

The reliability and validity of the Index of Complexity, Outcome and Need for determining treatment need in Dutch orthodontic practice

T. J. Louwerse*, I. H. A. Aartman**, G. J. C. Kramer* and B. Prahl-Andersen*

Departments of *Orthodontics and **Social Dentistry and Behavioural Sciences, Academic Centre for Dentistry Amsterdam (ACTA), The Netherlands

SUMMARY The Index of Complexity, Outcome and Need (ICON), based on international opinion, has been proposed as a multipurpose occlusal index. The aim of this study was to validate the ICON for treatment need in the Netherlands by relating it to Dutch orthodontic opinion. Furthermore, the reliability of this index was explored, for both a calibrated orthodontist and non-calibrated orthodontists. A sample of 102 patients was chosen which represented the actual distribution of severity of malocclusion experienced by orthodontists in every day practice. The ICON was scored, based on complete patients records of those 102 patients, by an examiner calibrated in the use of this index. The results were compared with the opinion about treatment need of seven Dutch orthodontists – the ‘gold standard’. Nine non-calibrated orthodontists also scored the ICON for 49 patients.

The intra-examiner agreement of both the non-calibrated and the calibrated orthodontists was moderate to high [0.52–0.86 and 0.89, respectively, measured with the Intraclass Correlation Coefficient (ICC)]. The inter-examiner agreement of the ICON score of the nine orthodontists was moderate measured with the single estimate of the ICC (0.60), and high measured with the average estimate (0.93). Spearman’s correlation coefficient between the ICON score (calibrated) and the gold standard was sufficient: 0.78. The sensitivity and specificity were 1 and 0.36, respectively. The best compromise between sensitivity and specificity was at a cut-off point of 52, instead of the international ICON cut-off point of 43. There was a significant difference in ICON score between the non-calibrated orthodontists and the calibrated orthodontist, mainly based on the Aesthetic Component (AC) of the Index of Orthodontic Treatment Need (IOTN). It can be concluded that the ICON needs to be adjusted when used to determine treatment need in the Dutch orthodontic population.

Introduction

Assessing orthodontic treatment need is a complex issue. When deciding whether or not a patient should be orthodontically treated, both the desire of the patient (and/or parent) and the opinion of the orthodontist must be taken into account. Further, when there is a limited number of orthodontists and limited resources, priority should be given to patients with the highest treatment need (Richmond *et al.*, 1994).

Numerous indices have been developed since the 1960s either to rank or score the severity of a malocclusion relative to a preconceived orthodontic ideal, or in terms of treatment need (Draker, 1960; Grainger, 1967; Salzmann, 1968; Summers, 1971; Linder-Aronson, 1974; Lundström, 1977; Brook and Shaw, 1989; Buchanan, 1991; Shaw *et al.*, 1991; Richmond *et al.*, 1992). None of these indices have been developed and validated for both the deviation from normal occlusion and treatment need. In addition, when an index is validated against specialists’ opinion (the gold standard) in a certain region, this does not indicate that the index will also be valid in other geographic regions. Richmond and Daniels (1998a,b) showed that the opinion of the orthodontic

professional depends on the region or country where the specialist practices. They found that Dutch orthodontists demonstrated the lowest recommended treatment rate in comparison with orthodontists in the USA and in nine other European countries, and that Dutch orthodontists, together with orthodontists from the USA, were the strictest in the assessment of the acceptability of the outcome of treatment. The geographically based difference in thinking and the fact that the known orthodontic indices did not incorporate outcome, need, and complexity of treatment, was one of the main reasons for developing a new index: the Index of Complexity, Outcome and Need (ICON; Daniels and Richmond, 2000).

For the development of the ICON, 97 orthodontists from nine different countries judged treatment need and acceptability of outcome of a heterogeneous sample of 240 initial and 98 study casts of treated patients. Five highly predictive occlusal traits were identified (Table 1) and then used to predict the specialists’ decision using regression analysis. The ICON requires that these five occlusal traits be scored, multiplied by their respective weights, and then summed. Cut-off values were determined for the

Table 1 Components of the Index of Complexity, Outcome and Need (ICON) with their scoring range and weights.

ICON components	Scoring range	Weight
Aesthetic Component of the Index of Orthodontic Treatment Need	1–10	7
Upper arch crowding	0–5	5
Crossbite	0–1	5
Anterior vertical relationship	0–4	4
Sagittal relationship of the buccal segment	0–2	3

dichotomous judgements by plotting specificity, sensitivity and overall accuracy. In this way, a sum score of 43 was found as the international cut-off point for treatment need: when the ICON score is higher than 43, treatment is indicated. A cut-off point of 31 was found for the outcome of treatment. With a score lower than 31 the treatment result is considered acceptable.

The index is based on 'average' international orthodontic opinion and could, as the authors claim, provide the means for comparison of treatment thresholds in different countries and serve as a basis for quality assurance standards in European orthodontics. Because it is an index of treatment need, an index of severity of malocclusion, and an index of treatment outcome, the ICON should offer significant advantages over other indices of treatment need. However, it remains to be seen whether the ICON is valid with respect to its assessment of need, outcome, and complexity in each country separately. An accepted manner to validate a new orthodontic index is by means of comparing it to a gold standard: a pooled decision of specialists about treatment need. Recently, it was found that the 'need part' of the ICON was valid with respect to its relationship with the opinion of specialists practicing in central Ohio, USA (Firestone *et al.*, 2001).

The aim of this study was to validate the ICON in a Dutch orthodontic setting by relating it to the orthodontic opinion [the pooled decision of seven Dutch orthodontists (clinical sense) as the gold standard] and to examine whether the indicated cut-off point is acceptable for the Dutch situation. Furthermore, the reliability of this index was explored, for both a calibrated orthodontist and non-calibrated orthodontists.

Materials and methods

Sample

The data of 102 patients from nine different orthodontic practices in the north-west of the Netherlands were used. The patients applied for treatment in the year 2000. The 102 subjects represented a spectrum of malocclusion types and severity as seen in the average orthodontic practice in the Netherlands. The distribution of the Dental Health Component (DHC) grades of the Index of Orthodontic Treatment Need (IOTN) are shown in Table 2. The study

Table 2 Distribution of the Dental Health Component grades of the Index of Orthodontic Treatment Need of the sample (total $n = 102$).

Grade	Treatment need	n
1	None	0
2	Little	2
3	Moderate	32
4	High	45
5	Very high	23

casts, lateral cephalometric and panoramic radiographs, and extraoral photographs of the patients, all taken before orthodontic treatment, were used to rate orthodontic treatment need.

Examiners

Nine volunteer orthodontists (eight males and one female) from a group of 14 who are members of a regional orthodontic association in the north-west Netherlands rated the records of the 102 subjects. The experience of these specialists ranged from 3 to 37 years and all had their own practice. In addition, one author (TJL), who was calibrated in the use for the ICON, scored the ICON.

Procedure

This procedure partly resembled that used by Firestone *et al.* (2001). The nine orthodontists were invited to the Department of Orthodontics of the Academic Centre for Dentistry Amsterdam (ACTA) for one day. The ICON was introduced by the calibrated orthodontist before they scored the treatment need of the cases. The casts, photographs, and radiographs were displayed in a fixed order on tables. Each examiner started with a different case. There was no time limit.

For all 102 cases treatment need was rated on a scale ranging from 1 (no/minimal need) to 7 (very high need) by the nine examiners. They were asked to rate treatment need without taking into account the source of treatment funding, the amount of money available for treatment, the treating orthodontist, and whether or not treatment was carried out. After assessing the inter-examiner agreement of these scores, they were averaged across the observers. The resulting score was called the 'clinical sense'. The participants were further asked to indicate which score on the 7-point scale indicated the cut-off point above which they thought orthodontic treatment was required. This was called the 'indicated treatment point'.

In addition, the nine orthodontists assessed the ICON for 49 of the 102 cases. The calibrated orthodontist scored all the 102 cases on the ICON. The scoring range of the ICON is shown in Table 1. Each case was categorized as 'treatment need according to the ICON' if the ICON score of the calibrated orthodontist was higher than 43 points (the

international cut-off point; Daniels and Richmond, 2000). The other cases were categorized as 'no treatment need according to the ICON'.

To determine intra-examiner agreement, five of the nine examiners were asked to rate 15 of the 49 cases again for both clinical sense and the ICON. This part of the study was undertaken in their own orthodontic practices and took place between 30 and 60 days after the first session. For the assessment of the intra-examiner agreement of the ICON rated by the calibrated orthodontist, 18 of the 102 cases were scored again after 30 days.

Statistical analysis

Agreement between and within observers was assessed by the Intraclass Correlation Coefficient (ICC) for ordinal or quantitative variables (Bartko, 1966; Fleiss and Cohen, 1973). The average estimate of the ICC for inter-examiner agreement between the nine orthodontists was calculated to show the reliability of the combined ICON, which was used additionally for the validity assessment. The single estimate was used to assess the inter-examiner agreement of an individual score of a non-calibrated orthodontist. The Kappa (κ) (Cohen, 1960) and percentage agreement were used for nominal scales. Percentage agreement was used in addition to κ , because κ can drop dramatically based on the prevalence of the variable involved (Altman, 1991). Since the cases in the present study were a representative sample of patients referred for orthodontic treatment, the prevalence of treatment need was high. Spearman's correlation coefficient, and the specificity (the ability to correctly identify the absence of treatment need) and sensitivity (the ability to correctly identify treatment need) were used to compare the ICON score of the calibrated orthodontist with the gold standard. The Dutch optimum cut-off point for the ICON concerning treatment need was assessed by plotting a receiver-operating characteristic curve (Metz, 1978). The comparison of the raw ICON scores of the calibrated orthodontist and the raw ICON scores of the non-calibrated orthodontists was carried out using a paired *t*-test. To assess agreement between the calibrated and non-calibrated specialists concerning the total score and its different components, the ICC, κ statistics, and agreement were used. These measurements were also used to assess the relationship of the different components of the ICON within the group of nine specialists.

When not mentioned in the text, the measurements were statistically significant at $P < 0.05$.

Results

Reliability of the gold standard

Inspection of the intra-examiner agreement of five orthodontists and the paired inter-examiner agreement of all nine orthodontists resulted in the exclusion of two. The inter-examiner agreement (ICC) for the remaining

seven experts was 0.90 and the ICC for the intra-examiner-agreement ranged from 0.60 to 0.86.

For each case the mean clinical sense determined by the seven orthodontists was compared with their mean indicated treatment point, which was 4.43. When this value was more than or equal to 4.43, the case was labelled as 'treatment need'. The others were labelled as 'no treatment need'. The examiner agreement (κ) of this dichotomous score ranged for the intra-examiner agreement from 0.33 ($P > 0.05$) to 0.58, and for the inter-examiner agreement from 0.02 ($P > 0.05$) to 0.49. These results are low due to the high prevalence of scores above 4.43 (Altman, 1991).

Reliability of the ICON

The intra-examiner agreement (ICC) of the raw ICON score of the calibrated orthodontist was 0.89. After dichotomizing the ICON based on a cut-off of 43, the agreement between the two scores of the examiner was 78 per cent. κ was -0.091 ($P > 0.05$).

The intra-examiner agreement of the raw ICON score of the five non-calibrated orthodontists ranged from 0.52 (range: 0.01–0.81) to 0.86 (range: 0.63–0.95). For the dichotomous ICON, κ ranged from 0.20 ($P > 0.05$) to 0.40.

The inter-examiner agreement of the raw ICON score was 0.93 measured by the average estimate of the ICC. The single estimate of the ICC was 0.60. The dichotomous ICON gave a reliability estimate of 0.20 ($P > 0.05$) to 0.60. As mentioned before, the intra- and inter-examiner results of the categorized scores were low due to the high prevalence of scores above 43 (Altman, 1991).

The inter-examiner agreement of the separate components of the ICON are shown in Table 3.

Validity of the ICON

Spearman's correlation coefficient between the raw ICON score of the calibrated orthodontist and the gold standard was 0.78 ($P < 0.001$). Twenty-five of the cases had a mean clinical sense for treatment need lower than the mean indicated treatment point (4.43), and 77 of the cases had a mean clinical sense for treatment need higher than 4.43. With a cut-off point of 43, these numbers were 9 and 93 for the dichotomous ICON score. The distribution of the categorized ICON scores over the categorized clinical sense for treatment need (the gold standard) are shown in Table 4. The sensitivity and specificity of the ICON were 1 and 0.36, respectively.

The best compromise for sensitivity and specificity in a Dutch orthodontic setting, compared with the Dutch gold standard, seemed to be at a cut-off point of 52 (Figure 1). With this cut-off point the sensitivity was 0.91, and specificity 0.84. When this same analysis was repeated using the ICON mean of the non-calibrated orthodontists, a cut-off point of 37 was found, with a sensitivity of 0.97 and a specificity of 0.86 (Figure 2). Spearman's correlation

Table 3 Inter-examiner agreement of the different components of the Index of Complexity, Outcome and Need (ICON) for the nine non-calibrated orthodontists.

ICON components	ICC average	Range ICC single	Range Kappa
Aesthetic Component of the Index of Orthodontic Treatment Need	0.93	0.44–0.77	
Upper arch crowding	0.90	0.26–0.80	
Crossbite			0.58–0.85
Anterior vertical relationship	0.96	0.52–0.85	
Sagittal relationship of the buccal segment			0.03*–0.47

ICC, intraclass correlation coefficient.

* $P > 0.05$. In all other cases statistically significant at $P < 0.05$.**Table 4** Distribution of the categorized calibrated scores of the Index of Complexity, Outcome and Need (ICON) (cut-off point = 43) over the categorized clinical sense for treatment need (the gold standard).

Gold standard	ICON		Total
	+	–	
+	77	0	77
–	16	9	25
Total	93	9	102

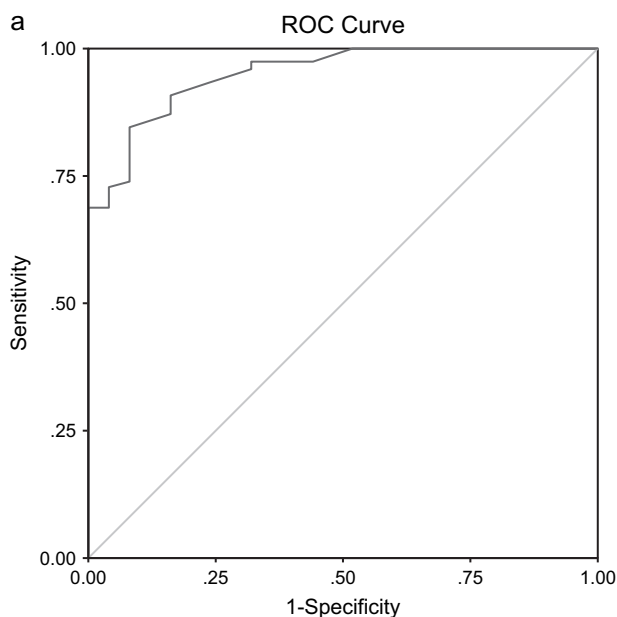
coefficient between the raw ICON score of the nine orthodontists and the gold standard was 0.79 ($P < 0.001$).

Comparison of the results of the calibrated and non-calibrated orthodontists

The mean of the calibrated raw ICON score was 63.88 (SD = 11.65) and for the non-calibrated raw ICON score 50.58 (SD = 15.30; $t = 8.00$, $df = 48$, $P < 0.001$). The distribution of the dichotomous ICON score (cut-off 43) of the non-calibrated orthodontists for the cases with and without treatment need according to the calibrated orthodontist (dichotomous version) is shown in Table 5. The ICC between the non-calibrated and calibrated raw ICON scores was 0.55. Table 6 shows the agreement of the different components of the ICON between the calibrated and non-calibrated specialists.

Discussion

The present study assessed the reliability and validity of the treatment need part of the ICON against Dutch orthodontic opinion, with seven experts as the gold standard. Two of the original nine orthodontists were excluded because of low intra- and inter-examiner agreement. The raw scores of the clinical sense of treatment need seemed to be similar to



Cut-off	Sensitivity	Specificity
43.5	1.00	.36
44.5	1.00	.40
45.5	1.00	.48
47.0	.97	.56
49.0	.97	.68
50.5	.96	.68
51.5	.94	.78
52.5	.91	.84
53.5	.90	.84
54.5	.87	.84
56.0	.84	.92
58.0	.82	.92
59.5	.78	.92
60.5	.74	.92
61.5	.73	.96
62.5	.70	.96
64.0	.69	.96
65.5	.69	1.00

Figure 1 (a) Sensitivity and specificity at different cut-off points for the calibrated specialist's score. Receiver-operating characteristic. (b) Different cut-off points with their sensitivity and specificity.

those in other studies that measured the validity of indices (Richmond *et al.*, 1992; Younis and Vig, 1997; Firestone *et al.*, 2001). The results of the dichotomous values, measured by κ , however, were low. As mentioned previously, this is most likely due to the high prevalence of scores above the indicated treatment point of 4.43 (Altman, 1991). In the present study a sample was chosen which resembled the patient population in an orthodontic practice, representing a full spectrum of malocclusion types. Thus, all patients were referred for orthodontic treatment and scored, as a consequence, high on treatment need. Such a sample was chosen in order to be able to assess the usefulness of the ICON in every day practice.

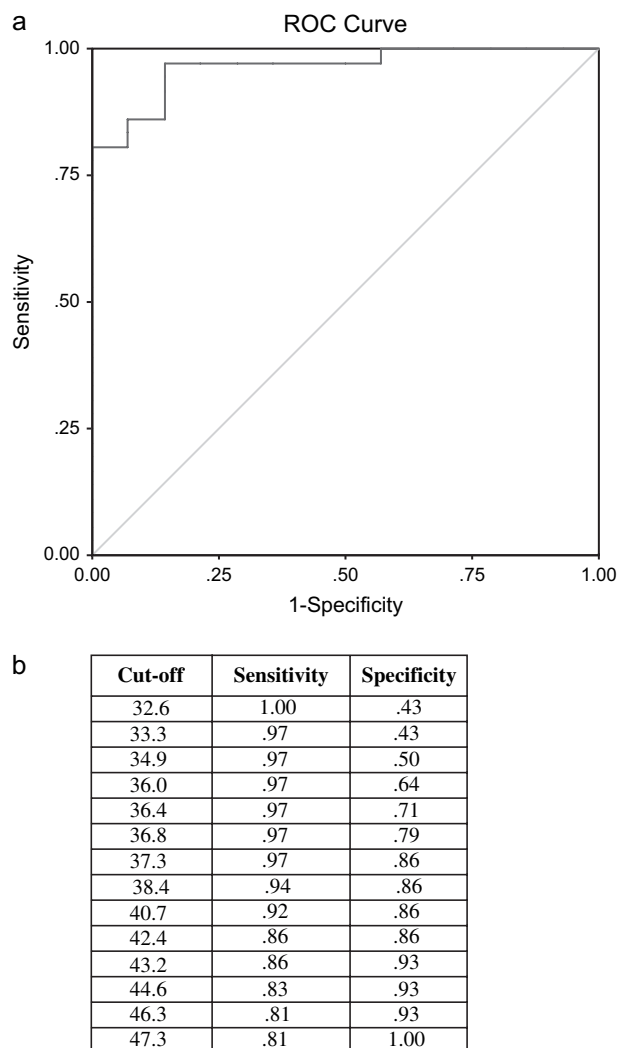


Figure 2 (a) Sensitivity and specificity at different cut-off points for the mean non-calibrated specialists' score. Receiver-operating characteristic. (b) Different cut-off points with their sensitivity and specificity.

The ICON showed a very good level of agreement for the calibrated orthodontist. The reliability of the pooled ICON score of the nine orthodontists together also appeared to be very good, and sufficiently reliable to use it to assess validity. However, the inter-examiner agreement in terms of planning to use the ICON as a single estimate by non-calibrated orthodontists was moderate to good. When used for the latter purpose it must be stated that the index was more reliable for the calibrated examiner. The intra-examiner agreement of the raw ICON score for the calibrated examiner (TJL) was comparable to that found by Firestone *et al.* (2001). As with the results of the dichotomous values of the clinical sense, the agreement of the dichotomous values of the ICON was low.

With respect to validity, the relationship between the ICON (raw score, applied by TJL) and the gold standard was comparable with the findings of Firestone *et al.* (2001)

Table 5 Distribution of the categorized non-calibrated scores of the Index of Complexity, Outcome and Need (ICON) over the categorized calibrated ICON scores.

Calibrated ICON	Non-calibrated ICON		Total
	+	–	
+	34	9	43
–	0	6	6
Total	34	15	49

Table 6 Agreement of the different components of the Index of Complexity, Outcome and Need (ICON) between the calibrated and the non-calibrated specialists.

ICON components	ICC	Kappa	Agreement (%)
Aesthetic Component of the Index of Orthodontic Treatment Need	0.63		
Upper arch crowding	0.74		
Crossbite		0.82	92
Anterior-vertical relationship	0.88		
Sagittal relationship of the buccal segment		0.5	65

and to the results of other indices (Beglin *et al.*, 2001). Spearman's correlation coefficient of 0.78 can be described as a substantial relationship. However, the results pertaining to the dichotomous scores were less satisfying. The sensitivity and specificity of the ICON, using the international cut-off point of 43, were 1.00 and 0.36, respectively, whereas Firestone *et al.* (2001) found a sensitivity of 0.94 and a specificity of 0.86 when compared with the local gold standard in Ohio. Their findings were comparable with those of Daniels and Richmond (2000). The best compromise for sensitivity and specificity (0.85 and 0.86, respectively) was found to be at a cut-off of 43. Based on this single study, the international cut-off was set at this value. The findings of the present investigation showed that for the Dutch situation, and for patients already referred for orthodontic treatment, the cut-off point was 52. With this adjusted cut-off point, the sensitivity became 0.91 and the specificity 0.84. This higher cut-off point is in agreement with the results of Richmond and Daniels (1998a) where the Dutch orthodontists showed the lowest recommended treatment rate in comparison with orthodontists from other countries. When used to determine treatment need in every day practice, it would appear that every country needs its own cut-off level.

The somewhat different results in this study may also be explained by the sample that was used. In this investigation, study casts, as well as lateral cephalometric and panoramic radiographs, and extraoral photographs of the patients were used. This is in contrast with the use of only casts in other

studies assessing the validity of an index (Richmond *et al.*, 1992; Younis and Vig, 1997; Daniels and Richmond, 2000; Firestone *et al.*, 2001). By presenting the orthodontists with the abovementioned information, the situation in every day practice, where the decision for treatment will also be based on complete patient information, was more closely resembled. Although Han *et al.* (1991) suggested that dental casts were the most important stimuli influencing orthodontists' treatment decisions, more so than radiographs and photographs, the incorporation of all information contributes to the external validity of the results.

There was a significant difference between the mean score of the nine orthodontists and that of the calibrated orthodontist. Related to this, a different cut-off value for the ICON was found depending on whether the ICON score of the calibrated or non-calibrated orthodontist was used. The difference seemed to be mainly caused by two of the five ICON-items: the Aesthetic Component (AC) and sagittal occlusion. The agreement between the mean non-calibrated and calibrated score for these two items was only moderate. Because of the much higher weight for this last item (7 instead of 3), the IOTN-AC seemed to be the largest cause of the difference in the total ICON score between the nine non-calibrated and the calibrated orthodontist. Within the group of non-calibrated orthodontists, the agreement of the AC was moderate to high. These results were only slightly lower than those in earlier studies (Evans and Shaw, 1987; Brook and Shaw, 1989; Shaw *et al.*, 1991). It should be stressed that in the calibrating process, the IOTN-AC proved to be difficult to learn. This may also be the case for other researchers. With this knowledge one would expect a low validity for the AC. Indeed, studies assessing the IOTN-AC demonstrated a moderate validity (Richmond *et al.*, 1995; Beglin *et al.*, 2001). Those authors found a respective percentage agreement of 67 and a κ of 0.70 between the IOTN-AC and their regional gold standard.

It remains a difficult issue whether or not to use an index to determine treatment need. The limitation of every index is that it is a better tool for epidemiological studies than for determining treatment need for individual subjects. If not adjusted, the ICON does not seem to be the ideal index for determining treatment need in Dutch orthodontic practice. The question still remains as to which norm must be adjusted; that of a calibrated orthodontist or the norm of several non-calibrated specialists who more reflect the reality of life. The ultimate consequence of the appropriate use of the ICON in its present form would be to calibrate all orthodontists.

Future studies should investigate the usefulness of the ICON in relation to other indices of treatment need in the Dutch situation.

Conclusions

The reliability of the ICON (concerning treatment need) in Dutch orthodontic practice is moderate to good

for non-calibrated orthodontists, and seems good for calibrated orthodontists; however, with respect to validity, some caution is indicated. The international cut-off value of 43 for the ICON did not appear to be useful in the Netherlands; a higher cut-off point, 52, was found. However, this was only valid when the index was scored by a person who had been calibrated. Finally, a difference emerged in the scores of the ICON between non-calibrated and calibrated specialists, which seems to be primarily due to one of the components of the index: the AC of the IOTN (Table 6).

Address for correspondence

Irene H. A. Aartman
Department of Social Dentistry and Behavioural Sciences
Academic Centre for Dentistry Amsterdam
Louwesweg 1
1066 EA Amsterdam
The Netherlands
E-mail: i.aartman@acta.nl

Acknowledgements

The authors wish to thank the four orthodontists and one dentist who took part in the pilot study, and the nine orthodontists for their cooperation during the research project. We would also like to thank C. P. Daniels, R. B. Kuitert, J. Kieft and A. E. V. Sint Jago for their help.

References

- Altman D G 1991 Practical statistics for medical research. Chapman and Hill, London
- Bartko J J 1966 The intraclass correlation coefficient as a measure of reliability. *Psychological Reports* 19: 3–11
- Beglin F M, Firestone A R, Vig K W L, Beck F M, Kuthy R A, Wade D 2001 A comparison of the reliability and validity of three occlusal indexes of orthodontic treatment need. *American Journal of Orthodontics and Dentofacial Orthopedics* 120: 240–246
- Brook P H, Shaw W C 1989 The development of an Index of Orthodontic Treatment Priority. *European Journal of Orthodontics* 11: 309–320
- Buchanan I B 1991 The evaluation and initial testing of an index to assess orthodontic treatment standards: the PAR index. Thesis, University of Manchester
- Cohen J 1960 A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20: 37–46
- Daniels C P, Richmond S 2000 The development of an index of complexity, outcome and need (ICON). *Journal of Orthodontics* 27: 149–162
- Draker H L 1960 Handicapping labio-lingual deviations: a proposed index for public health purposes. *American Journal of Orthodontics* 46: 295–315
- Evans R, Shaw W 1987 Preliminary evaluation of an illustrated scale for rating dental attractiveness. *European Journal of Orthodontics* 9: 314–318
- Firestone A R, Beck F M, Beglin F M, Vig K W L 2001 Validity of the Index of Complexity, Outcome, and Need (ICON) in determining orthodontic treatment need. *Angle Orthodontist* 72: 15–20

- Fleiss J L, Cohen J 1973 The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement* 33: 613–619
- Grainger R M 1967 Orthodontic Treatment Priority Index Public Health Service Publication No 1000, Series 2, No. 25. Government Printing Office, Washington DC, USA
- Han U K, Vig K W, Weintraub J A, Vig P S, Kowalski C J 1991 Consistency of orthodontic treatment decisions relative to diagnostic records. *American Journal of Orthodontics and Dentofacial Orthopedics* 100: 212–219
- Linder-Aronson S 1974 Orthodontics in the Swedish Public Dental Health System. *Transactions of the European Orthodontic Society*, pp. 233–240
- Lundström A 1977 Need for treatment in cases of malocclusion. *Transactions of the European Orthodontic Society*, pp. 111–123
- Metz C E 1978 Basic principles of ROC analysis. *Seminars in Nuclear Medicine* 8: 283–298
- Richmond S D, Daniels C P 1998a International comparison of professional assessment in orthodontics. Part 1: Treatment need. *American Journal of Orthodontics and Dentofacial Orthopedics* 113: 180–185
- Richmond S D, Daniels C P 1998b International comparisons of professional assessments in orthodontics. Part 2: Treatment outcome. *American Journal of Orthodontics and Dentofacial Orthopedics* 113: 324–328
- Richmond S D *et al.* 1992 The development of the PAR index (Peer Assessment Rating): reliability and validity. *European Journal of Orthodontics* 14: 125–139
- Richmond S, Roberts C T, Andrews M 1994 Use of the Index of Orthodontic Treatment Need (IOTN) in assessing the need for orthodontic treatment pre-coordinates of the curve and post-appliance therapy. *British Journal of Orthodontics* 21: 175–184
- Richmond S *et al.* 1995 Calibration of dentists in the use of occlusal indices. *Community and Dental Oral Epidemiology* 23: 173–176
- Salzmann J A 1968 Handicapping malocclusion assessment to establish treatment priority. *American Journal of Orthodontics* 54: 749–765
- Shaw W C, Richmond S, O'Brien K D, Brook P, Stephens C D 1991 Quality control in orthodontics: indices of treatment need and treatment standards. *British Dental Journal* 170: 107–112
- Summers C J 1971 The occlusal index: a system for identifying and scoring occlusal disorders. *American Journal of Orthodontics* 59: 552–567
- Younis J W, Vig K W L 1997 A validation study of three indexes of orthodontic treatment need in the United States. *Community Dentistry and Oral Epidemiology* 25: 358–362

Copyright of European Journal of Orthodontics is the property of Oxford University Press / UK and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.