

Comparison of three indices in evaluation of orthodontic treatment outcome

Kurt Bergström and Arne Halling

Departments of Orthodontics and Pediatric Dentistry, School of Dentistry, Karolinska Institute, Huddinge; Centre for Public Health Sciences, Linköping University, Linköping; and Orthodontic Department of the Public Dental Health Service, Växjö; Sweden

Bergström K, Halling A. Comparison of three indices in evaluation of orthodontic treatment outcome. *Acta Odontol Scand* 1997;55:36–43. Oslo. ISSN 0001–6357.

Three indices measuring outcome of orthodontic treatment were evaluated and compared in 39 specialist-treated individuals (group 1) and in 20 specialist-treated individuals who had undergone orthognathic surgery (group 2). The outcome was evaluated in accordance with the modified Indication Index, the modified ISMHB (Index of the Swedish Medical Health Board), and the PAR (Peer Assessment Rating) Index. Pre- and post-treatment need scores on the basis of the three indices were estimated. The ranking order of the individuals by the different indices was compared in the two groups separately, before and after treatment. Associations between the Indication Index and the ISMHB were consistently stronger than associations between the PAR Index and the two other indices, for which the variations were great. The reasons for the great variations seemed to be the different bases for assessment used and the different constructions of the indices: the PAR Index estimates deviations from an ideal occlusion, and the other indices evaluate treatment need. The outcome described by the indices differed with regard to the pre- and post-treatment scores and the ranking order of the individuals. Of the three indices, the Indication Index was found to be the one most suitable for estimating treatment outcome in terms of decrease of treatment need, and the PAR Index was unsuitable for estimating treatment need. □ *Dental occlusion; orthodontics; orthognathic surgery; outcome measures; treatment need*

Kurt Bergström, Orthodontic Department of the Public Dental Health Service, Klostergatan 16 B, S-352 31 Växjö, Sweden

Occlusal indices may be defined as methods of determining the level of treatment need or the amount of deviation from a normal occlusion. They can be used both for the evaluation of individual patients and populations and as indicators of the clinical outcome of orthodontic treatment. Economic restraints in recent years have resulted in an increasing interest in the evaluation of orthodontic treatment outcome, while quality assurance has become more important.

Evaluation of orthodontic treatment results (= outcome) has been described by Eismann (1) and Gottlieb (2). They limited their evaluations to measurements of plaster casts. Richmond et al. (3, 4) introduced the Peer Assessment Rating (PAR) Index, designed for the evaluation of treatment outcome. The amount of deviation from a normal occlusion is scored. The PAR Index is also based on measurements on study casts or on centric occlusion in the mouth. Berg (5) used a wider base, in which even photographs and radiographs were considered, and the treatment results were assessed on the basis of defined criteria.

Another way of estimating treatment outcome is to measure treatment need before and after treatment. Several indices for estimating the need for treatment have been introduced over the years (6). Grainger (7) introduced the orthodontic Treatment Priority Index (TPI), and Salzmann (8) presented the Handicapping Malocclusion Assessment Record (HMAR). Summers (9) created the Occlusal Index (OI), and Cons et al. (10)

the Dental Aesthetic Index (DAI). The OI has been used by several authors to assess the treatment outcome (11–14). Brook & Shaw (15) introduced the Index of Orthodontic Treatment Need (IOTN), which has been used to assess the need for orthodontic treatment before and after appliance therapy (16). The Dental Health Component of this index is similar to the Swedish ISMHB, an index created by the Swedish Medical Health Board (17). Another index used in Sweden is the Indication Index developed by Lundström (18).

In Sweden the outcome of orthodontic treatments in three counties was estimated in accordance with the Indication Index (19). The treatments were also estimated on the basis of the ISMHB (17), whereby similar results were found. It would be valuable to compare the results in this study with other studies, especially when the PAR Index has been used. This index seems to have gained in popularity, and several studies have been published in which this index has been used (20–26). To what extent is such a comparison possible?

Comparisons of different indices have not often been made. Which qualities are most important? Are indices describing treatment need or indices describing deviations from ideal tooth relationships to be preferred? Are indices with smaller bases and higher reproducibility sufficient, or are indices with wider bases more valid?

Our aim was to compare three indices in the evaluation of orthodontic treatment outcome with regard to the pre- and post-treatment scores and the

Table 1. Summary of characteristics of the different indices and factors on which they are based

	Indication Index	ISMHB	PAR Index
Aim of the index			
To describe treatment need	+	+	-
To describe deviations from an ideal occlusion	-	-	+
Basis for assessment			
Ability to consider expected disadvantages	+	-	-
Study casts			
Measurements of deviations	-	-	+
Evaluation of malocclusions	+	+	-
Photographs of facial appearance			
Evaluation of profile	+	-	-
Evaluation of soft tissues	+	-	-
Evaluation of lip closure	+	-	-
Case histories			
Notes about forced bites	+	+	-
Notes about lip dysfunction	+	+	-

ranking order of the individuals on the basis of the indices and to establish recommendations for clinical use.

Materials and methods

From two samples of 19-year-olds (337 and 302 individuals) in the counties of Stockholm and Kronoberg (27), 39 regular patients in specialist care were selected (group 1). From a sample of 48 individuals who had undergone surgical (orthognathic) treatment, another 20 cases were selected (group 2) (mean age, 27.9 years; range, 16–44 years). They were regular patients in specialist care in Stockholm, where the orthognathic treatment had been performed. All individuals satisfying the following inclusion criteria were selected for the

study: a well-documented case history and pre- and post-treatment study casts and photographs were available. The material in the two groups consisted of a total of 59 well-documented, orthodontically treated individuals.

The three indices used were a modified Indication Index (28), a modified 5-grade ISMHB index, as stated by the Swedish Medical Health Board (17), and a PAR Index (3). Estimation of treatment need in accordance with the Indication Index and ISMHB was made separately by two orthodontic specialists (K. Bergström and L. Johansson), who both have more than 20 years of clinical experience. When the estimated score differed between the two examiners (which occurred in 14% of the judgements), a new estimation was made by consensus. For each of the 59 cases the two indices were estimated on the same occasion. The PAR Index was calculated after measurements on study casts by an orthodontic specialist (L. Johansson) in accordance with the guidelines of the authors.

The characteristics of the three indices and the factors on which they are based are summarized in Table 1.

A summary of the guidelines for evaluating treatment need and deviations from an ideal occlusion

Indication Index. A wide base for assessment is used (see Table 1). The disadvantages of the different malocclusions are evaluated—such as health hazards or conditions detrimental to the well-being of the individual—and even expected disadvantages are taken into account. Aesthetic, prophylactic, and functional—actual or expected—disadvantages of malocclusions are estimated and marked in one of three columns for very great, great, or moderate/uncertain disadvantage (Fig. 1). The marks in the different columns are summed up, and the total score is expressed as a three-digit

Treatment indications		Disadvantage (actual or expected)		
		very great	great	moderate / uncertain
Aesthetic:	Aesthetic/Psychosocial			X
	Plaque retention			X
	Gingival impingement			X
Prophylactic:	Cranio-mandibular disorders			
	Root resorption			
	Traumatic injury			
Functional:	Mastication deficiency			
	Speech defect			
	Mouth-breathing			
Indication Index:		0	0	3

Fig. 1. An example of an estimation using the Indication Index (28). The actual or expected disadvantage for each component is either marked with a cross in one of the three columns from left to right, 'very great', 'great', or 'moderate/uncertain', or left blank if the estimated disadvantage is judged to be little or none. The crosses in each column are summed, and the Indication Index will be 003, which is converted to 2 on the basis of the global scale (Table 2).

Table 2. Global scale for conversion of Indication Index. On the basis of the number of markings in the different columns the treatment need is expressed as a number between 0 and 10

Scale	Very great treatment need	Great treatment need	Moderate/uncertain treatment need
0	0	0	0
1	0	0	1–2
2	0	0	≥3
3	0	1	0–2
4	0	1	≥3
5	0	2	0–2
6	0	2	≥3
7	0	≥3	Any
8	1	0–2	Any
9	1	≥3	Any
10	≥2	Any	Any

number. The Indication Index is converted to a global scale (Table 2). The treatment need is expressed as a number between 0 and 10.

Modified Index of the Swedish Medical Health Board, ISMHB. The treatment need of a malocclusion is, with the given examples as guidelines, assigned to one of five categories as no, little, moderate, urgent, or very urgent need (Table 3). The treatment need is expressed as a number between 0 and 4.

PAR Index. Displacements of teeth are recorded as the distance between contact points of adjacent teeth. The greater the displacement, the greater the PAR score (see Table 4). Buccal occlusion is recorded with regard to the three planes of space. Overjet is recorded, positive as well as negative, and overbite and center line discrepancies are recorded. The individual scores are summed to obtain an overall total, representing the degree of deviation from normal alignment and occlusion. A score of 0 indicates good alignment, and higher scores indicate increased levels of irregularity. The components of the PAR Index have been weighted to reflect current British orthodontic opinion more closely (3). The PAR Index is expressed as a number, rarely beyond 50.

Pre- and post-treatment need scores according to the three indices were estimated in all 59 cases. The associations between the indices with regard to the

Table 4. PAR Index form (3) and an example of recordings. Scores are applied to each feature of the malocclusion in accordance with special guidelines. The unweighted score is then multiplied by a weighting factor

	Unweighted score	Weighting factor	Weighted score
Upper anterior segment alignment	1	×1	1
Lower anterior segment alignment		×1	
Right buccal occlusion	1	×1	1
Left buccal occlusion		×1	
Overjet	3	×6	18
Overbite	1	×2	2
Center line	2	×4	8
		Total	30

ranking order of the individuals were compared in the two groups separately, before and after treatment.

Reliability/statistics

After 2 months the cases were again estimated in the same manner on the basis of the Indication Index and ISMHB, and recordings in accordance with the PAR Index were made. The reliability was calculated, using the gamma statistic G (29). The Spearman rank-order correlation coefficient r_s was used to compare the rankings of individuals by different indices. A probability at the 5% level or less was considered statistically significant.

Results

The correlations or degree of agreement (gamma statistic G) between the two estimations were, before treatment, 0.98 for the Indication Index, 0.94 for the ISMHB, and 0.98 for the recordings of the PAR Index and, after treatment, 0.89 for the Indication Index, 0.89 for the ISMHB, and 0.98 for the recordings of the PAR Index.

According to all three indices used, the individuals in

Table 3. A modified 5-grade index (ISMHB) for orthodontic treatment need, stated by the Swedish Medical Health Board (17)

4	Very urgent need. Aesthetically and/or functionally handicapping anomalies, such as cleft lip and palate, extreme postnormal or prenormal occlusion, retained upper incisors, extensive aplasia.
3	Urgent need, such as prenormal forced bite, deep bite with gingival irritation not only on papilla incisiva, large overjet with lower lip behind upper centrals, extremely open bite, cross-bite causing transverse forced bite, scissors bite interfering with articulation, severe frontal crowding or spacing, retained canines, aesthetically and/or functionally disturbing rotations.
2	Moderate need, such as aesthetically and/or functionally disturbing proclined or retroclined incisors, deep bite with gingival contact but without gingival irritation, severe crowding or spacing, infraocclusion of deciduous molars and permanent teeth, moderate frontal rotations.
1	Little need. Mild deviations from normal (ideal) occlusion, such as prenormal occlusion with little negative overjet, postnormal occlusion without other anomalies, deep bite without gingival contact, open bite with little frontal opening, cross-bite without forced bite, mild crowding or spacing, mild rotations of only little aesthetic and/or functional significance.
0	No need. Normal (ideal) occlusion without deviations.

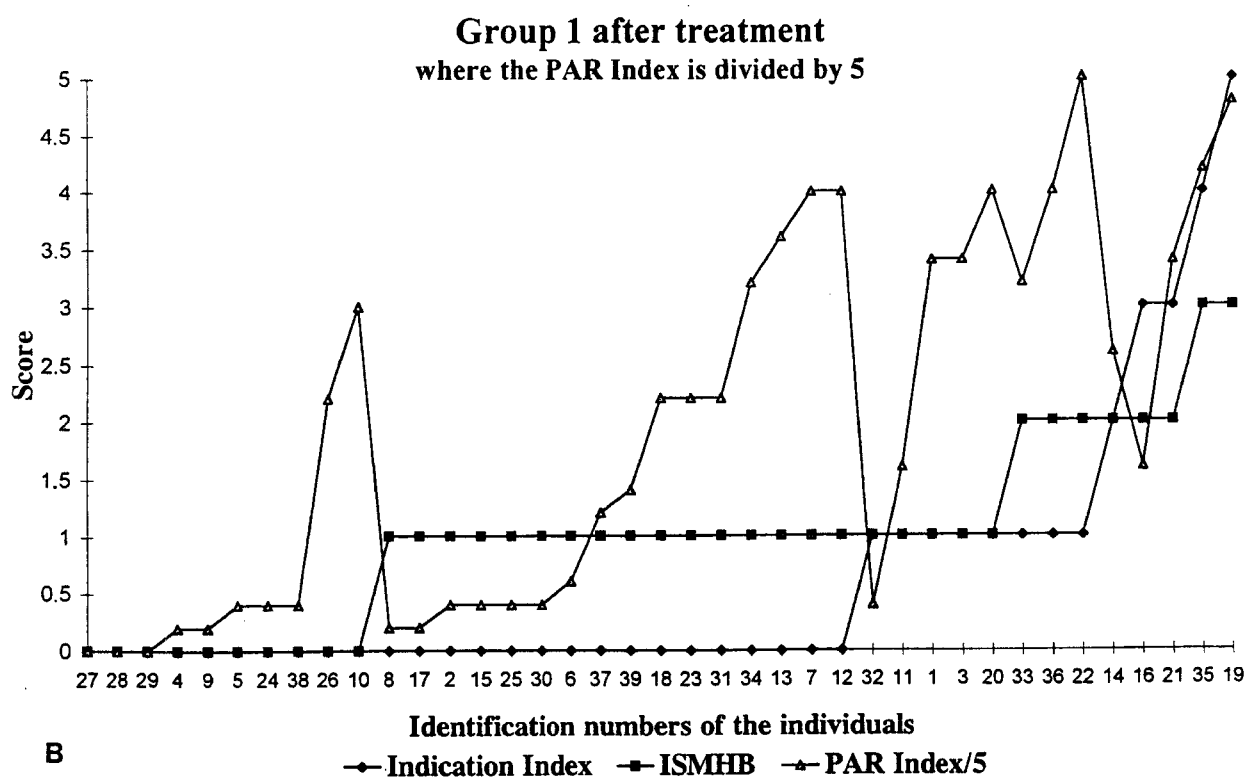
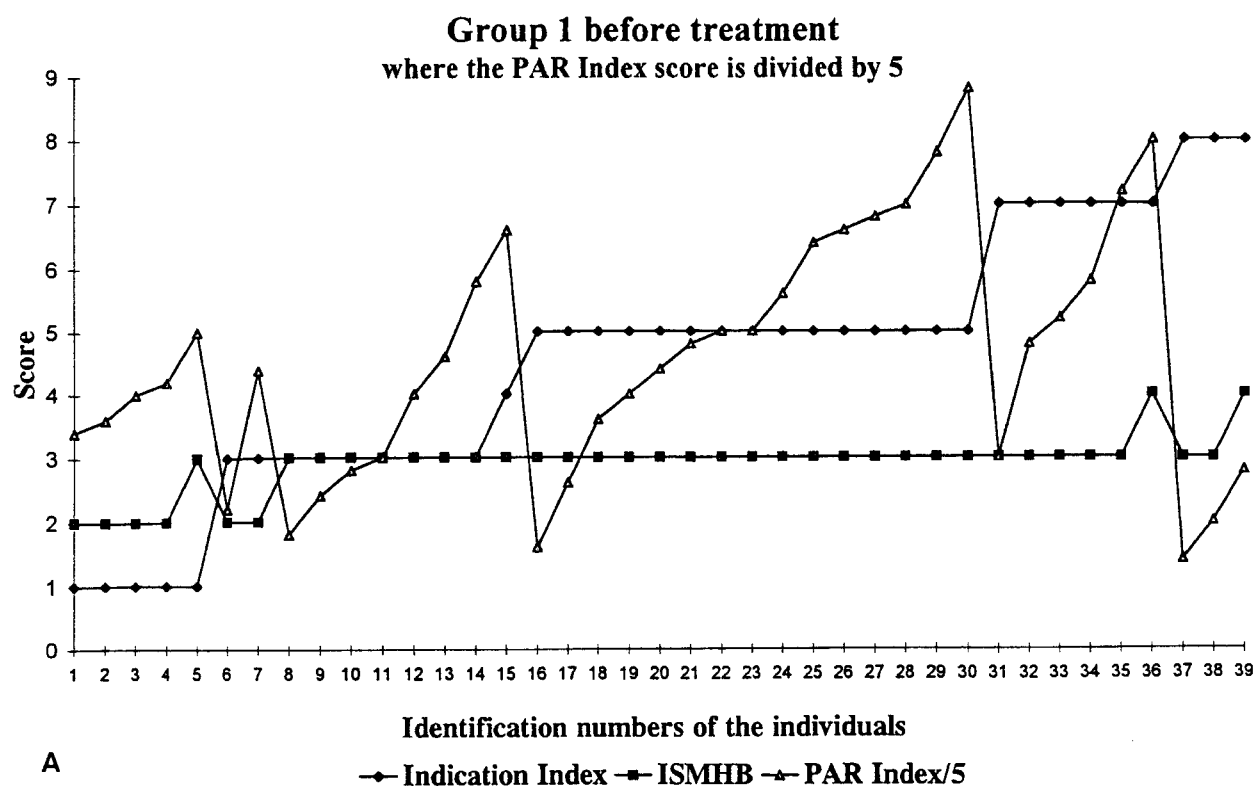
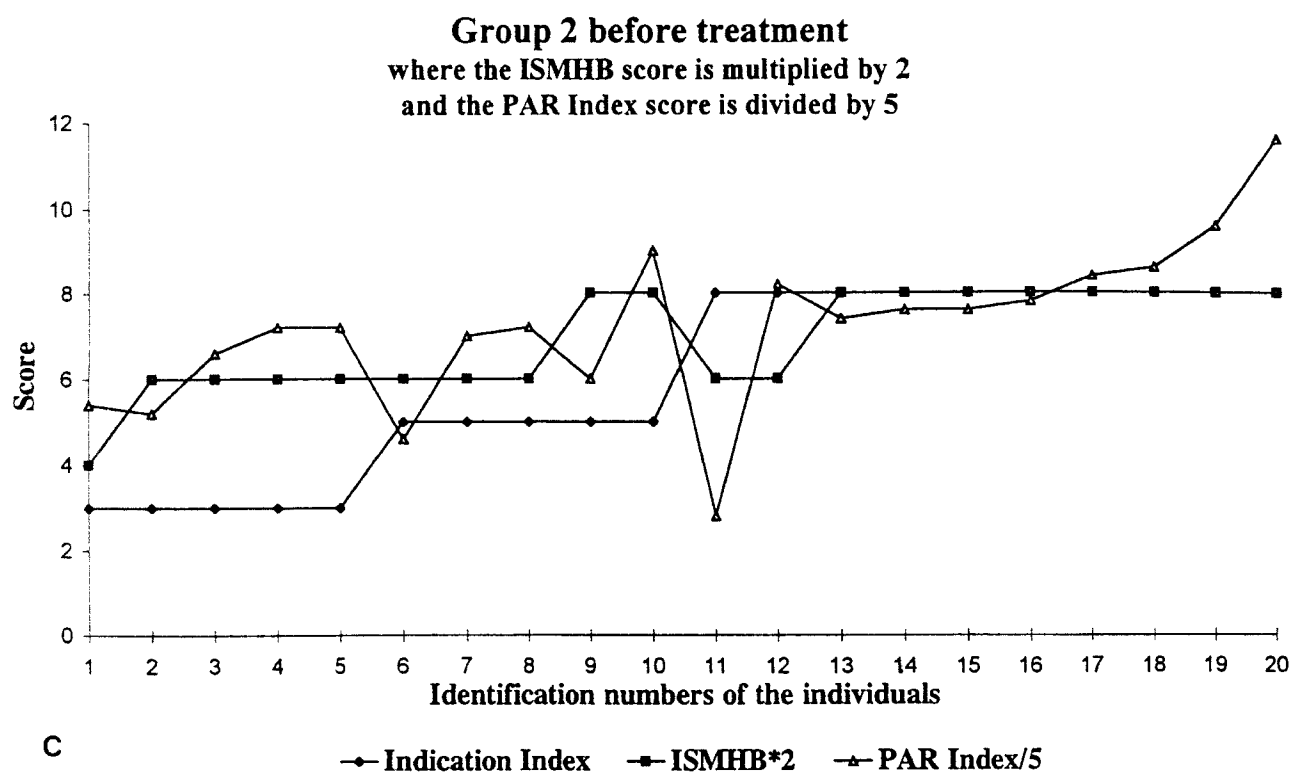
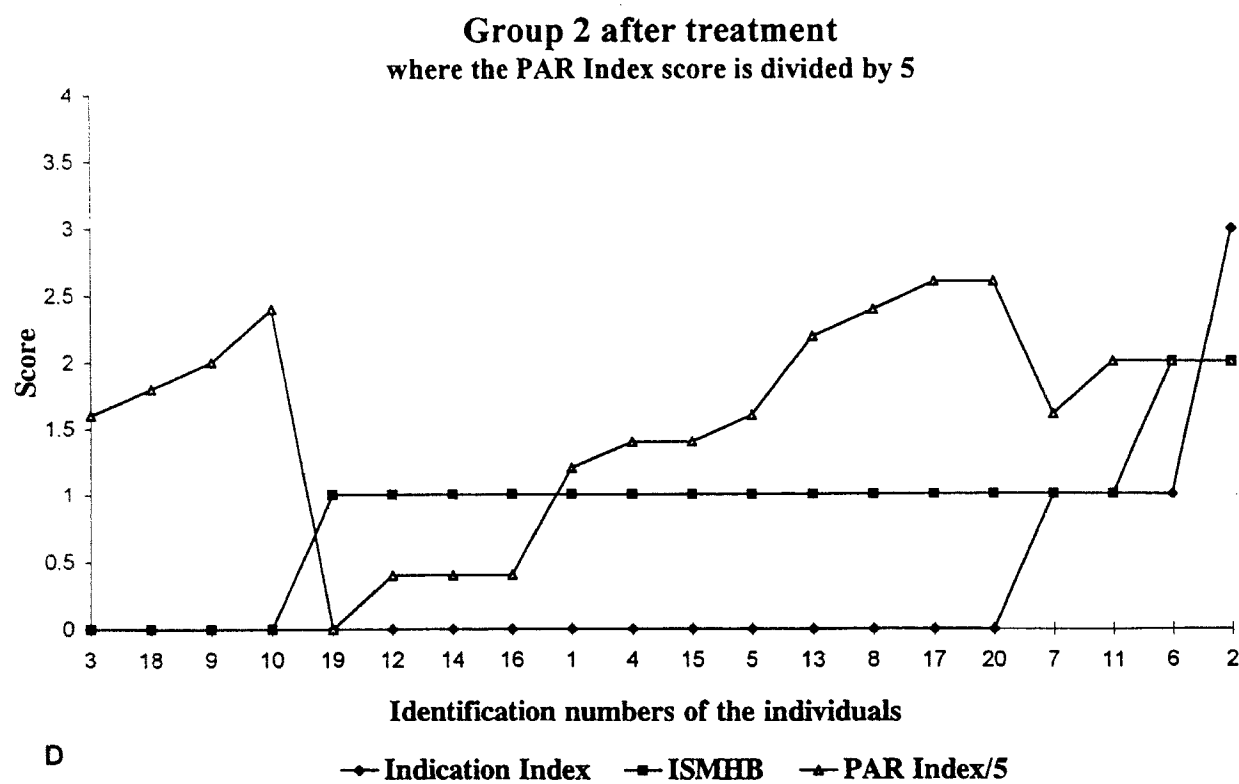


Fig. 2A and B. Individuals ranked in order by 1) the Indication Index, 2) the ISMHB, and 3) the PAR Index. The scores are multiplied with a suitable constant to obtain overlapping to a greater extent. 2A. Group 1 ($n = 39$) before treatment. 2B. Group 1 after treatment.



C



D

Fig. 2C and D. 2C. Group 2 ($n = 20$) before treatment. 2D. Group 2 after treatment.

Table 5. Pre- and post-treatment associations between the different indices in the two groups by the Spearman rank-order correlation coefficient, r_s

	Group 1, $n = 39$		Group 2, $n = 20$	
	Before treatment	After treatment	Before treatment	After treatment
Indication Index/ISMHB	0.64	0.73	0.68	0.56
Indication Index/ PAR Index	0.16	0.57	0.58	0.17
ISMHB/ PAR Index	0.22	0.65	0.70	-0.06

group 2 had on an average the most severe anomalies before treatment. The scores after treatment were in general considerably lower, and especially according to the Indication Index in group 2 (Fig. 2A-D).

The associations between the different indices in estimating the two groups selected, before and after treatment, are shown in Table 5, in which the Spearman rank-order correlation coefficient, r_s , is surveyed, and in Fig. 2A-D, in which the individuals are ranked in order by 1) the Indication Index, 2) the ISMHB, and 3) the PAR Index. The associations between the Indication Index and the ISMHB were consistently stronger than between the PAR Index and the two other indices, especially before treatment in group 1 and after treatment in group 2. An Indication Index score of 0, which in principle means no or little treatment need, corresponds, for example, to PAR Index scores of up to 20 (Fig. 2B).

Discussion

The reliability of the estimations with the Indication Index and the ISMHB was high. According to the guidelines of Landis & Koch (30), values in this range may be interpreted as showing almost perfect reliability. Malmgren (31) found that the level of agreement between judgments was higher after joint training. The two examiners in this study both had long clinical experience and a great deal of joint training. They were also very familiar with both the Indication Index and the ISMHB. These circumstances may explain the high level of agreement between the two estimations. The reliability of the PAR Index was, as expected, very high, as the measurements are easy to perform and reproducible to a high extent. The reliability of the estimations in this study may thus be considered good.

The validity (the extent to which an index measures what it was designed for) of the evaluations with the different indices is, however, more difficult to determine (32). The assessment base should basically be wide enough to enable considerations important for the case to be taken into account. This is a component of vital importance for the validity. If an index has a narrow

base, essential factors may be disregarded in the assessment. The Indication Index has the widest base for assessment of the three indices compared (Table 1).

The individuals in the study were all regular patients in specialist care, but surgical (orthognathic) treatment had also been performed in group 2. The anomalies to be estimated represented different grades of severity, from the most severe anomalies in group 2 before treatment to the least severe in group 2 after treatment.

There was a great variation in the association, r_s , between the indices: -0.06 was the weakest and 0.73 the strongest association. This implies that one index will in part select different individuals for treatment than another. This is in accordance with earlier reports (6, 33).

In evaluating treatment outcome, the different indices disagree to the same extent. The individuals are ranked in order by the indices, before and after treatment, and there is disagreement. The association between the Indication Index and the ISMHB was fairly strong in both groups. This may be because both indices are designed to measure treatment need. They were also both very familiar to the examiners. The differences existing between the indices may be explained by the possibilities to evaluate soft tissues and future disadvantages with the Indication Index and not with the ISMHB.

The ISMHB is a fairly insensitive measure (34). Most cases in group 1 were given a score of 3 before treatment, and nearly all the estimations in group 2 after treatment resulted in a score of 1 (Figs. 2A and D). This means that the ISMHB may not be considered so suitable for selecting patients for treatment or for evaluation of outcome. A wider grading from 0 to 10, used in Finland, was found to make the rating of the cases easier (35).

The associations between the PAR Index and the two Swedish indices varied more in general. The reason may be that the PAR Index was designed to measure deviations of teeth from an ideal position, and the other indices to measure treatment need. Most important, however, the bases of evaluation are different (Table 1). Evaluations using the PAR Index are based exclusively on study casts, and functional and aesthetic considerations of the face are disregarded completely. The Indication Index has the widest base for evaluation. The soft profile, the masking influence of the soft parts, and lip dysfunctions as well as expected normal or abnormal changes can be evaluated and included in this index. Although an evaluation of these essential factors will include a component of subjectivity, the validity might be increased.

The PAR Index is weighted 'to reflect current British orthodontic opinion'. These weightings seem in many cases to give a real effect, and agreement with indices for estimating treatment need are obtained. In other cases, however, the weightings seem to result in scores that are high and unmotivated. An example of this is the

large weighting of overjet ($\times 6$). An overjet > 3 mm yields 6 points, and > 7 mm 18 points. An overjet can be aesthetically very disturbing but is not necessarily so. In some individuals an overjet may be considered becoming, and when the lip profile is concave, it could be unwise to reduce an overjet. In the Indication Index such information is considered, and this most likely decreases the association with the PAR Index. Other deficiencies of the PAR Index have also been reported (21, 25), such as the inability to consider cases with limited treatment objectives, when final treatment must wait.

Desirable qualities of an index are as follows: high validity—that is, the index measures what it was designed for; wide assessment base; high reliability—that is, the evaluations should be reproducible; and easy to learn and use.

When comparing the different indices, it is important to consider the assessment base used, and the design of the indices. The PAR Index is constructed for measuring deviations from ideal tooth arrangement and seems to be unsuitable for measuring treatment need. The ISMHB is a fairly insensitive measure, and is not suitable for estimating treatment outcome. The Indication Index seems to be better than the other two indices and is recommended. Actual or expected disadvantages of the malocclusion are assessed, and a wide base is used. The reliability is good, and the index is easy to learn and use.

Conclusions

Comparison of studies that have used different indices to evaluate treatment outcome should be done with caution. The outcomes described by the different indices may in part differ. To estimate treatment outcome in terms of decrease of treatment need, the Indication Index seems to be better than the other two indices and is recommended.

Acknowledgements.—Dr. Lennart Johansson is thanked for giving time and energy in assessment of the individuals. This study was supported by grants from the Kronoberg County Council.

References

1. Eismann D. A method of evaluating efficiency of orthodontic treatment. *Trans Eur Orthod Soc* 1974;223–32.
2. Gottlieb EL. Grading your orthodontic treatment results. *J Clin Orthod* 1975;IX:156–61.
3. Richmond S, Shaw WC, O'Brien KD, Buchanan IB, Jones R, Stephens CD, et al. The development of the PAR Index (Peer Assessment Rating): reliability and validity. *Eur J Orthod* 1992;14:125–39.
4. Richmond S, Shaw WC, Roberts CT, Andrews M. The PAR Index (Peer Assessment Rating): methods to determine outcome of orthodontic treatment in terms of improvement and standards. *Eur J Orthod* 1992;14:180–7.
5. Berg R. Post-retention analysis of treatment problems and failures in 264 consecutively treated cases. *Eur J Orthod* 1979;1:55–68.
6. So LL, Tang EL. A comparative study using the Occlusal Index and the Index of Orthodontic Treatment Need. *Angle Orthod* 1993;63:57–64.
7. Grainger RM. Orthodontic treatment priority index. Toronto: Faculty of Dentistry, University of Toronto, 1967.
8. Salzmann JA. Handicapping malocclusion assessment to establish treatment priority. *Am J Orthod* 1968;54:749–65.
9. Summers CJ. The Occlusal Index: a system for identifying and scoring occlusal disorders. *Am J Orthod* 1971;59:552–67.
10. Cons NC, Jenny J, Kohout FJ. DAI: the dental aesthetic index. College of Dentistry, University of Iowa, 1986.
11. Pickering EA, Vig P. The Occlusal Index used to assess orthodontic treatment. *Br J Orthod* 1975;2:47–51.
12. Elderton RJ, Clark JD. Orthodontic treatment in general dental service assessed by the Occlusal Index. *Br J Orthod* 1983;10:178–86.
13. Elderton RJ, Clark JD. An investigation of treatment provided in the General Dental Service for patients with class II, division 1 malocclusions. *Br J Orthod* 1984;11:2–8.
14. Tang EL, Wei SHY. Assessing treatment effectiveness of removable and fixed orthodontic appliances with the Occlusal Index. *Am J Orthod Dentofac Orthop* 1990;98:550–6.
15. Brook PH, Shaw WC. The development of an index of orthodontic treatment priority. *Eur J Orthod* 1989;11:309–20.
16. Richmond S, Roberts CT, Andrews M. Use of the Index of Orthodontic Treatment Need (IOTN) in assessing the need for orthodontic treatment pre- and post-appliance therapy. *Br J Orthod* 1994;21:175–84.
17. Swedish Medical Health Board. Kungl. Medicinalstyrelsens cirkulär den 13 december 1967 angående anvisningar för journalföringen inom folktankvårdens tandregleringsvård. Stockholm:MF nr 71, 1967.
18. Lundström A. Need for treatment in cases of malocclusion. *Trans Eur Orthod Soc* 1977;111–23.
19. Bergström K, Halling A. Orthodontic treatment outcome in three Swedish counties with different orthodontic resources. *Swed Dent J* 1996;20:51–60.
20. Richmond S. Personal audit in orthodontics. *Br J Orthod* 1993;20:135–44.
21. Kerr WJS, Buchanan IB, McColl JH. Use of the PAR Index in assessing the effectiveness of removable orthodontic appliances. *Br J Orthod* 1993;20:351–7.
22. O'Brien KD, Shaw WC, Roberts CT. The use of occlusal indices in assessing the provision of orthodontic treatment by the Hospital Orthodontic Service of England and Wales. *Br J Orthod* 1993;20:25–35.
23. Richmond S, Andrews M. Orthodontic treatment standards in Norway. *Eur J Orthod* 1993;15:7–15.
24. Richmond S, Andrews M, Roberts CT. The provision of orthodontic care in the General Dental Services of England and Wales: extraction patterns, treatment duration, appliance types and standards. *Br J Orthod* 1993;20:345–53.
25. Fox NA. The first 100 cases: a personal audit of orthodontic treatment assessed by the PAR (peer assessment rating) Index. *Br Dent J* 1993;174:290–7.
26. John W, Kerr S, Buchanan IB, McNair FI, McColl JH. Factors influencing the outcome and duration of removable appliance treatment. *Eur J Orthod* 1994;16:181–6.
27. Bergström K, Halling A. Orthodontic care delivered by general practitioners and specialists in three Swedish counties with different orthodontic specialist resources. *Swed Dent J* 1996;20:35–50.
28. Medin L, Bergström K, Filipsson R, Hannerz H, Jensen R. Need of orthodontic treatment. A suggestion to modify Indication Index according to Anders Lundström. *Kieferorthopaed Mitt* 1995;9:167–83.
29. Siegel S, Castellan NJ. In: Anker JD, editor. Nonparametric

- statistics for the behavioral sciences. 2nd ed. New York: McGraw-Hill Inc, 1988.
30. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
 31. Malmgren O. Studies on the need and demand for orthodontic treatment [thesis]. *Swed Dent J* 1980 Suppl 6.
 32. Carlos J. Evaluation of different indices of malocclusion. *Int Dent J* 1970;20:606–17.
 33. Järvinen S, Vaätäjä P. Variability in assessment of need for orthodontic treatment when using certain treatment-need indices. *Community Dent Oral Epidemiol* 1987;15:245–8.
 34. Brattström V. Socialstyrelsens behandlingbehovsindex i ortodonti—en kritisk granskning. *Tandlakartidningen* 1987;79:704–7.
 35. Heikinheimo K. Need of orthodontic treatment and prevalence of craniomandibular dysfunction in Finnish children [thesis]. Turku: University of Turku, 1989.

Received for publication 25 March 1996

Accepted 11 June 1996