European Journal of Orthodontics 20 (1998) 159–167

The influence of operator changes on orthodontic treatment times and results in a postgraduate teaching environment

Niall J. McGuinness and James P. McDonald*
Departments of Orthodontics, Merlin Park Hospital, Galway, Ireland and *Victoria Hospital, Kirkcaldy, UK

SUMMARY There is some concern that patients treated by postgraduate students during their speciality training may be disadvantaged in the quality of treatment result and subject to longer treatment times if they are treated by more than one operator. This study analysed the pre- and post-treatment study models of orthodontic patients from a district general hospital department, debonded by one junior member of junior staff in the calendar years 1991 and 1992. All patients were treated using the same pre-adjusted edgewise appliance in both arches. Two groups of patients were identified: patients whose treatment was started and finished by the same operator (registrar A), and those whose treatment was begun by another operator (registrar B), but finished by registrar A. A random sample of 30 patients from each group was selected and the study models at the start and end of treatment were scored using the PAR (Peer Assessment Rating) Index.

No significant differences were found between the average PAR scores for each group at the beginning of treatment, or between groups A and B at the end of treatment; the average treatment time for the patients treated by one operator was 17.67 months (SD 4.15 months), while the average treatment time for the patients treated by more than one operator was 26.1 months (SD 6.78 months). Statistically, this difference was highly significant ($P < 0.001$). A highly linear relationship ($R^2 = 0.92$) was found between the percentage PAR score reduction and the initial PAR score. The findings and their implications are discussed.

Introduction

There exists a widely-held impression that a change of operator during a course of orthodontic treatment results in a longer treatment time, but no data appear to be available to support this. In teaching and district general hospitals in the United Kingdom, junior staff changes occur frequently and arrangements must be made for patients' treatment to be carried on by another staff member. In the senior registrar grade (which is generally for 3–4 years), an opportunity exists to accumulate a sizeable number of completed cases before moving to a consultant post, thereby giving the clinician a chance to assess the results of treatment and provide an opportunity for self-audit.

In order to assess the outcome of treatments in medicine and dentistry, indices of various types have been developed. Indices have been used to categorize medical and dental disorders for the purposes of epidemiology and research, and to allocate patients into categories of treatment need. Malocclusion is unique in that it presents a number of unrelated traits to which different patients may react in different ways. In order to categorize dental malocclusions, several indices have been developed (McGuinness and Stephens, 1994). All occlusal indices should have a number of characteristics, the most important of which are: (a) validity; (b) reproducibility; (c) ease of use; and (d) acceptability to the profession and public alike.

The Peer Assessment Rating (PAR) Index assigns numerical values to various traits that

make up a malocclusion and the sum of these represents the degree of departure of the occlusion from the ideal (i.e. where there is an excellent occlusion the score is near zero) (Richmond *et al.*, 1992). This index has been validated by a panel of 74 dentists, comprising 22 consultant orthodontists, 22 specialist practitioners, 11 community dental officers, 15 general dental practitioners, and four others (advisers from the Dental Practice Board and junior orthodontic hospital staff) to ensure that they reflected contemporary orthodontic opinion in the UK (Shaw *et al.*, 1991). The PAR Index has been used extensively since to assess the standard of treatment results in a number of studies in the UK (McGuinness and Stephens, 1993; O'Brien *et al.*, 1993), Norway (Richmond and Andrews, 1993) and has been adapted for use in the North American context (DeGuzman *et al.*, 1995).

The standard of orthodontic treatment results in relation to treatment duration has been examined by Fink and Smith (1992). In their study of patients who had undergone fixed appliance therapy, it was found that the time spent in detailed finishing by individual clinicians was an important source of unexplained variation in treatment duration. Kelly (1994), and Kelly and Springate (1996) examined the results of fixed upper and lower appliance orthodontic treatment from 10 specialist practitioners, each of whom supplied 20 sets of study models of their most recently treated cases. It was found that the average treatment time was not related to the extraction pattern, but to the use of headgear. Kerr *et al.* (1994), in a study of removable appliance cases, concluded that the initial severity of the malocclusion (as measured by the Peer Assessment Rating Index of Malocclusion) was a good predictor of the duration of treatment. However, no studies have examined the influence of change of operator during treatment on the standard of treatment result or the duration of treatment.

The objectives of this study were:-

1. To determine if there was any difference between the standard of treatment result in two groups of patients, one of which had their treatment carried out in its entirety by a single operator, the other by two operators, using the PAR (Peer Assessment Rating) Index. (Richmond *et al.*, 1992).
2. To determine if the average length of treatment time for the above two groups was different, and if this difference in treatment time influenced the standard of treatment result.

Materials and methods

The clinical records of all patients treated with fixed appliances by one particular junior staff member (registrar A) in a district general hospital orthodontic department for the calendar years 1991 and 1992 were examined. One-hundred-and-eighty-three patients were debonded during this period, and two groups were identified: a smaller one of 36 patients (Group A) who had their fixed appliance treatment started and completed by registrar A; and a larger group of 147 patients (group B) who had their treatment started by registrar A's immediate predecessor (registrar B), but whose treatment was taken over by and who were debonded by registrar A. Continuity of treatment for all patients was maintained in that registrar A succeeded registrar B immediately registrar B vacated the post. The clinic was supervised at all times by a consultant orthodontist, from whom advice and guidance on treatment could be sought.

Before selecting samples from each of these groups, the following patients were eliminated: (a) patients who had orthognathic surgery as part of treatment; (b) patients with cleft lip and palate; and (c) patients who had one-arch fixed appliance therapy only. All patients were treated with the same type of pre-adjusted edgewise appliance, and attended for appliance adjustment every 5–6 weeks during treatment. Thirty patients were then randomly selected from each group, and the following data obtained: (a) the PAR score of the initial (pre-treatment) study models; (b) the length of treatment time (from date of initial bonding to date of debond); (c) the PAR score of the final (day of debond) study models.

The number of broken appointments was collated and found to be the same for both groups. The PAR scores were obtained by use of the Peer

Assessment Rating transparent ruler (Richmond *et al.*, 1992); in all, a total of 60 sets of models were scored. The reproducibility of the scores was tested by re-scoring a random selection of 25 per cent of the study models, and a paired *t*-test was performed, which confirmed reproducibility of scoring. The data for treatment times and the PAR scores were analysed, and the mean and standard deviation (SD) obtained; the figures were further analysed for any significant differences by use of a *t*-test and Mann–Whitney *U*-test to establish if there was any difference between the two groups at the beginning and end of treatment. Pearson and Spearman correlation coefficients were then used to determine if there was a relationship between the initial PAR score and the change in PAR score, as well as to establish if there was any relationship between the initial PAR score and the length of treatment, and whether the average length of treatment was different in groups A and B.

Results

The PAR scores and percentage reductions for groups A and B are given in Tables 1 and 2. Table 3 shows the differences in PAR scores at the start of treatment between groups A and B analysed using a *t*-test, and Mann–Whitney *U* (Wilcoxon Rank Sum W) test. No significant differences were found between the two groups. In Table 4, these tests were repeated for the final PAR scores of groups A and B. In this instance, it was found using Levene's test statistic ($F = 11.29$, $P = 0.001$) that there was sufficient evidence to reject the null hypothesis; however, the Mann–Whitney *U*-test showed that this difference was not significant ($Z = -0.814$, $P = 0.42$).

Table 1 Group A patients: one operator treatment.

Patient no.	PAR score (start)	PAR score (end)	Reduction (per cent)	Treatment time (months)
1	16	11	31.3	20
2	47	21	55.3	23
3	37	11	70.3	20
4	23	2	91.3	11
5	31	1	96.8	25
6	30	3	90.0	20
7	46	12	73.9	22
8	33	3	90.9	13
9	31	12	61.3	17
10	35	2	94.3	15
11	12	5	58.0	20
12	46	4	95.0	19
13	18	16	11.1	8
14	27	2	92.6	17
15	15	2	86.7	12
16	22	4	81.8	20
17	23	10	56.5	15
18	21	10	52.4	12
19	43	5	88.4	16
20	38	2	94.7	24
21	14	6	57.1	16
22	24	2	91.7	17
23	38	2	94.7	20
24	22	11	50.0	15
25	20	2	90.0	24
26	22	11	50.0	21
27	13	11	15.4	15
28	32	11	65.6	20
29	24	11	54.2	18
30	39	6	84.6	15

Table 2 Group B patients: treatment by more than one operator.

Patient no.	PAR score (start)	PAR score (end)	Reduction (per cent)	Treatment time (months)
31	23	6	73.0	19
32	40	4	90.0	36
33	38	8	78.9	26
34	39	11	71.8	29
35	33	2	93.9	20
36	34	10	70.6	27
37	27	13	51.8	23
38	48	2	95.8	33
39	36	5	86.1	33
40	28	3	89.3	22
41	18	4	77.7	34
42	41	4	90.2	21
43	39	2	94.9	13
44	30	2	93.3	31
45	41	2	95.1	32
46	25	4	84.0	24
47	22	9	59.1	16
48	22	4	81.8	38
49	46	4	91.3	29
50	27	5	81.5	26
51	24	2	91.7	22
52	23	4	82.6	29
53	38	3	92.1	23
54	25	8	68.0	14
55	28	10	64.3	30
56	31	9	71.0	20
57	17	4	76.5	40
58	20	10	50.0	28
59	22	6	72.7	29
60	35	7	80.0	26

The changes in PAR scores during treatment were similarly tested, and no significant differences were found between groups A and B (Table 5). Examining the data for the length of treatment for the two groups (Table 6) it was found that the average length of treatment in group A was 17.67 months (SD 4.17 months) and in group B, 26.1 months (SD 6.78 months). This difference was found to be statistically significant using Levene's statistic ($F = 6.07$, $P = 0.000$) and highly significant using the Mann–Whitney U-test ($Z = -4.728$, $P = 0.000$).

Pearson and Spearman correlation coefficients (rankings) were calculated for the variables (Figure 1) and it was found that a highly linear relationship ($R^2 = 0.92$) existed between the initial PAR score and the change in PAR score for both groups, suggesting that the overall change in PAR score can be predicted before treatment from the initial score.

There was a reduction of 70.85 per cent in the average PAR score for group A, while group B showed a reduction of 79.95 per cent (Figure 2). When groups A and B were compared, it was found that 50.0 per cent of group A patients fell into the 'greatly improved' category, 43.3 per cent in the 'improved' category, and 6.7 per cent in the 'worse-no different category'. The corresponding figures for the group B patients were 56.7, 43.3, and 0 per cent, respectively.

Discussion

This study has shown that there is a highly significant difference in treatment times between two groups of patients, one of which (group A)

Table 3 PAR scores before treatment: means, standard deviations, and comparison of intergroup scores.

(a) *t*-Test for independent samples of group

Variable PAR start	No. of cases	Mean	SD	SE of mean
Group A	30	28.06	10.41	1.9
Group B	30	30.66	8.51	1.55

Mean difference = –2.6.
Levene's test for equality of variances: $F = 1.657$, $P = 0.203$ (NS).

(b) *t*-Test for equality of means

Variances	*t*-Value	df	Two-Tailed significance	SE of difference	95% CI for difference
Equal	–1.06	58	0.294	2.455	(–7.514, 2.314)
Unequal	–1.06	55.8	0.294	2.455	(–7.518, 2.318)

(c) Mann–Whitney U–Wilcoxon Rank Sum W test

Mean rank	Cases			
27.80	30	Group = 1 Group A		
33.20	30	Group = 2 Group B		
	—			
	60 Total			
U	*W*	*Z*	Two-Tailed *P*	
369.0	834.0	–1.1989	0.2306	

was treated from start to finish by one clinician (registrar A) and the second of which (group B) had their treatment started by another clinician (registrar B), but whose treatment was completed by registrar A. All patients were recalled every 5–6 weeks for appliance adjustment during treatment. The average treatment time in group A was 17.67 months (SD 4.15 months), while that in group B was 26.1 months (SD 6.78 months). This difference was found to be highly significant, using both Levene's and Mann–Whitney statistical tests. It would appear, therefore, that change of operator has a profound effect on the total treatment time, increasing it by an average of 8.43 months. However, it was found that increased treatment time had no effect on the standard of orthodontic treatment result as assessed by the PAR score. The average PAR score for group A at the end of treatment was 7.03, while that for group B was 5.57. This difference was not found to be statistically significant. The difference in average PAR scores between the two groups at the beginning of treatment was also found not to be statistically significant.

However, there are many aspects of malocclusion that are not considered by the PAR Index. These include the angulation of the upper incisors, the angulation of the buccal segments

Table 4 PAR scores after treatment: means, standard deviations, and comparison of intergroup scores.

(a) *t*-Test for independent samples of group

Variable PAR start	No. of cases	Mean	SD	SE of mean
Group A	30	7.033	5.096	0.93
Group B	30	5.5667	3.17	0.579

Mean difference = 1.4667.
Levene's test for equality of variances: $F = 11.292$, $P = 0.001$.

(b) *t*-Test for equality of means

Variances	*t*-Value	df	Two-Tailed significance	SE of difference	95% CI for difference
Equal	1.34	58	0.186	1.096	(–0.726, 3.66)
Unequal	1.34	48.52	0.187	1.096	(–0.726, 3.669)

(c) Mann–Whitney U–Wilcoxon Rank Sum W test

Mean rank	Cases			
32.32	30	Group = 1 Group A		
28.68	30	Group = 2 Group B		
	—			
	60 Total			
U	*W*	*Z*	Two-Tailed *P*	
395.5	969.5	–0.8144	0.4154 (NS)	

and root parallelism, the spacing or crowding of the buccal segments, facial aesthetics, iatrogenic effects of treatment, and the potential for long-term relapse (Richmond *et al*., 1993). Further developments are needed in order to make the PAR Index more sensitive to treatment outcomes.

The relationship between the initial PAR score at the start of treatment and the reduction in PAR score during treatment was found to be highly linear using the Pearson correlation coefficient ($R^2 = 0.92$). This suggests that the change in PAR score during treatment can be predicted from the initial PAR score. Kerr *et al.* (1994) in a study of removable appliance treatment, suggested that the initial PAR score was a good predictor of the length of treatment time where removable appliance therapy was used. In the present study, the use of the same type of fixed appliance system for each patient treated probably contributes to the highly linear relationship between initial PAR score and the PAR score reduction observed.

In examining the results of this study, it is interesting to note that group A had 50 per cent of patients in the 'greatly improved' category, 43.3 per cent in the 'improved category', and 6.7 per cent were in the 'worse-no different'

Table 5 Changes in PAR scores.

(a) *t*-Test for independent samples of group

Variable PAR start	No. of cases	Mean	SD	SE of mean
Group A	30	21.033	11.385	2.079
Group B	30	25.1	9.65	1.762

Mean difference = –4.0667
Levene's test for equality of variances: $F = 1.061$, $P = 0.307$ (NS)

(b) *t*-Test for equality of means

Variances	*t*-Value	df	Two-Tailed significance	SE of difference	95% CI for difference
Equal	–1.49	58	0.141	2.725	(–9.521, 1.388)
Unequal	–1.49	56.48	0.141	2.725	(–9.524, 1.391)

(c) Mann–Whitney U–Wilcoxon Rank Sum W test

Mean rank	Cases			
27.37	30	Group = 1 Group A		
33.63	30	Group = 2 Group B		
	—			
	60 Total			
U	*W*	*Z*	Two-Tailed *P*	
356.0	821.0	–1.3911	0.1642	

category. The corresponding figures for the group B patients were 56.7, 43.3, and 0 per cent, respectively. The figures for group A compare favourably with the 12 per cent in the 'worse-no different' category given in the study by Fox (1993). However, it should be noted that the figures by Fox were for a registrar training for a specialist qualification in orthodontics, while the registrars in the present study (registrars A and B) had completed their 3-year training before commencing treatment of the patients in groups A and B. In order for a case to achieve 'greatly improved' status, the PAR score reduction must be reduced by 22 points or 30 per cent, whichever is the greater. It would appear that the longer treatment time for group B has resulted in no cases falling into the 'worse-no different' category, but this difference cannot be detected for the overall reductions by the statistical tests.

In the analysis of the results, it would appear that there is no difference between the two groups (A and B) in the standard of treatment result. However, the average treatment time for group A was almost 9 months shorter than those in group B, the patients in group B being treated by two different clinicians. The influence of the pre-adjusted appliance system used in all cases cannot be underestimated—it is probably very

Table 6 Length of treatment.

(a) *t*-Test for independent samples of group

Variable PAR start	No. of cases	Mean	SD	SE of mean
Group A	30	17.6667	4.147	0.757
Group B	30	26.1	6.779	1.238

Mean difference = –8.433.
Levene's test for equality of variances: $F = 6.067$, $P = 0.017$.

(b) *t*-Test for equality of means

Variances	*t*-Value	df	Two-Tailed significance	SE of difference	95% CI for difference
Equal	–5.81	58	0.000	1.451	(–11.338, –5.529)
Unequal	–5.81	48.04	0.000	1.451	(–11.35, –5.516)

(c) Mann–Whitney U–Wilcoxon Rank Sum W test

Mean rank	Cases		
27.80	30	Group = 1 Group A	
33.20	30	Group = 2 Group B	
	—		
	60 Total		
U	*W*	*Z*	Two-Tailed *P*

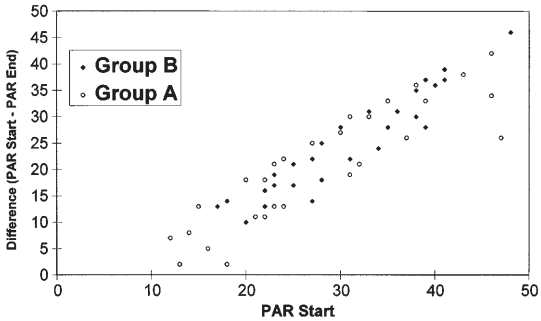

Figure 1 Relationship between initial PAR score and PAR score reduction for groups A and B ($R^2 = 0.92$).

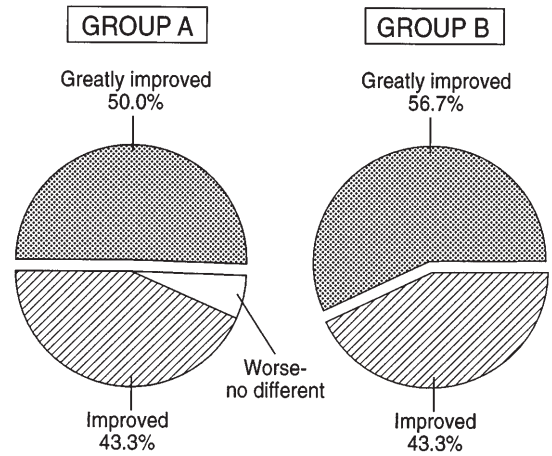

Figure 2 Pie charts showing the reduction in PAR scores for group A (one operator only) and group B (two operators).

likely that this has allowed a consistently high standard of treatment result in both cases.

Conclusions

1. Change of operator contributes significantly to a lengthening of treatment times in fixed orthodontic appliance therapy, in this study by an average of 8.43 months. Thus, patients who for some reason are treated by more than one postgraduate student are compromised with respect to treatment duration.
2. No significant differences were found in the standard of orthodontic treatment results between the two groups of patients, as assessed by the PAR (Peer Assessment Rating) Index. However, according to the conventions of use of the PAR Index, 6.7 per cent of patients in group A fell into the 'worse-no different' category, compared with 0 per cent in group B.
3. No correlation appears to exist between the length of treatment time and the final standard of orthodontic treatment result.
4. The study suggests that the overall reduction in PAR score during treatment can be predicted from the initial PAR score. This observation may be a function of the use of the same fixed appliance system for every patient included in the study.
5. The study highlights the fact that patients who are treated by orthodontic postgraduates should, as far as possible, be completed by that postgraduate and not transferred to another operator.

Address for correspondence

Mr N. McGuinness
Merlin Park Regional Hospital
Dublin Road
Galway
Ireland

Acknowledgements

I would like to thank Miss Helen Knight, Consultant Orthodontist, for allowing me access to her department and study models; Professor C. D. Stephens, for his advice and constructive comments during the writing of this paper; and Mr James Boyd, Medical Statistician, Victoria Hospital, Kirkcaldy, for his assistance with the statistical analysis.

References

DeGuzman L, Bahiraei D, Vig K W, Vig P S, Weyant R J, O'Brien K 1995 The validation of the Peer Assessment Rating index for malocclusion and treatment difficulty. American Journal of Orthodontics and Dentofacial Orthopedics 107: 172–176

Fink D F, Smith R J 1992 The duration of orthodontic treatment. American Journal of Orthodontics and Dentofacial Orthopedics 102: 45–51

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

Kelly B M 1994 An investigation into the treatment times and treatment standards in specialist orthodontic practice in the General Dental Service. MSc Thesis, University of London

Kelly B M, Springate S D 1996 Specialist orthodontics in the General Dental Service. British Dental Journal 180: 209–215

Kerr W J S, Buchanan I B, McNair F I, McColl J H 1994 Factors influencing the outcome and duration of removable appliance treatment. European Journal of Orthodontics 16: 181–186

McGuinness N J, Stephens C D 1993 Holograms and study models assessed by the PAR (Peer Assessment Rating) index of malocclusion—a pilot study. British Journal of Orthodontics 20: 123–129

McGuinness N J, Stephens C D 1994 An introduction to indices of malocclusion. Dental Update 21: 140–144.

O'Brien K, Shaw W C, Roberts C T 1993 The use of occlusal indices in assessing the provision of orthodontic treatment by the hospital orthodontic service of England and Wales. British Journal of Orthodontics 20: 25–35

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

Richmond S, Andrews M 1993 Orthodontic treatment standards in Norway. European Journal of Orthodontics 15: 7–15

Richmond S, Shaw, W C, Stephens C D, Webb W G, Roberts C T, Andrews M 1993 Orthodontics in the general dental services of England and Wales: a critical assessment of standards. British Dental Journal 174: 315–329

Shaw W C, Richmond S, O'Brien K, Brook P, Stephens C D 1991 Quality control in orthodontics: indices of orthodontic treatment need and treatment standards. British Dental Journal 170: 107–112