 2001

ORIGINAL ARTICLE

G. Willems · R. Heidbüchel · A. Verdonck · C. Carels

Treatment and standard evaluation using the Peer Assessment Rating index

Received: 31 July 2000 / Accepted: 26 September 2000

Abstract The aim of this retrospective study was to determine the outcome of orthodontic treatment carried out on patients by postgraduate students at the Katholieke Universiteit Leuven, Belgium. The treatment outcome of 292 'final examination' patients and of 287 'control' patients was compared by means of the Peer Assessment Rating (PAR) index. The sample consisted of dental casts representing a wide range of malocclusions at the start of treatment and post-treatment. All patients received non-surgical treatment between 1987 and 1996 by one of 18 different postgraduate students. The data were analysed with a variant of the analysis of covariance. A significantly higher (P<0.001) treatment standard was found for final examination patients compared to the control, indicated by the mean percentage PAR score reduction of 79.1% and 70.7 %, respectively. When the results are expressed in terms of treatment outcome, 44.5% of the examination patients and 44.0% of the controls were allocated to the 'Greatly improved' group, while 3.1 % of the patients examined and 7.3% of the control patients were classified as 'Worse or no different'.

Keywords Orthodontics · Treatment outcome · PAR index

Introduction

Until the early 1990s no index had been developed to record treatment success. Until that time the Occlusal Index, initially designed for other tasks, was used to evaluate the success of treatment [4, 12]. It was only with the recent development of the Index of Orthodontic Treatment Need and the Peer Assessment Rating Index (PAR) [2, 14] that there has been increased interest in the use of indices as tools for assessing treatment outcome.

In recent years several investigators all over the world have carried out similar studies to determine the effectiveness of orthodontic treatment by use of the PAR index [5, 6, 10, 15–18].

In a study that was performed in Norway, where nearly all orthodontic treatment was undertaken by specialists, the standard of orthodontic treatment outcome was better than in the United Kingdom. The mean percentage reduction in PAR score was 78% and only 4% of cases were categorized as 'Worse or no different' [13].

Finally, Al Yami [1] evaluated the overall quality of orthodontic treatment at the University of Nijmegen in the Netherlands. He evaluated 1870 patients using the PAR index. The results showed that the mean weighted PAR score was 28.0±10.0 and 8.0±6.1, for the pretreatment and post-treatment dental casts, respectively. The mean percentage improvement was 69%, 43% of the sample was 'Greatly improved', 49% was 'Improved' and 8% was 'Worse or no different'.

The aim of the present study was to evaluate the orthodontic outcome of patients treated by postgraduate students during their orthodontic specialist training at the Katholieke Universiteit Leuven. Because postgraduate students have to present a number of treated patients at the end of their orthodontic specialist training, there could be a difference in treatment standard between these examination patients and those patients not belonging to the final examination group.

Materials and Methods

The Peer Assessment Rating Index

Evaluation of the treatment standard was undertaken by applying the weighted PAR index. Details of the PAR index have been published elsewhere [14] but in summary the index applies a score to the following occlusal features: upper anterior segment alignment (1), lower

G. Willems (✉) · R. Heidbüchel · A. Verdonck · C. Carels
Department of Orthodontics, School of Dentistry,
Oral Pathology and Maxillofacial Surgery,
Faculty of Medicine, Katholieke Universiteit Leuven,
Kapucijnenvoer 7, 3000 Leuven; Belgium
e-mail: guy.willems@med.kuleuven.ac.be
Tel.: +32 16 332439, Fax: +32 16 332435

anterior segment alignment (1), left and right buccal occlusions (1), overbite and open bite (2), overjet and anterior crossbite (6), centreline (4). The number in the parentheses represents the weighting given to each feature of the malocclusion. The weighted individual scores for the various components are finally calculated to obtain a total score (PAR score) that represents the degree a case deviates from normal alignment and occlusion. Malocclusion which deviates markedly from normal would attract a high weighted PAR score (rarely above 50) and an almost ideal occlusion will score less then 5. The effectiveness of treatment can be assessed by comparing the pretreatment and post-treatment PAR scores. The greater the reduction in PAR score, the greater the improvement achieved. A malocclusion is defined as 'Greatly Improved' when the post-treatment PAR score is at least 22 points lower than the pretreatment PAR score. The malocclusion is defined as 'Improved' when the post-treatment PAR score is 30% lower than the pretreatment PAR score, and cases showing a change in the PAR of less than 30% are defined as 'not improved'. A high standard of treatment is achieved when the proportion of cases falling into the 'Worse or No different' category of an individual's case load is neglegible and the mean percentage reduction in weighted PAR score is high (e.g. greater than 70%). It is unrealistic to expect all malocclusions to be treated to produce an ideal occlusion. The outcome of treatment is often dependent on many factors, e.g. complexity of the case, treatment methods, expertise of the practitioner and patient compliance [10, 14, 15].

Selection of the patients' records

All the available patients' dental plaster casts in the archives of the Department of Orthodontics at the Katholieke Universiteit Leuven were examined. To be included in the study, patients had to be treated orthodontically (without surgery) by one of 18 different postgraduate students and both the pretreatment and post-treatment casts had to be in good condition. The sample of 579 cases represented a full range of malocclusions including increased overbites and overjets, partially erupted teeth, crowding, spacing, open-bites and crossbites.

After examining the dental casts, the samples were categorized in two different groups. Group 1, the final examination cases, consisted of 292 patients (117 males and 175 females) with a mean age of 14.0±6.8 years and 17.0±6.7 years at the pretreatment and post-treatment stages, respectively. Group 2, the control cases (patients not belonging to the final examination group), consisted of 287 patients (123 males and 164 females) with a mean age of 13.0±4.7 years and 16.0±4.1 years at the pretreatment and post-treatment stages, respectively. Due to administrative methods of patient allocation, randomization of the cases for both groups in respect to severity was certified.

The patients were finally categorized according to the year in which they finished their active orthodontic treatment between 1987 and 1996.

Data collected

All the dental casts were alphabetically examined by one investigator who was calibrated in the use of the PAR index. This made it possible to compare the results in this study with those of other workers similarly calibrated in the use of these indices.

In addition, the following details were collected from each patient's record card: (1) gender, (2) date of birth, (3) appliances used, (4) postgraduate student, (5) examination patient or not, (6) Angle classification, (7) date of the pretreatment and post-treatment records and (8) date of start of active treatment.

Data analysis

Twenty cases were scored twice with an interval of 3 weeks to allow for assessment of reliability by means of a one-sample *t*-test applied to identify any systemic difference in scoring between the two media. Random errors are important in that measurements with high random errors in relation to their total variability will be of little value.

Mean and standard deviation of the PAR scores were calculated at the pretreatment stage and at the end of active treatment. The percentage reduction in the weighted PAR was calculated to assess the improvement or treatment standards. In addition to the basic descriptive analysis of the data, the Cochran-Mantel-Haenszel test or Fisher's Exact test was carried out to compare the two groups for individual pretreatment and post-treatment PAR variables. Analyses of variance and the Wilcoxon test were applied to compare the quality of treatment and investigate and separate the effects of the various factors which might influence the level of treatment outcome as measured by the PAR index. A Chi-squared analysis was used to relate the appliance type to pretreatment and post treatment PAR scores, PAR score change and percentage PAR score change (SAS Statistical Software Package, SAS Institute, Cary, NC, USA).

Results

Error analysis

The magnitude of the intra-observer duplicate (random) error was low, 1.47, indicating a high level of reproducibility of the PAR index.

Treatment duration

The mean treatment duration from placement of the appliance to removal was 28.8±11.8 months for the exami-

Table 1 Change in Peer Assessment Rating (PAR) Index for the different groups

	Group 1 (*n*=292)				Group 2 (*n*=287)			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Pretreatment PAR	25.3	9.1	4.0	53.0	27.1	9.7	5.0	56.0
Post-treatment PAR	4.9	4.5	0.0	27.0	7.2	5.6	0.0	32.0
Numerical PAR change	20.5	9.5	1.0	53.0	19.8	10.3	0.0	54.0
Percentage PAR change	79.1	18.7	12.0	100.0	70.7	23.1	0.0	100.0

Table 2 Percentage of perfect post-treatment scores of the individual components of the PAR Index with standard errors (*SE*)

Component	Percentage scorers			SE		
	Group 1	Group 2	Total	Group 1	Group 2	Total
Maxillary anterior teeth	85.6	87.8	86.7	0.2	0.2	0.1
Mandibular anterior teeth	81.9	81.2	81.5	0.2	0.1	0.2
Right and left occlusion	5.1	3.8	4.5	0.1	0.1	0.1
Overjet/anterior crossbite	79.8	53.7	65.3	0.2	0.3	0.2
Overbite/open bite	79.8	77.4	78.6	0.2	0.2	0.2
Centreline	95.2	91.3	93.3	0.1	0.2	0.1

nation patients (with a maximum of 48.2 months and a minimum of 8.0 months) and 35.9±14.9 months for the control patients (with a maximum of 63.8 months and a minimum of 5.9 months). Fixed appliance treatment (28.9 months) was longer compared to removable appliance treatment (22.1 months). No significant differences in treatment time were found between males and females (28.9 and 26.5 respectively).

The changes in occlusion and alignment due to treatment as assessed by the PAR index

The results by comparing the two different groups (examination versus control patients)

The range of start and finish PAR scores for the two different groups are shown in Table 1. The mean pretreatment PAR score was 25.3 for the final examination patients and 27.1 for the control patients. The latter was significantly higher ($P<0.05$) although no significant differences could be found in the two groups for the individual pretreatment PAR variables (Cochran-Mantel-Haenszel test).

The occlusal result as determined by mean post-treatment PAR index was 4.9 in the final examination group, which was significantly better than 7.2 for the control group ($P<0.05$). The final examination group had a significantly more ideal overjet ($P<0.001$), sagittal and transversal occlusion ($P<0.05$). Table 2 shows the percentage of perfect scores (score=0) of the different post-treatment PAR index components.

The mean percentage improvement of PAR score was 79.1 in the final examination group and 70.7 in the control group which was significantly different ($P<0.001$). Still 3.1% from the patients of the final examination group and 7.3% of the control group fell into the 'Worse or no different' category. Further, Group 1 counted 44.5% in the 'Greatly improved' and 52.4% in the 'Improved' category, whereas for Group 2 these numbers were 44.0% and 48.8%, respectively.

Table 3 Change in Peer Assessment Rating (PAR) for the group as a whole

	Group l + group 2 (*n*=579)		
	Mean	S.D.	Min
Pretreatment PAR score (T1)	26.2	9.4	4.0
Post-treatment PAR score (T2)	6.0	5.2	0.0
PAR score change (Tl–T2)	20.2	9.9	0.0
Percentage PAR score change	74.9	21.4	0.0

The results of the group as a whole

The mean PAR score for the whole sample was initially 26.2±9.4 and dropped to 6.0±5.2 (Table 3).

This resulted in a mean percentage improvement of 74.9%, which represents a high standard of treatment with 44.2% of cases being 'Greatly improved' and 50.6% cases being 'Improved'. In the whole group still 5.2% cases fell into the 'Worse or no different' category.

A significant correlation was found between the pretreatment and post-treatment PAR scores in both groups: the higher the pretreatment PAR score, the higher the post-treatment PAR score ($P<0.001$).

Variation in PAR with sex

Significant differences between males (n=240) and females (n=339) were found for the mean PAR at the pretreatment stage, the scores being higher in males (27.8 for males and 25.1 for females; $P<0.05$). Taking account of the previous correlation, a higher mean post-treatment PAR score was also expected in males. However, only a

Table 4 Details of appliances used for all patients

Treatment method	Number of Patients	Pretreatment PAR	Post-treatment PAR	Change	% change
Two-arch fixed	279	26.4	4.4	22.0	81.3
Single-arch fixed	94	23.0	7.5	15.5	65.1
Removable	46	21.2	8.1	13.1	60.1
Functional	160	28.9	7.0	21.9	74.6

Table 5 Peer Assessment Rating outcome by treatment method

Treatment method	Worse/no different		Improved		Greatly improved	
	n	%	*n*	%	*n*	%
Two arch fixed	5	1.8	138	49.8	134	48.4
Single arch fixed	11	11.7	60	63.8	23	24.5
Removable	7	15.2	33	71.7	6	13.0
Functional	7	4.4	81	50.6	72	45.0

small, but not significant difference was found in mean post-treatment PAR score (6.3 for males and 5.9 for females). Although there was a higher pretreatment PAR score in males, no significant difference in percentage reduction of PAR score with sex was found (74.2% for females and 75.9% for males). Males, however, were categorized more often as 'Greatly improved' than females (49.7% and 40.5%, respectively).

Variation in PAR with treatment methods

Table 4 illustrates the differences in treatment success using different treatment methods. The categorical representation of these results are shown in Table 5. To provide a more representative sample size, the group treated with a removable appliance in combination with a single arch fixed appliance was combined with the single arch fixed appliance group. All cases treated with a functional appliance were also combined. The pretreatment PAR scores for patients treated with functional appliances were higher than those treated with the other treatment methods. The use of full upper and lower fixed appliances gave the lowest mean post-treatment PAR score and the highest mean percentage PAR score reduction. Functional appliances scored also highly and single arch fixed appliances achieved a greater percentage reduction in PAR than removable appliances.

It can be seen that the appliance type had a significant effect on the post-treatment PAR score and a highly significant effect on the percentage PAR change (Table 6).

Variation in PAR with Angle classification

In the different groups of Angle classification the treatment outcome for the final examination patients was also significantly higher ($p<0.001$), but there was no significant difference in treatment outcome between the various Angle classification groups (Table 7).

Table 6 Results of Chi-square analysis on factors affecting PAR scores

	Appliance type (Chi-a wared value)	Degrees of freedom
Pretreatment PAR score	37.7	3
Post-treatment PAR score	44.2*	3
PAR change	56.9**	3
Percentage PAR change	59.2***	3

*$P<0.05$; **$P<0.01$; ***$P<0.001$

Table 7 Differences among malocclusion severity, occlusal result and relative improvement

Class	*n*	Pre-PAR	Post-PAR	%Δ PAR
I	96	24.6	5.1	75.1
II	187	25.4	6.3	72.3
III	19	29.6	4.9	81.3

Table 8 Results for the final examination group (1) and the control group (2) according to treatment period

End of active treatment (year)	Group	*n*	PAR pre Mean	PAR post Mean	% change Mean
1991	1	103	25.0	5.3	76.8
	2	103	26.8	6.2	74.9
1993	1	53	24.4	5.2	77.1
	2	53	26.3	7.7	67.9
1995	1	90	26.0	4.7	80.3
	2	91	28.8	8.3	67.0
1996	1	46	26.2	3.7	84.3
	2	40	28.5	6.9	71.7

Variation in PAR with time

A trend in the quality of treatment was noted in the later years for the control and especially for the final examination group (Table 8). The percentage PAR change var-

Table 9 Mean percentage reduction in weighted PAR score related to appliance type in different studies

Appliance group	Richmond et al. [14]		O'Brien et al. [10]		Fox [7]		Kelly and Springate [9]		Buchanan et al. [3]	
	n	%red	*n*	%red	*n*	%red	*n*	%red	*n*	%red
Two-arch fixed appliance	196	71.4	934	75.5	44	78.9	200	89.0	82	74.0
Single-arch fixed appliance	149	54.6	458	59.4	27	51.9	–	–	–	–
Removable appliance	559	49.8	238	51.8	12	46.3	–	–	–	–
Functional appliance	–	–	–	–	9	72.3	–	–	–	–

ied from 76.8% to 84.3% between the different time periods, the quality of treatment increased over time (non-parametric ANOVA, t=0.009). For the group as a whole, a significantly increased percentage reduction of PAR score from 1991 to 1996 was found (non-parametric ANOVA, t=0.008).

Discussion

In general a significantly higher treatment standard was found for the final examination patients. An explanation for this difference could be that more time was spent on the examination group with longer appointments and less time between two appointments, even their mean treatment duration from placement of the appliance to removal was shorter.

At the individual practitioner level, it was interesting to find the contrary for three postgraduate students. They produced better treatment outcome for the control group.

In the examination group, 3.1% of the cases remained 'no different' after orthodontic treatment or they became worse. This could be explained by the presence of an open bite due to the interposition of the tongue. Because of the large weighting of this variable (×2) the post-treatment PAR score remains relatively high.

Especially in the control group, the presence of a remaining overjet (weighted ×6) could explain the large group of no different or worse cases (Table 2). The treatment results of five postgraduate students individually showed a significantly higher overjet at post-treatment time in the controls than in the examination patients, although in general no significant difference was established between the two groups.

Residual crowding (mostly in the lower jaw, when an upper fixed appliance was used) and poor sagittal occlusion, were also post-treatment factors resulting in an increase in the post-treatment PAR score. As can be seen from Table 2, there was a low percentage of cases with a perfect lateral occlusion at the end of active treatment and this was significantly worse in the control group. This may be due to the very sensitive score for the lateral occlusion to deviations from normal: a very minor deviation from full interdigitation is scored as a non-optimal occlusion. On the other hand, a perfect center-line and alignment of the upper and lower anterior dentition was found in more than 80% of the cases in both groups.

One patient in the control group had no reduction in the PAR score. Despite the use of a fixed appliance in the upper jaw which resolved the initial crowding in the upper jaw, the deep bite, the poor sagittal occlusion and the overjet remained the same and a midline discrepancy (weighted ×4) occurred. Although the patient had no dento-occlusal PAR score improvement, he was satisfied with the increased dental aesthetics.

When the standard of treatment was assessed in relation to the appliance, it was found that the use of upper and lower fixed appliances produced the best standard of treatment. This is in agreement with the findings of other authors (Table 9). The influence of the treatment method on outcome is due to the greater control of tooth movement achieved with fixed appliances and underlines the efficiency of treatment methods that rely upon simple methods and mechanics. In the study carried out by Buchanan et al. [3] two different types of fixed appliance were compared in a group of 82 cases, treated by one consultant orthodontist (or under his supervision). It was found that the pre-adjusted Edgewise group (pretreatment PAR score=28) achieved a significantly greater reduction in PAR score than the Begg group (pretreatment PAR score=32) (81% reduction for the first versus 65% for the latter).

Functional appliances scored highly due to the marked effect on overjet (weighted ×6) and overbite (weighted ×2).

It seems that the group of patients with an Angle Class III classification had the highest treatment outcome. However, because of the low number of patients in this group no significant difference in treatment outcome between the various Angle classification groups was found. All subjects with Class III malocclusions underwent more than 70% reduction in PAR score.

Compared with the results of Wenger et al. [19], where data were analyzed derived from the Universities of Pittsburgh and Ohio State during 1971–92, the mean pretreatment severity of Class I, II and III malocclusions were different and increased with class. The occlusal results as determined by post-treatment PAR scores are similar for all classes. Otuyemi and Jones [11] evaluated 50 Class-II division 1 malocclusions treated in the Orthodontic Department of the Eastman Dental Hospital in London. They all underwent more than 70% reduction in PAR score. However, maintenance of post-treatment results 1 and 10 years post-retention was only achieved in 60% and 38% of the cases, respectively.

Although the PAR index was found to be highly reproducible in this study, it has obvious limitations in assessment of treatment outcome. Furthermore, the index only measures tooth position which, although important, is not the only factor in orthodontic treatment. The index is time dependent, a case apparently 'Greatly improved' at the end of treatment may relapse and end up as 'Improved' or even 'Worse or no different'. On the contrary, cases made 'Worse' may actually end up being 'Improved' following post-retention changes (i.e. relapse of a Class III results in a reduction of the overjet and overbite, causing the PAR score to decrease).

The PAR index does not measure inclinations/angulation of teeth, residual buccal spacing (extraction space closure), posterior alignment, or changes in arch dimensions, and it does not penalize orthodontic treatments involving inappropriate dental arch expansion [8]. Iatrogenic damage such as decalcification, root resorption, gingival recession, periodontal breakdown and facial aesthetics are obviously not measured in any way although they undoubtedly contribute to the 'quality' of treatment. The PAR index also fails to evaluate treatment suitability/motivation, the functional occlusion, the temporomandibular joint and patient satisfaction.

However, the PAR index is an epidemiological tool and was validated against a cross-section of dental opinion over a wide selection of cases [15] and if used as intended, to assess samples from case loads rather than individual cases, it is a reliable tool in assessing performance of practitioners or services.

Conclusions

The results of this study suggest that when the PAR index is used as a measure of outcome in terms of dentoocclusal change, the postgraduate students provide treatment of a high standard (>70% mean PAR score reduction).

The 'final examination' group had a significantly higher treatment outcome than the control group.

The choice of treatment method had a great influence upon the standard of treatment. Two arch fixed appliances were the most effective treatment method. Functional appliances were more effective than single arch fixed appliance and removable appliances.

The quality of the orthodontic treatment outcome in the Postgraduate Training Program in Leuven has improved since 1991.

References

1. Al Yami EA (1997) Orthodontics: treatment need and treatment outcome. Thesis, Universiteit Nymegen, The Netherlands
2. Brook PH, Shaw WC (1989) The development of an index of orthodontic treatment priority. Eur J Orthodont 11:309–320
3. Buchanan IB, Rusell JI, Clark JD (1996) Practical application of the PAR index: an illustrative comparison of the outcome of treatment using two fixed appliance techniques. Br J Orthodont 23:351–357
4. Elderton RJ, Clark JD (1983) Orthodontic treatment in the General Dental Services assessed by the Occlusal Index. Br J Orthodont 10:178–186
5. Feghali R, Nelson S, Afsharpanah A, Hans MG (1995) Assessment of an orthodontic clinic sample using the the PAR index. J Dent Res 74:139 (Abstract)
6. Feghali R, Afsharpanah A, Hans MG, Nelson S, Hassanein R (1996) Assessing orthodontic treatment outcome from 1980–1985 using the PAR index. J Dent Res 75:363 (Abstract)
7. Fox NA (1993) The first 100 cases: a personal audit of orthodontic treatment assessed by the PAR (Peer Assessment Rating) index. Br Dent J 174:290–297
8. Hinman C (1995) The Dental Practice Board: orthodontics – the current status. Practice management forum. Br J Orthodont 22:287–290
9. Kelly BM, Springate SD (1996) Specialist orthodontics in the General Dental Service. Br Dent J 180:209–215
10. O'Brien KD, Shaw WC, Roberts CT (1993) The use of occlusal indices in assessing the provision of orthodontic treatment by the hospital orthodontic services of England and Wales. Br J Orthodont 20:28–38
11. Otuyemi OD, Jones SP (1995) Long-term evaluation of treated Class II Division 1 malocclusions utilizing the PAR index. Br J Orthodont 22:171–178
12. Pickering EA, Vig P (1976) The Occlusal Index to assess orthodontic treatment. Br J Orthodont 2:47–51
13. Richmond S, Andrews M (1993) Orthodontic treatment standard in Norway. Eur J Orthodont 15:7–15
14. Richmond S, Shaw WC, O'Brien KD, Buchanan IB, Jones R, Stephens CD, Roberts CT, Andrews M (1992a) The development of the PAR index: reliability and validity. Eur J Orthodont 14:125–139
15. Richmond S, Shaw WC, Roberts CT, Andrews M (1992b) The PAR index (Peer Assessment Rating): methods to determine outcome of orthodontic treatment in terms of improvement and standards. Eur J Orthodont 14:180–187
16. Richmond S, Shaw WC, Stephens CD (1992c) Orthodontics in the General Dental Services of England and Wales, the provision of treatment. Br Dent J 172:150
17. Rinaldi A, Beck M, Vig K, Beglin F, Mayers M, VIG P (1996) Increased orthodontic treatment efficiency over two decades. J Dent Res 75:437 (Abstract)
18. Shaw WC, Richmond S, O'Brien KD, Brook P (1991) Quality control in orthodontics: indices of treatment need and treatment standards. Br Dent J 170:107–112
19. Wenger R, Douangpanya S, Vig K, Beck M, Vig P (1996) Classes I, II and III differences in severity, duration and orthodontic results. J Dent Res 75:437 (Abstract)