European Journal of Orthodontics 21 (1999) 181–192

An appraisal of the Peer Assessment Rating (PAR) Index and a suggested new weighting system

A. M. Hamdan* and W. P. Rock**
*Department of Orthodontics, University of Jordan, Amman, Jordan and
**Orthodontic Unit, University of Birmingham, UK

SUMMARY The PAR Index was developed to measure treatment outcome in orthodontics. Validity was improved by weighting the scores of some components to reflect their relative importance. However, the index still has limitations, principally due to the high weight assigned to overjet. Difficulties also arise from the application of one weighting system to all malocclusions, since occlusal features vary in importance in different classes of malocclusion.

The present study examined PAR Index validity using orthodontic consultant assessments as the 'Gold standard' and clinical ranking of occlusal features and statistical modelling to derive a new weighting system, separate for each malocclusion class.

Discriminant and regression analyses were used to derive new criteria for measuring treatment outcome. As a result a new and more sensitive method of assessment is suggested which utilizes a combination of point and percentage reductions in PAR scores. This was found to have better correlations with the 'Gold standard' than the PAR nomogram.

Introduction

Comparison of pre-, post-treatment, and post-retention dental casts helps to improve the quality of future treatments (Hickham, 1975), and occlusal indices have been developed to assess treatment standards and success (Eismann, 1974, 1980; Gottlieb, 1975; Berg and Fredlund, 1981). In order to overcome difficulties created by using indices inappropriately, Richmond (1990) introduced a new method for measuring treatment standards, the Peer Assessment Rating (PAR) Index.

Unlike previous indices the PAR Index was carefully tested for reliability and validity, being developed over a series of meetings with a group of experienced orthodontists (British Orthodontic Standards Working Party, 1986). Over 200 study casts were discussed until agreement was reached regarding individual features considered to be important in obtaining an estimate of malocclusion. A score was then allocated to each feature that deviated from the ideal, and component scores were summed to obtain a total score representing the degree of malocclusion. The index was validated using assessments of deviation from normal occlusion as the 'Gold standard' (Richmond *et al.*, 1992a) and validity was improved by assigning multipliers or 'weightings' to each component to reflect relative importance and produce a new weighted PAR total score. This is the final form in which the index was introduced.

The PAR Index is used to measure treatment outcome by comparing pre- and post-treatment weighted scores for point and percentage reductions. Improvement is categorized into three grades according to specific criteria; 'Greatly improved' requiring a score reduction of at least 22 points, 'Improved' requiring a reduction of at least 30 per cent, and 'Worse or no different' categorized by a reduction of less than 30 per cent. These criteria are graphically represented using the 'PAR nomogram' (Richmond *et al.*, 1992b).

Limitations of the PAR Index

Studies have identified limitations associated with PAR scoring (Fox, 1993; Kerr *et al.*, 1993). Problems relate mainly to the generic weighting system, particularly the respective weightings for

(a) (b)

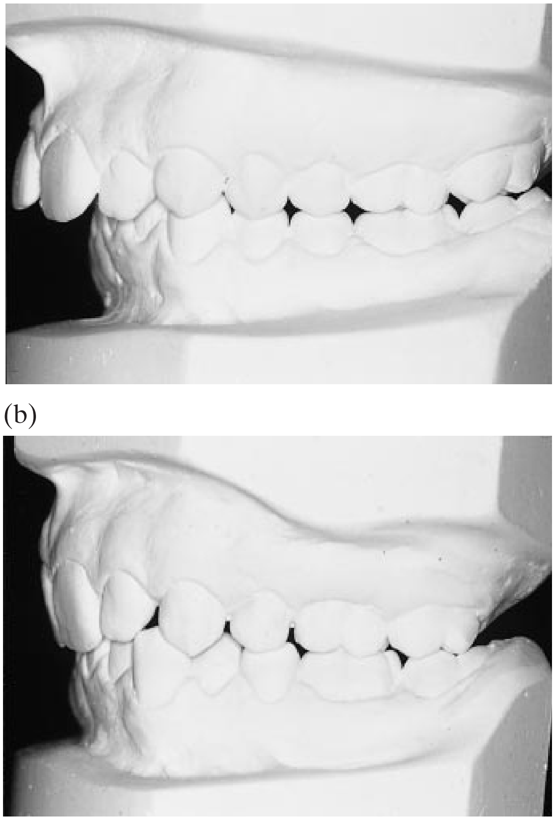

Figure 1 Treatment of an increased overjet that produced a PAR score reduction of 18 points by simple incisor retroclination.

(a) (b)

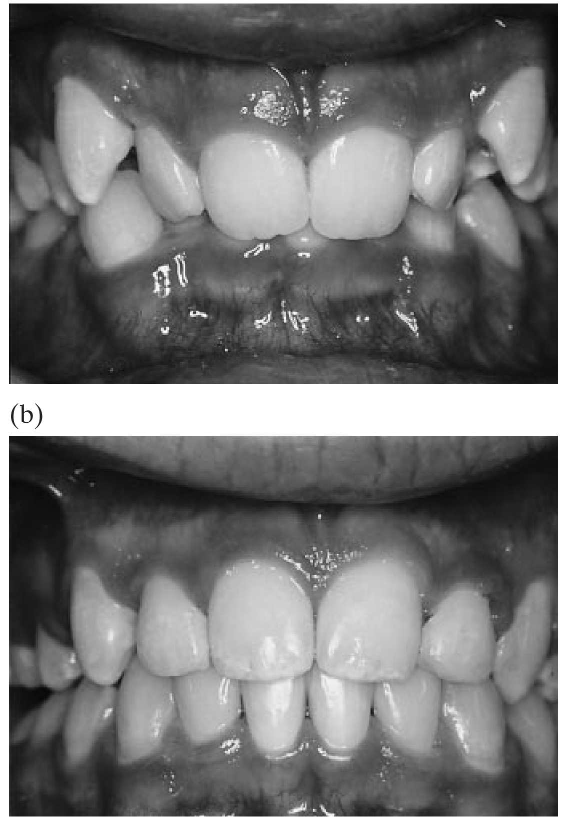

Figure 2 Treatment of traumatic overbite that reduced PAR score by only 6 points.

overjet and overbite. The high weighting for overjet may influence the index to such an extent that it is unduly sensitive in any malocclusion where overjet is increased. For example reduction of an increased overjet from 8 to 2 mm by retroclining the upper incisors (Figure 1) will reduce the PAR score by 18 points, only 4 points from a 'Greatly improved' result according to the PAR nomogram. However, the aesthetic and functional benefit of such treatment may be questioned.

On the other hand, the weighting for overbite is low, so that correction of a complete and traumatic overbite (Figure 2) merits a reduction of only 6 points, failing to represent treatment value in terms of function and appearance.

A further limitation of the PAR Index is that occlusions with initial scores of less than 22 points cannot become 'Greatly Improved' after treatment.

Finally, the zero weighting allocated to 'Displacements' in the buccal segments which include impacted teeth (Brook and Shaw, 1989) indicates that such irregularities are disregarded even though their correction may have a significant effect on treatment outcome.

Aims of the present study

1. To re-test the validity of the PAR Index against assessments by West Midlands Consultant Orthodontists.
2. To improve the validity of the index by deriving separate weighting formulae for each malocclusion class.

3. To compare the validity of three new weighting systems.
4. To apply the best new weighting system (W_{NEW}) to unweighted PAR scores and examine the effect for each malocclusion class.
5. To define criteria for allocating completed cases into different treatment outcome grades using the new weighting system, and to compare them with those of the original PAR Index.

Materials and methods

Eighty sets of pre- and post-treatment dental casts, representing equal numbers of Class I, Class II division 1, Class II division 2, and Class III cases were randomly selected from those treated at the Birmingham Dental Hospital.

PAR scoring for all 160 sets of casts was carried out by one author (AMH) whose reliability was tested against a trained and calibrated examiner who had attended calibration sessions at the Manchester Dental School with the original author of the PAR Index. A good level of agreement was found, the mean difference between scores being 0.03 unweighted PAR points (SD = 2.1; $P > 0.05$; Altman, 1991; Lowe, 1993).

Fourteen out of 16 consultant orthodontists in the Region also participated in the study. They were first asked to examine the casts, and identify important or 'Key occlusal features' from a list of eight:

1. Buccal occlusion in all three planes.
2. Centreline coincidence.
3. Crowding.
4. Impacted/missing teeth.
5. Overbite/open bite.
6. Overjet/reverse overjet.
7. Spacing.
8. Tooth morphology: abnormalities in the shape and/or size of the teeth.

The examiners were asked to identify features as either important or unimportant, and to list important features in descending order. Important occlusal features were allocated rank values ranging from 8 points for the most important, to 1 point for the least important. Unimportant occlusal features were all allocated the same rank value, equal to the average of those remaining. An example is given below;

Occlusal features of importance		**Points**
1st	Overjet	8
2nd	Overbite	7
3rd	Crowding	6
4th	______	
5th	______	
6th	______	
7th	______	
8th	______	

Unimportant features		**Points**
1.	Buccal occlusion	3
2.	Centreline	3
3.	Impacted/missing teeth	3
4.	Spacing	3
5.	Tooth morphology	3
6.	______	
7.	______	
8.	______	

The relative importance of the eight occlusal features contributing to the 80 malocclusions was expressed in two ways. 'Average rank score' was the mean of the scores allocated to each occlusal feature by the 14 examiners. 'Mean rank value' (MRV) represented the mean of the 'Average rank scores' for each malocclusion class.

Occlusal feature rankings were evaluated using principal component (PC) analysis, also known as modelling. This is a statistical method used to combine a number of variables into fewer derived variables in an attempt to provide succinct information concerning the way in which one individual differs from another (Mardia *et al.*, 1979; Chatfield and Collins, 1980; Kendal, 1980; Armitage and Berry, 1988). Plots then show whether there are natural tendencies for data to group together according to these variables, indicating that it would be inappropriate to deal with the sample as a single entity and that analysis should be carried out separately for each group. Since six variables were included in the present study, each would be expected to contribute at least one-sixth (16.7 per cent) to the total variance: any component that explained less could be excluded from the analysis.

Table 1 Mean unweighted pre-treatment PAR scores.

PAR components	Mean unweighted pre-treatment PAR scores				
	Total sample	Class I	Class II/1	Class II/2	Class III
Buccal occlusion	4.2	3.9	4.1	3.7	5.1
Centreline	0.5	0.4	0.5	0.5	0.8
Displacements					
Upper anterior	7.3	6.6	6.3	6.3	10.0
Lower anterior	3.4	3.5	3.4	3.5	3.5
Posterior (buccal)	7.3	5.9	7.7	7.3	8.2
Overbite/open bite	1.0	0.3	1.2	2.3	0.3
Overjet/reverse overjet	2.0	0.9	2.5	1.6	3.2
Total unweighted score	25.7	21.5	25.7	25.2	31.1

Table 2 Mean weighted pre-treatment PAR scores.

PAR components	Mean weighted pre-treatment PAR scores				
	Total sample	Class I	Class II/1	Class II/2	Class III
Buccal occlusion	4.2	3.9	4.1	3.7	5.1
Centreline	2.0	1.6	1.0	2.0	3.2
Displacements					
Upper anterior	7.3	6.6*	6.3	6.3	10.0
Lower anterior	3.4	3.5	3.4	3.5	3.5
Posterior (buccal)	Nil	Nil	Nil	Nil	Nil
Overbite/open bite	2.0	0.6	2.4	4.6	0.6
Overjet/reverse overjet	12.0*	5.4	15.0*	9.6*	19.2*
Total weighted score	30.9	21.6	33.2	29.7	41.6

*Highest mean sub-component score for each group.

The casts were then scored for deviation from normal occlusion on 9- and 5-point scales, anchored at one end with the term 'None', which corresponded to a score of nil, and at the other with 'Great'.

Treatment outcome was scored on 9- and 5-point scales by comparing pre- and post-treatment casts. Scales were anchored at one end with the term 'Markedly worse' which corresponded to a score of nil, and at the other with 'Greatly improved'.

Results

Pre-treatment unweighted and weighted PAR scores are shown as Tables 1 and 2. Comparison of these values shows the dramatic effect of the multipliers on weighted components, especially 'Overjet/reverse overjet', the score which increased six-fold as a result of weight multiplication. PAR scores were increased by 21 per cent for the total sample, 29 per cent for Class II division 1, 18 per cent for Class II division 2, and 32 per cent for Class III cases. Class I scores were little affected.

Occlusal feature ranking

'Crowding' and 'Overjet/reverse overjet' were the most important occlusal features for the sample as a whole (Table 3). Differences in MRVs between the whole sample and individual malocclusion classes were examined using a

Table 3 Mean rank values for the total sample and individual malocclusion classes.

Occlusal features	Mean rank values (MRVs)				
	Total sample	Class I	Class II/1	Class II/2	Class III
Buccal occlusion	4.9	4.9	4.9	4.8	4.8
Centreline	4.1	4.4	3.8	4.0	4.3
Crowding	6.6*	6.6*	5.8*	6.6*	7.2*
Impacted/missing teeth	3.8	4.5	4.0	3.7	3.0
Overbite/open bite	4.5	3.7	4.2	5.8*	4.4
Overjet/reverse overjet	5.7*	4.5	6.9*	4.8	6.5*
Spacing	3.3	3.9	3.4	3.1	2.7
Tooth morphology	3.2	3.5	3.1	3.2	3.0

*Most important occlusal features.

Table 4 Principal component analysis of average ranks.

Principal Component Analysis						
Eigen analysis of the Correlation Matrix						
Eigenvalue	1.8263	1.3623	1.1150	0.9594	0.5287	0.2084
Proportion	0.304	0.227	0.186	0.160	0.088	0.035
Cumulative	0.304	0.531	0.717	0.877	0.965	1.000
Variable	PC1	PC2	PC3	PC4	PC5	PC6
Buccal occlusion	0.325	–0.252	0.406	–0.673	0.425	0.180
Centreline	–0.026	–0.663	–0.195	0.484	0.478	0.244
Crowding	0.597	–0.256	0.035	0.115	–0.632	0.406
Impacted/missing teeth	–0.671	-0.038	0.186	–0.148	–0.205	0.670
Overbite	0.135	0.236	–0.820	–0.327	0.152	0.351
Overjet	0.261	0.612	0.298	0.412	0.350	0.414

Mann–Whitney test (two-sample Wilcoxon rank sum test). The following MRV 'row differences' were significant:

1. 'Crowding' in Class II division 1 cases was significantly less important than for the whole sample, Class I, and Class III cases ($P < 0.05$).
2. 'Impacted/missing teeth' was significantly less important in Class III cases than for others ($P < 0.05$).
3. 'Overbite/open bite' was significantly more important in Class II division 2 cases than for others ($P < 0.005$).
4. 'Overjet/reverse overjet' was significantly more important in Class II division 1 and Class III cases than for others ($P < 0.005$).

The results for PC analysis of occlusal feature ranking are shown as Table 4. 'Spacing' was scored as important in only 11.3 per cent of cases (nine cases) and 'Tooth morphology' in seven (8.8 per cent) of cases. These features were therefore excluded and PC analysis was carried out on the average ranks of the remaining six occlusal features.

The 'Eigenvalues' of the correlation matrix represent the relative contribution of each PC to total variance. PCs 1–3 accounted for 30.4, 22.7, and 18.6 per cent, respectively, to give a combined contribution of 72 per cent. The remaining PCs each explained less than 16.7 per cent of the variance and were therefore excluded from the analysis.

Table 5 An example of the calculation of principal component scores prior to three-dimensional plotting.

Variables	Average rank (14 examiners)	Average rank × PC1	Average rank × PC2	Average rank × PC3
Buccal occlusion	3	3 × 0.325	3 × –0.252	3 × 0.406
Centreline	3	3 × –0.026	3 × –0.663	3 × –0.195
Crowding	6	6 × 0.597	6 × –0.256	6 × 0.035
Impacted/missing teeth	3	3 × –0.671	3 × –0.038	3 × 0.186
Overbite	7	7 × 0.135	7 × 0.236	7 × –0.820
Overjet	8	8 × 0.261	8 × 0.612	8 × 0.298
Spacing*	3			
Tooth morphology*	3			
Total PC score		5.50	2.15	–2.00

*Spacing and tooth morphology were excluded from the analysis.

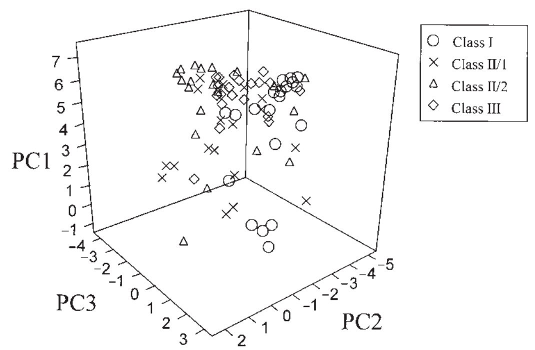

Figure 3 A three-dimensional plot of main principal components.

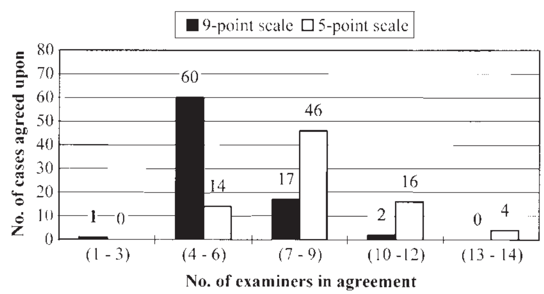

Figure 4 Inter-examiner agreements for assessment of deviation from normal occlusion.

For each of the 80 cases the average rank of 14 examiners for each variable was multiplied by its corresponding PC coefficient. The products were then summed to produce PC totals for each case (Table 5). Using the three PC scores for each case as co-ordinates, a three-dimensional plot was constructed for all 80 cases. There was a tendency for the symbols representing each malocclusion class to form clusters (Figure 3). Such-dimensional clustering (Minitab Reference Manual, 1996) supports the hypothesis that it is inappropriate to group all orthodontic cases together to derive a generic weighting formula, and that each class of malocclusion should be considered separately.

Deviation from normal occlusion

The 14 consultants agreed more closely when using the 5-point than the 9-point scale (Figure 4). Agreement was best for extreme cases, that is those with very mild or very great deviation from normal.

Treatment outcome

Treatment outcome scores were calculated for each case from the scores of the 14 consultants (Table 6). Reductions in PAR totals as a result of treatment placed most results between the 'Improved' and 'Greatly improved' categories.

Table 6 Treatment outcome scores.

		Treatment outcome	
		9-point scale	5-point scale
Total sample (80 cases)	Mean	6.4	3.3
	SD	0.9	0.5
	Median	6.6	3.4
	Range	3.8–7.7	1.9–4.0
Class I (20 cases)	Mean	6.3	3.3
	SD	0.8	0.4
	Median	6.4	3.4
	Range	4.4–7.5	2.3–3.9
Class II division 1 (20 cases)	Mean	6.1	3.2
	SD	1.0	0.6
	Median	6.5	3.4
	Range	4.2–7.3	2.0–3.9
Class II division 2 (20 cases)	Mean	6.6	3.4
	SD	1.0	0.5
	Median	6.8	3.6
	Range	4.0–7.7	1.9–3.9
Class III (20 cases)	Mean	6.4	3.4
	SD	1.0	0.5
	Median	6.6	3.5
	Range	3.8–7.9	1.9–3.9

Derivation of new weightings

During validation of the PAR Index the original authors felt that direct summing of component scores might not produce the best index, as the profession might place greater importance on certain aspects of malocclusion than others. Multiple regression (MR) was therefore used to derive weightings for components as a means of improving validity and a generic weighting formula was derived.

PC analysis in the present study has indicated that it is inappropriate to use a generic weighting formula for all malocclusions. Possible new systems and weightings were therefore investigated for each class of malocclusion using weightings derived from mean rank values, and from principal component and multiple regression analysis, respectively.

Mean rank value (MRV) weighting (Table 7)

MRVs were modified to be used as weightings for occlusal features by, first, rounding to the nearest whole number. Features with an MRV of four or less were allocated a weighting of one. Occlusal features with higher MRVs were allocated progressively higher weightings to derive a total W_{MRV} score

PC coefficient weighting (Table 8)

Weightings were derived by multiplying PC coefficients by 10 and rounding to the nearest whole number. The resultant ‘PC Weightings’ (W_{PC}) were used as numerical values irrespective of sign so that negative scores did not reduce the total and, thus, appear to indicate a decrease in the severity of malocclusion. Percentage contributions to total variance were higher when weights were derived for malocclusion classes separately than for the whole sample.

Table 7 Mean rank values and derived weights for six occlusal features.

Occlusal features	Mean rank values (MRV) and derived weights									
	Total sample		Class I		Class II/1		Class II/2		Class III	
	MRV	Wt	MRV	Wt	MRV	Wt	MRV	Wt	MRV	Wt
Buccal occlusion	5	2	5	2	5	2	5	2	5	2
Centreline	4	1	4	1	4	1	4	1	4	1
Crowding	7	4	7	4	6	3	7	4	7	4
Impacted/missing teeth	4	1	4	1	4	1	4	1	3	1
Overbite/open bite	5	2	4	1	4	1	6	3	4	1
Overjet/reverse overjet	6	3	5	2	7	4	5	2	7	4

Table 8 Principal component weights (W_{PC}).

Occlusal features	Principal component weightings (W_{PC})				
	Total sample	Class I	Class II/1	Class II/2	Class III
Buccal occlusion	3	3	2	4	1
Centreline	0	2	4	2	4
Crowding	6	6	4	6	3
Impacted/missing teeth	7	6	6	6	6
Overbite/open bite	1	1	1	2	3
Overjet/reverse overjet	3	3	5	3	6
Eigenvalue	1.8	2.3	2.4	1.9	2.0
Percentage contribution to total variance	30	38	40	32	33

Table 9 Multiple regression weights (W_{MR}). Also adopted as (W_{NEW}).

PAR components	Multiple regression (MR) weightings				
	Total	Class I	Class II/1	Class II/2	Class III
Buccal occlusion	1	1	1	1	1
Centreline	1	4	2	2	1
Displacements					
Upper anterior	1	2	1	2	1
Lower anterior	1	1	1	1	1
Posterior (buccal)	0	0	0	0	0
Overbite/open bite	2	3	2	5	6
Overjet/reverse overjet	2	2	6	1	5
R^2	45.2%	84.1%	72.8%	57.2%	62.3%

Multiple regression weighting (Table 9)

Median 5-point scale assessments were chosen to derive MR weights (W_{MR}). Partial regression coefficients of each PAR component were multiplied by 10 and rounded to the nearest whole number irrespective of sign. R^2 values were higher when weightings were derived separately for each malocclusion class. To reflect the importance of buccal occlusion irregularities and lower incisor displacements base weights of one were applied to these scores.

Weightings derived by multiple regression had the best correlation with consultant orthodontic opinion in terms of deviation from normal occlusion (Table 10). They were therefore used to calculate new weighted scores for each malocclusion class (Table 11). Comparison of Tables 2 and 11 shows that the same PAR components had the highest weighted scores for all but Class II division 2 cases. The greatest contribution to Class II division 2 totals came from 'Overbite/open bite'.

Allocation of treatment outcome grades using the new weighting system

Both 9- and 5-point scales were collapsed into three treatment outcome categories, and discriminant analysis was used to create a linear separation between them and achieve best fit with the 'Gold standard'. Three predictor variables were used to identify criteria for allocation into each grade using the new weighting system (W_{NEW}). These were total weighted post-treatment score, point

Table 10 Validation of different PAR Index weighting systems according to Pearson's Correlation Coefficients.

		9-point scale deviations	
		Mean	Median
W_{ORG}	Total sample	0.62	0.63
W_{MRV}	Total sample	0.53	0.51
(Table 7)	Class I	0.66*	0.73*
	Class II division 1	0.54	0.53
	Class II division 2	0.34	0.21
	Class III	0.53	0.49
W_{PC}	Total sample	0.51	0.49
(Table 8)	Class I	0.67*	0.73*
	Class II division 1	0.39	0.39
	Class II division 2	0.35	0.24
	Class III	0.67*	0.66*
W_{MR}	Total sample	0.66*	0.68*
(Table 9)	Class I	0.68*	0.75*
	Class II division 1	0.75*	0.78*
	Class II division 2	0.73*	0.70*
	Class III	0.81*	0.80*

*Correlations higher than corresponding original PAR Index weightings.

reduction in weighted scores, and percentage reductions in scores as a result of treatment.

Individual variables were found to be poor predictors of treatment outcome and the best fit with the 9-point assessment scale was obtained by using combinations of predictors. There were significant differences between the means of predictor variables for all treatment outcome groups except pre-treatment totals ($P < 0.05$), although the ranges overlapped considerably between groups. Tukey pairwise comparisons indicated that the greatest overlap was between the 'Improved' and 'Greatly improved' categories (Table 12).

Prediction of treatment outcome using pre- and post-treatment scores had a Spearman's rank correlation coefficient (r_s) of 0.57 with consultant gradings. When using point and percentage score reductions as predictor variables correlation with consultant gradings was 0.67 (Figure 5). The majority of 'Greatly improved' cases were clustered in the upper right-hand corner of the plot, indicating a combined reduction of more than 22 points or 87 per cent. The majority of 'Improved' cases had a percentage reduction of more than 30 per cent, but since the points reductions were 22 or less they were not allocated to the 'Greatly improved' category. All but one of the cases assessed as 'Worse/no different' were located below the 30 per cent reduction mark.

Application of the original criteria for categorization of treatment outcome (PAR nomogram)

The original criteria for categorization of treatment outcome (Richmond *et al.*, 1992b) were applied to newly-derived pre- and post-treatment scores (Figure 6). The greatest overlap was between the 'Improved' and 'Greatly improved'

Table 11 Mean pre-treatment scores using the suggested new weights.

PAR components	New PAR weighting system; mean pre-treatment weighted PAR scores			
	Class I	Class II/1	Class II/2	Class III
Buccal occlusion	3.9	4.1	3.7	5.1
Centreline	1.6	0.8	1.0	0.8
Displacements				
Upper anterior	13.2*	6.3	12.6*	10.0
Lower anterior	3.5	3.4	3.5	3.5
Posterior (buccal)	Nil	Nil	Nil	Nil
Overbite/open bite	0.9	2.4	11.5	1.5
Overjet/reverse overjet	1.8	15.0*	1.6	16.0*
Total weighted score	24.9	32.0	33.9	36.9

*Highest sub-component scores for each group.

Table 12 Descriptive statistics of the 4 predictor variables classified according to the 9-point scale 'Gold standard'.

	Pre-treatment total	Post-treatment total	Point reduction	Percentage reduction
Worse/no different (5 cases)				
Mean	24.6	21.8	3	10
Standard deviation	5.9	8.6	7	30
Minimum	15	12	–4	–20
Maximum	31	35	14	54
Improved (67 cases)				
Mean	25.5	6.3	19	77
Standard deviation	10.5	7.6	10	25
Minimum	6	0	–15	–48
Maximum	49	46	44	100
Greatly improved (8 cases)				
Mean	32.5	1.4	31	96
Standard deviation	6.3	1.8	6	5
Minimum	23	0	23	87.5
Maximum	40	4	39	100

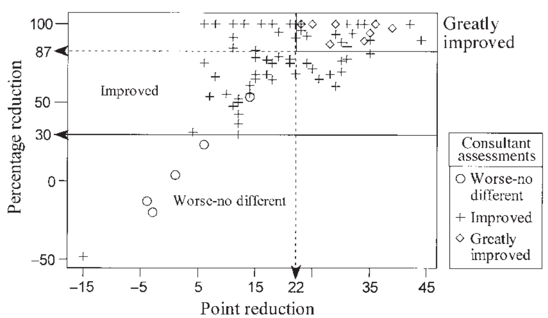

Figure 5 Categorization of treatment outcome using point and percentage reductions in PAR scores.

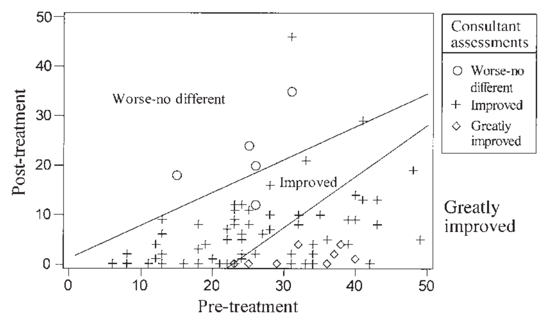

Figure 6 Categorization of treatment outcome according to the original PAR nomogram.

categories. This method had a correlation (r_s) of 0.56 with consultant assessments.

Discussion

The availability of a definition for 'Normal' occlusion (Andrews, 1972) set a standard for comparison of treatment outcomes. On the other hand, lack of a clear definition for what constitutes deviation from normal occlusion may have increased the subjectivity of assessments. In the present study, discussions with examiners after assessments were completed revealed that many had been reluctant to give the maximum deviation score, reserving it only for the most severe cases. Furthermore, they would have liked the opportunity to go back and re-score certain cases pre-treatment, once they had an idea of the full range of deviations from normal within the sample.

Using the standard PAR scoring system, treatment results for each malocclusion class were heavily dependent on the number of 'Greatly improved' cases in the group. Pre-treatment scores made a significant contribution to treatment outcome since the higher the pre-treatment score, the easier it was to achieve the 22 point reduction needed for a 'Greatly improved' result. Class III

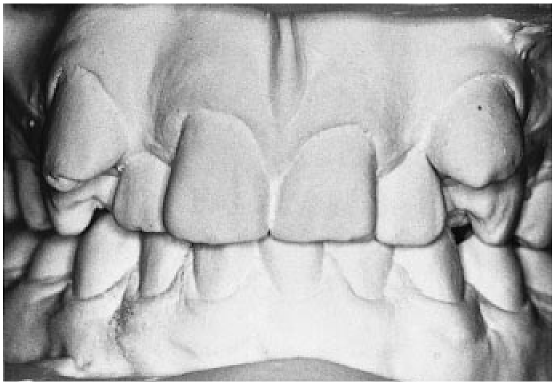

Figure 7 A Class I malocclusion with 'Great' deviation from normal occlusion, but a weighted PAR score of only 15.

cases had the highest mean pre-treatment score (41.6 points) and Class I cases the lowest (21.6 points), since three weighted occlusal features (overjet/reverse overjet, overbite/open bite, and centreline) made a minimal contribution to the final score in this group. The mean pre-treatment score for Class I fell short of the minimum 22-point reduction needed to achieve a 'Greatly improved' result. Figure 7 illustrates an example of this problem; this Class I malocclusion was scored as having a 'Great' deviation from normal occlusion by the consultant panel, yet its weighted pre-treatment PAR score was only 15 points. Correction of the malocclusion to an ideal occlusion (reduction of PAR score to nil) would still fall short of a 'Greatly improved' result.

The multiple regression weighting system (W_{MR}) had the best validity of the four methods tested and the values for coefficients were higher when individual malocclusions were considered than when the sample was taken together (Table 10). The suggested new weighting system is in accordance with the clinical characteristics of each malocclusion class in that 'Overbite/open bite' becomes the occlusal feature with the highest weighting in Class II division 2 and Class III malocclusions, whilst for Class I cases, 'Centreline' has the greatest weight. 'Overjet/reverse overjet' remained the most important occlusal feature in Class II division 1 malocclusions.

The new system also goes some way to reducing the range between pre-treatment PAR scores for the four main malocclusion classes. The average Class I case now has a chance to become 'Greatly improved' as a result of treatment.

Conclusions

The results support the hypothesis that it is inappropriate to group all orthodontic cases together to derive a *generic* weighting formula, and that weightings should be derived separately for each malocclusion class since:

1. Individual malocclusion class PC weightings had higher percentage contributions to the total variance than generic PC weightings, suggesting that scoring would be more valid if each class of malocclusion was dealt with separately.
2. R^2 values were substantially increased when individual malocclusion class MRV weightings were compared with generic MRV weightings.
3. The validity of multiple regression weightings for individual malocclusion classes was higher than those of the total sample weightings.

The most valid PAR index weightings were derived by multiple regression, modified by the addition of baseweights for buccal occlusion and lower anterior displacements.

Assessments of treatment outcome using point and percentage reductions were more valid than using the original PAR nomogram.

Address for correspondence

Dr Peter Rock
School of Dentistry
St Chad's Queensway
Birmingham B4 6NN, UK

Acknowledgements

We wish to record our grateful thanks to the 14 consultants who so generously took the time and trouble to make this work possible. The statistical advice of Tim Marshall was invaluable. Verona Needham gave expert secretarial help.

References

Altman D G 1991 Practical statistics for medical research. Chapman and Hall, London

Andrews L 1972 The six keys to normal occlusion. American Journal of Orthodontics 62: 296–309

Armitage P, Berry G 1988 Statistical methods in medical research, 2nd edition. Blackwell Scientific Publications, Oxford

Berg R, Fredlund A 1981 Evaluation of orthodontic treatment results. European Journal of Orthodontics 3: 181–185

British Orthodontic Standards Working Party 1986 Second report. British Journal of Orthodontics 13: 165–173

Brook P, Shaw W 1989 The development of an index of orthodontic treatment priority. European Journal of Orthodontics 11: 309–320

Chatfield C, Collins A J 1980 Introduction to multivariate analysis. Chapman and Hall, London

Eismann D 1974 A method of evaluating the efficiency of orthodontic treatment. Transactions of the European Orthodontic Society pp. 223–232

Eismann D 1980 Reliable assessment of morphological changes resulting from orthodontic treatment. European Journal of Orthodontics 2: 19–25

Fox N A 1993 The first 100 cases: a personal audit of orthodontic treatment assessed by the PAR (Peer Assessment Rating) Index. British Dental Journal 174: 290–297

Gottlieb E 1975 Grading your orthodontic treatment result. Journal of Clinical Orthodontics 9: 155–161

Hickham J H 1975 Directional edgewise orthodontic approach. Journal of Clinical Orthodontics 9: 143–154

Kendal M G 1980 Multivariate analysis, 2nd edition. Griffin, London

Kerr W J S, Buchanan I B, McColl J H 1993 Use of the PAR Index in assessing the effectiveness of removable orthodontic appliances. British Journal of Orthodontics 20: 351–357

Lowe D 1993 Planning for medical research: a practical guide to research methods. Astraglobe Ltd., Middlesborough, UK

Mardia K V, Kent J T, Bibby J M 1979 Multivariate analysis. Academic Press, London

Minitab Reference Manual 1996 Release 10 for Windows. Minitab, Boston, USA

Richmond S 1990 A critical evaluation of orthodontic treatment in the General Dental Services in England and Wales. PhD Thesis, University of Manchester

Richmond S *et al.* 1992a The development of the PAR Index (Peer Assessment Rating): reliability and validity. European Journal of Orthodontics 14: 125–139

Richmond S, Shaw W C, Roberts C T, Andrews M 1992b The PAR Index (Peer Assessment Rating): methods to determine outcome of orthodontic treatment in terms of improvement and standards. European Journal of Orthodontics 14: 180–187