

SCIENTIFIC
SECTION

Measuring failure of orthodontic treatment: a comparison of outcome indicators

N. A. Fox and J. R. Chapple

Orthodontic Department, James Cook University Hospital, Middlesborough, UK

Objective: To compare treatment failure rates on a sample of completed and discontinued orthodontic treatment cases as measured by PAR, IOTN and ICON to determine whether the use of a sole index would suffice.

Subjects and method: All patients completing or discontinuing orthodontic treatment in the hospital orthodontic departments in the Northern Region during two calendar months were identified and pre and post-treatment models were analysed. PAR scores, IOTN (DHC and AC) and ICON were recorded.

Results: One-hundred-and-forty-five cases were identified and 15 had incomplete records on the day of analysis. The final sample was therefore 130. The overall treatment discontinuation rate of these 130 patients was 24.6%. The treatment failure rate with respect to occlusal improvement varied from 3.1% when measured by PAR and 10.0% using ICON. With respect to residual need it varied from 0.77 to 20.1% with respect to IOTN depending on the criteria examined. The residual treatment need with respect to ICON was 17.2%.

Conclusions: Different occlusal indices give differing failure rates when used on the same patients with ICON being the most critical index. We felt that ICON was the most valid with respect to identifying treatment failure. Its use would enable international comparison of results.

Key words: Occlusal indices, ICON, treatment outcomes

Received 10th February 2004; accepted 8th April 2004

Introduction

It is unreasonable to expect that any medical or dental intervention will always be 100% successful. Indeed, percentage failure rates for various procedures carried out in the National Health Service (NHS) form part of the assessment of NHS Trust hospitals. In orthodontics, various outcome indicators have been in use for some time. The PAR index¹ has been used for around 15 years to measure occlusal improvement. IOTN has been used to measure need with respect to dental health and aesthetics.² The Index of Complexity, Outcome and Need (ICON) has been developed more recently.³ This index purports to measure need, outcome, complexity and degree of improvement. It is quicker to apply than Peer Assessment Rating (PAR) and Index of Treatment Need (IOTN), and it has been previously suggested that it could replace these indices when measuring treatment outcomes, although ICON appeared to require a more stringent standard than PAR.⁴

The null hypothesis is that PAR, IOTN and ICON will give the same results with respect to treatment

outcomes and residual treatment need with respect to the analysis of study models.

The aim of this study was to compare treatment failure rates and residual treatment need on a sample of completed and discontinued orthodontic treatment cases as measured by PAR, IOTN and ICON to determine whether the use of a sole index would suffice.

Materials and method

All orthodontic consultants in the Northern Region were contacted and agreed to take part in the study. All patients who either completed or discontinued active orthodontic treatment within the hospital service during the calendar months of May and June 2003 were identified. The definition of completed treatment was that all treatment aims had been achieved and active appliance treatment was stopped, and retainers if appropriate were fitted. With respect to functional appliance treatment, completed treatment was defined as when overjet reduction was complete. If functional

appliance patients were going on to fixed appliance treatment, this was not regarded as completed treatment and they were not included in the study. Discontinued treatment was defined as early termination of active treatment for any reason whatsoever with failure to complete all the aims of treatment. Final study models were taken for discontinued treatment wherever possible. The models were analysed by a calibrated examiner who visited the units in Carlisle, Newcastle, Sunderland and Middlesborough.

The study models were then analysed and the following recorded for both pre- and post-treatment cases:

- IOTN (DHC);
- IOTN (AC);
- PAR;
- ICON.

Statistical analysis

Repeat scoring of 20 cases chosen at random by computer checked intra-examiner reliability. The reliability of PAR and ICON scoring was checked via a comparison of their respective means and variances. IOTN scores were compared using the unweighted Kappa statistic.

Results

All consultants in the region submitted cases for analysis. The records of 145 cases were submitted, of these, 58 were male (40%) and 87 were female (60%). The mean age of the sample was 17.1 years (SD 4.2 years), 107 cases were completed and 38 were discontinued resulting in a discontinuation rate of 26.2%. The mean treatment time was 23.3 months for completed cases (standard deviation 10.2). The mean time before treatment was discontinued was 20.0 months (SD 9.0).

At the time we visited the units to examine the study models, 130 cases (89.7%) had both pre- and post-treatment models available. Of the 15 cases where either the pre- or post-treatment models were missing of which 9 were completed and 6 were discontinued. These were excluded from the study and analysis limited to the 130 cases with complete records. The discontinuation rate of the 130 records examined was 24.6%.

The results of the intra examiner reliability analysis revealed unweighted kappa values of 0.87, 1.0 respectively for IOTN (AC), IOTN (DHC) representing a high level of agreement. The results of the reproducibility data for ICON and PAR are shown in Table 1. The results show almost identical means for both examinations. In addition, the variances are almost identical. The variance of the difference between the two examinations was less than 10% of the variance of the initial measurement for ICON and less than 1% for PAR. This shows that ICON was less reproducible than PAR, but it was still within acceptable limits. The mean differences for PAR and ICON were 1.8 and 3.5, respectively.

Figure 1 shows the pre- and post-treatment IOTN (DHC) scores. Figure 2 shows the pre- and post-treatment IOTN (AC) scores. Whilst IOTN was designed as an indicator of treatment need and not outcome, one can argue that IOTN (AC) is used informally by the patient as a measure of success and, therefore, the results are included. Figure 3 shows the degree of improvement as measured by ICON. This is calculated by taking the pretreatment ICON score and subtracting four times the post-treatment ICON score as detailed elsewhere.³ This gives a scale of the degree of occlusal improvement achieved. Figure 4 shows the percentage improvement with respect to PAR. To be described as 'improved', an improvement of 30% should be achieved.¹

Table 1 Results of the reproducibility test for PAR and ICON

	N	Mean	SE	SD	Variance
Start ICON (1st examination)	20	89.10	2.99	13.37	178.73
Start ICON (2nd examination)	20	88.00	2.96	13.23	174.95
Difference	20	3.5	0.78	3.47	12.05
Finish ICON (1st examination)	20	25.65	2.75	12.32	151.71
Finish ICON (2nd examination)	20	25.50	2.39	10.70	114.47
Difference	20	1.85	0.74	3.31	10.98
Start PAR (1st examination)	20	35.90	2.52	11.25	126.62
Start PAR (2nd examination)	20	35.45	2.48	11.10	123.10
Difference	20	1.35	0.21	0.93	0.87
Finish PAR (1st examination)	20	6.60	1.27	5.70	32.46
Finish PAR (2nd examination)	20	6.8	1.24	5.54	30.69
Difference	20	0.80	0.12	0.52	0.27

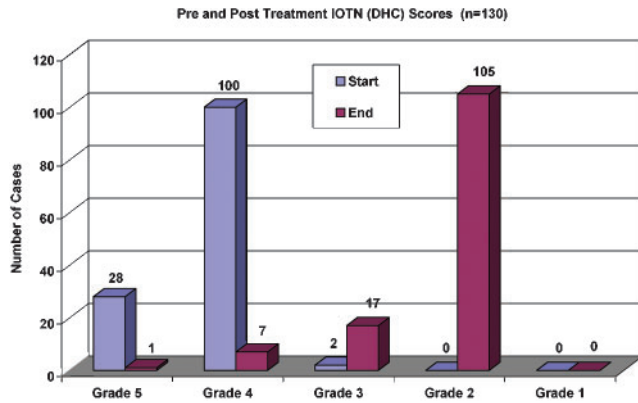


Figure 1 A plot of pre- and post-treatment IOTN (DHC) scores

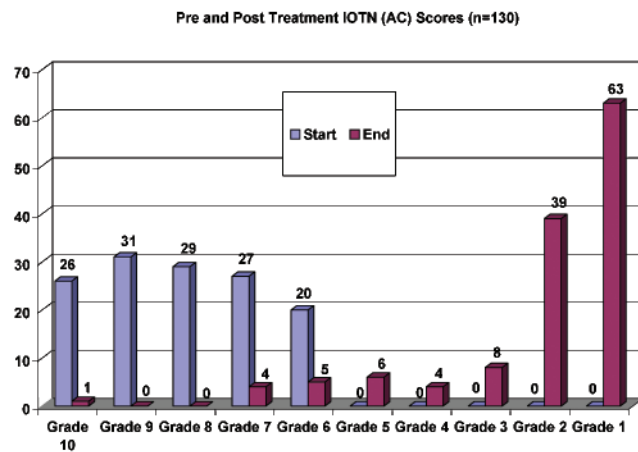


Figure 2 A plot of pre- and post-treatment IOTN (AC) scores

Table 2 shows the failure rates with respect to residual treatment need as measured by IOTN and ICON. It can be seen that there is a wide variation in failure rate depending on the criteria examined but ICON is comparable with IOTN if residual borderline need is regarded as failure. Table 3 shows the treatment failure rate with respect to occlusal improvement. Once again, there is a difference with ICON demanding a more stringent improvement.

Table 2 Failure rates of treatment with respect to residual definite treatment need (n=130)

Criteria examined failure rate	Number of cases	Percentage
IOTN (AC) > 7 at end of treatment	1	0.77%
IOTN (AC) >4<=7	17	13.1%
IOTN (DHC) > 3 at end of treatment	8	6.2%
IOTN (DHC) >2 at end of treatment	27	20.1%
Post-treatment ICON >43	23	17.2%

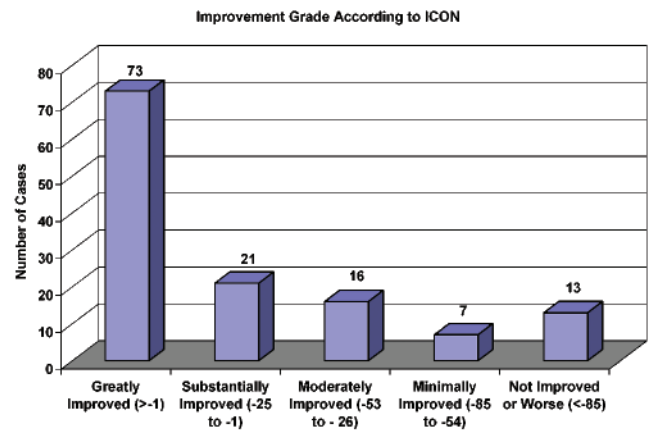


Figure 3 A plot to show the degree of improvement as measured by ICON. This is calculated by taking the pretreatment ICON score and subtracting four times the post-treatment ICON which are the values quoted on the x-axis

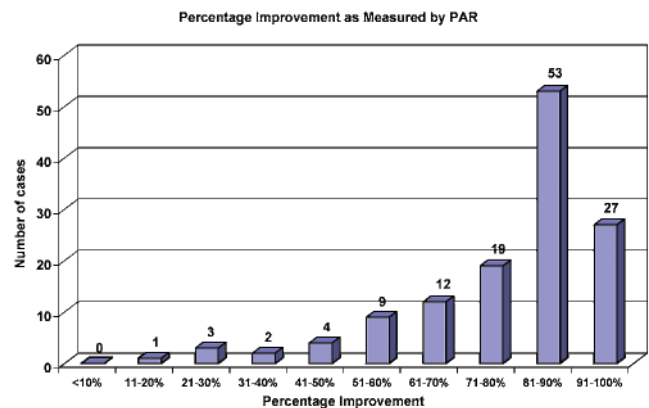


Figure 4 A plot showing the percentage improvement with respect to PAR. To be described as 'improved', an improvement of 30% or more should be achieved.

Discussion

It has previously been suggested that ICON may be able to replace PAR and IOTN as an indicator of outcome and need.⁴ It is true that ICON can be used to measure outcome, need and complexity from a single index, and

Table 3 Failure rates of treatment with respect to occlusal improvement (n=130)

Criteria examined	Number of cases	Percentage
<30% reduction in PAR score worse/no different)	4	3.1%
Pretreatment ICON minus 4 times post-treatment ICON <-85 (not improved or worse)	13	10.0%

our results suggest that, with respect to treatment need, if borderline need (IOTN DHC grade 3 and IOTN AC Grade 5-7) is included, ICON can indeed replace IOTN.

With respect to PAR, there is more of a difference. PAR evolved as a peer assessment tool validated in the UK by dentists of varying orthodontic experience when British GDS orthodontics still involved a large amount of removable appliance treatment only. PAR does not include measurement of residual spacing in the buccal segments. With the development of specialist orthodontic practices and a shift to fixed appliances, we feel it may be time for a change. One drawback of PAR is that different countries use different weightings to suit their own validations. International comparison of results is therefore difficult. An international panel of specialist orthodontists validated ICON and we feel this strengthens the index considerably. Certainly in our own subjective assessment of treatment success, we felt the failure rate identified by ICON (10%) was fair, but 3.1% suggested by PAR was too lenient. We also noticed that when scoring the models, ICON was much quicker to apply than PAR and IOTN.

One factor that this study did highlight was the widely differing rates of treatment failure implied by the different indices particularly when different criteria were applied. There needs to be some agreement within the orthodontic profession about this and should obviously be considered carefully when comparing results between studies.

Conclusions

1. ICON can be used as a substitute for IOTN to identify residual need in patients after orthodontic treatment if borderline need is included.
2. With respect to occlusal improvement, ICON is harsher than PAR, but its use is appropriate in identifying treatment failure. Its use would enable international comparison of results. It is quicker to use on models than PAR and IOTN and the data produced substitutes for these indices.
3. ICON is significantly quicker to apply to models than PAR.

Acknowledgements

We would like to thank Nigel Carter, Alison Downing, Peter Gordon, Ross Hobson, Rye Mattick, Lyn May, Dennis McConnell, Ian Shaw, Ron Tyrell and Steve Whitehurst, our consultant colleagues in the Northern Region for taking part in this study and allowing us to examine the records of patients under their care. We would like to thank Peter Gordon for his statistical advice

Authors and contributors

Nigel Fox was responsible for study design, analysis of the study models, initial data analysis and final approval of the article. Jonathan Chapple was responsible for the literature search and contributed to the writing of the article, including analysis and data interpretation, revisions and final approval of the version to be published. The guarantor is Nigel Fox.

References

1. Richmond S, Shaw WC, O'Brien KD, Buchanan IB, Jones R, Stephens CD, Roberts CT, Andrews M. The development of the PAR Index (Peer Assessment Rating): reliability and validity. *Eur J Orthod* 1992; **14**: 125–39.
2. Brook PH, Shaw WC. The development of an index of orthodontic treatment priority. *Eur J Orthod* 1989; **11**: 309–20.
3. Daniels C, Richmond S. The development of the index of complexity, outcome and need (ICON). *J Orthod* 2000; **27**: 149–62.
4. Fox N, Daniels C, Gilgrass T. A comparison of the Index of Complexity Outcome and Need (ICON) with the Peer Assessment Rating (PAR) and the Index of Orthodontic Treatment Need (IOTN). *Br Dent J* 2002; **193**: 225–30.