

ORIGINAL ARTICLE

Validity assessment and determination of the cutoff value for the Index of Complexity, Outcome and Need among 12–13 year-olds in Southern Chinese

Zheng-Yu Liao^{1,2,3}, Fan Jian^{1,2}, Hu Long^{1,2}, Yun Lu^{1,4}, Yan Wang^{1,2}, Zhi Yang^{1,5}, Yu-Wei He^{1,2}, Peter Wamalwa^{1,6}, Jing Wang^{1,2}, Nian-Song Ye^{1,2}, Sheng Wang^{1,2} and Wen-Li Lai^{1,2}

To validate the use of the Index of Complexity, Outcome and Need (ICON) in assessing orthodontic treatment need among 12–13 year-olds in southern China, we determined the threshold value of ICON based on Chinese orthodontists' judgments. The samples consisted of 335 students in grade 7 from 16 randomly selected middle schools in Chengdu, China. Three associate professors provided ICON scores for each participant and the results were compared with the gold standard judgments from 25 experts on treatment needs. Based on the gold standard, 195 casts belonged to the treatment category, while the rest 140 belonged to the no-treatment category. With the international cutoff point of 43, the sensitivity and specificity of the ICON score were 0.29 and 0.98. The best compromise between sensitivity and specificity in Chengdu, compared with the gold standard, was found at a cutoff point of 29, and the sensitivity and specificity were 0.88 and 0.83. When used to evaluate the treatment need of 12–13 year-olds in southern China, the international ICON cutoff value did not correspond well with Chinese orthodontists' judgments; a lower cutoff value of 29 offered a greater sensitivity and specificity with respect to expert orthodontists' perception of treatment need.

International Journal of Oral Science (2012) 4, 88–93; doi:10.1038/ijos.2012.24; published online 1 June 2012

Keywords: Index of Complexity, Outcome, and Need; occlusal index; treatment need

INTRODUCTION

Numerous indices have been developed since 1960s to rank/score the severity of a malocclusion relative to a preconceived orthodontic ideal or in terms of treatment need.¹ Some indices have been developed to assess orthodontic treatment need, such as the Index of Orthodontic Treatment Need (IOTN),² the Handicapped Labio-Lingual Deviation index³ and the Dental Aesthetic Index,⁴ while others have been developed to evaluate orthodontic treatment outcome, such as the Peer Assessment Rating.⁵ The above-mentioned indices only involve one aspect of orthodontic treatment; however, the treatment need and treatment outcome are also considered important by orthodontists and patients. Therefore, the Index of Complexity, Outcome and Need (ICON),⁶ which assesses both elements, was developed and used widely in recent years.

With social improvements and economical developments in China, more and more people are receiving orthodontic treatments. There is a need in China for an overall objective assessment of orthodontic treatment need based on malocclusion severity. The ICON can achieve this purpose by evaluating the orthodontic treatment needs, complexity and outcomes. If this system is adopted, it will help maximise the benefit of orthodontic treatment by guiding the distribution of limited

health resources based on those who need it most. In a recent study determining the malocclusion complexity and orthodontic treatment needs in urban Iranian schoolchildren, the authors used the ICON and IOTN to assess the relationship between the two indices and found that ICON is a good substitute for IOTN, yet ICON results in a lower treatment-need threshold.^{7–8}

There is evidence that geographic location may affect orthodontists' determination of treatment need and outcome.^{9–10} It is necessary to validate any index purporting to identify treatment need with the opinions of orthodontists practicing within a limited geographic region. Therefore, the primary aim of this study was to validate ICON as an index of treatment need in southern China and to investigate whether the indicated cutoff point was acceptable to the orthodontists in southern China. The secondary aim was to plot a receiver operating characteristic (ROC) curve, and to find a new cutoff point for the Chinese population.

MATERIALS AND METHODS

Study samples

The participants were randomly selected from grade-7 students at 16 randomly selected middle schools in the main city and suburbs of

¹Department of Orthodontics, West China School of Stomatology, Sichuan University, Chengdu, China; ²State Key Laboratory of Oral Diseases, Sichuan University, Chengdu, China; ³Department of Orthodontics, College of Stomatology, Nanchang University, Nanchang, China; ⁴College of Stomatology, Dalian Medical University, Dalian, China; ⁵Joint Orthodontics-orthognathic Surgery Center, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University, Shanghai, China and ⁶Clinical Supervisor and Lead Dentist, Latrobe Community Health Service, Morwell, Australia

Correspondence: Professor WL Lai, Department of Orthodontics, West China School of Stomatology, Sichuan University, No 14, Section 3, Ren Min Nan Road, Chengdu 610041, China

E-mail: wenlilai@hotmail.com

Received 26 December 2011; Accepted 15 February 2012

Chengdu, the capital of Sichuan Province, China. Participants were enrolled from January 2008 to March 2008. The initial number of selected students was 350. This study was approved by the Committee for the Use of Human Subjects in Research, Sichuan University.

The exclusion criteria were a history of orthodontic treatment, physical and mental impairments, cleft lip and/or palate, dentofacial deformities and the presence of mixed dentition. Overall, 335 out of the 350 (95.7%) participants met the inclusion criteria, with 174 boys (51.9%) and 161 girls (48.1%). Their ages ranged from 12 to 13 years. The subjects represented a full range of malocclusions and severities. The orthodontic study casts were prepared for each subject and used for further assessments.

Examiners

Volunteers were selected from specialist orthodontists; all of them are working in the Department of Orthodontics, West China School of Stomatology. Finally, three experienced associate professors were invited to judge the ICON. They received extensive training and underwent the calibration processes.^{6,11}

Procedure

In the first session, the three calibrated examiners assessed the 335 study casts under strict adherence to the ICON guidelines. Study casts were displayed on tables in a fixed order. Each examiner started with a different case. The examiners were allowed to work at their own pace with no time limit. This procedure was similar to that described by Louwse *et al.*¹ Five components of ICON (Table 1) were scored, multiplied by respective weights and summed up.¹²

At the second session, organized 4 weeks later, a random set of 50 models was assessed by the three examiners in order to test their reliability. The kappa value for the three examiners was calculated by assessing 50 sets of models twice within a 4-week interval.

Two experts from the Department of Orthodontics, West China School of Stomatology, Sichuan University, were asked to score the same 335 casts and record the need for treatment of each with a score on a scale from 1 to 7. The experts were senior orthodontists in China and had made great contributions in the development of orthodontics in China. They used the above-mentioned scale to evaluate the casts, with 1 denoting ‘none/minimal’ need and 7 denoting ‘very great’ need.

This procedure was similar to that used by Firestone *et al.*¹³ The experts were required to rate treatment need without considering the treatment cost or orthodontists’ skills required to treat the case.

The average score of the two experts was used. The resulting score was designated as the ‘clinical sense’. The raters were further asked to point out which score on the seven-point scale indicated the cutoff point above which they felt orthodontic treatment was indicated. This score was termed the Indicated Treatment Point (ITP) score, i.e., the ‘gold standard’.¹⁴ The method of determining the gold standard was similar to the approach employed by Firestone *et al.*¹³

To assess the representativeness and reliability of the gold standard, we randomly selected 50 casts 8 months later and asked another 23 orthodontic experts from domestic and board to give clinical sense to the casts using the method mentioned above.

Statistical analysis

The Shapiro–Wilk test¹⁵ was used to test the normality of the ICON data. The rank-sum test¹⁶ was used to test for gender differences when using ICON in determining orthodontic need. The statistical analysis was done using SPSS 15.0 software (SPSS Inc, Chicago, IL, USA).

For evaluation of the reliability of the gold standard, the mean ITP and kappa value were used. The correlation between and within raters was estimated by kappa coefficient. The gold standard was determined in the following way; first, the ITP for the two experts was calculated. Then, the mean rater score of the two experts for each cast was calculated, and finally, the mean score for each cast was compared with the ITP score. If the cast score was equal to or greater than the ITP value, the case was assigned to the ‘treatment’ category, and if the cast score was less than the ITP score, the cast was assigned to the ‘no treatment’ category.

To determine the reliability of ICON, correlations between and within raters were estimated by the kappa value. For evaluation of the validity of ICON, Spearman’s correlation coefficient¹⁷ was applied to find the correlation between the ICON score and the gold ITP score. The overall agreement (simple kappa coefficient) of ICON with the gold standard (the decisions of the orthodontists) was calculated. Sensitivity and specificity were used to compare the casts’ ICON scores recorded by the three examiners with the gold standard. The sensitivity was the percentage of those who were identified as needing treatment

Table 1 ICON Scoring System

| Component | Score | | | | | | Weight |
|--|--|--|----------------------|---------------|------------|--------------------------|--------|
| | 0 | 1 | 2 | 3 | 4 | 5 | |
| Aesthetic assessment | Score from 1 to 10 according to the AC of the IOTN | | | | | | 7 |
| Upper arch | | | | | | | |
| Crowding | <2 mm | 2.1–5 mm | 5.1–9 mm | 9.1–13 mm | 13.1–17 mm | >17 mm or impacted teeth | 5 |
| Spacing | <2 mm | 2.1–5 mm | 5.1–9 mm | >9 mm | | | 5 |
| Crossbite | No crossbite | Crossbite present | | | | | 5 |
| Incisors’ bite | | | | | | | |
| Open bite | Edge to edge | <1 mm | 1.1–2 mm | 2.1–4 mm | >4 mm | | 4 |
| Overbite | 1/3 lower incisor coverage | 1/3 to 2/3 coverage | 2/3 to fully covered | Fully covered | | | 4 |
| Buccal segment antero-posterior relationship | Cusp to embrasure only; class I, II or III | Any cusp relation but not including cusp to cusp | Cusp to cusp | | | | 3 |

AC, aesthetic component; ICON, Index of Complexity, Outcome and Need; IOTN, Index of Orthodontic Treatment Need.

The ICON has five components with a weighting for each: (1) aesthetic component of IOTN (weight 7); (2) upper arch crowding or spacing (weight 5); (3) crossbite (weight 5); (4) overbite or open bite (weight 4); and (5) buccal segment antero-posterior relationship (weight 3).

among all cases needing treatment. The specificity was the percentage of those who were identified as not needing treatment among all cases not needing treatment.

The Spearman's correlation coefficient was used to evaluate the correlation between ICON and the gold standard. The Chinese optimum cutoff point for ICON regarding treatment need was assessed by plotting a ROC curve.¹⁸⁻²⁰ The area under the ROC curve was used to discriminate the randomly chosen cases in need of orthodontic treatment from those cases with no treatment needs.²¹

As an important supplementary part of this study, the revised gold standard (i.e., the clinical sense of the 25 experts for the 50 casts) was averaged. The correlation between the gold standard and the ICON scores was evaluated. The ROC curve for the 50 casts was also drawn.

RESULTS

Analysis and results of the 335 casts

ICON scores. The ICON scores ranged from 7 to 102, with the mean value being 35 (Figure 1). Study samples comprised 174 boys (51.9%) and 161 girls (48.1%). The distribution of ICON scores across genders is presented in Figure 2. There were no statistically significant differences in orthodontic treatment need between genders.

Reliability of the gold standard. The kappa coefficient for the two experts was 0.87. The mean indicated treatment point score was 4.20±1.95 (mean±SD) points. The casts with mean scores equal to or greater than 4.20 points were classified to the 'treatment' category. The remaining casts, with scores less than 4.20 points, were assigned to the 'no treatment' category. One hundred and ninety-five (58%) casts belonged to the 'treatment' category, while 140 (42%) were categorized as 'no treatment' category.

Reliability of ICON. The kappa coefficient of the ICON scores of the three ICON raters was 0.82. For intra-rater reliability, the kappa coefficients were 0.89, 0.87 and 0.92. These results suggest a relatively high reliability of the ICON score in our study.

Validity of ICON. The Spearman's correlation coefficient between the ICON score and the gold ITP score was 0.83. As mentioned before, 140 casts had mean scores below 4.20, and 195 casts had scores above 4.20. With the cutoff point of 43, these numbers were 245 and 90,

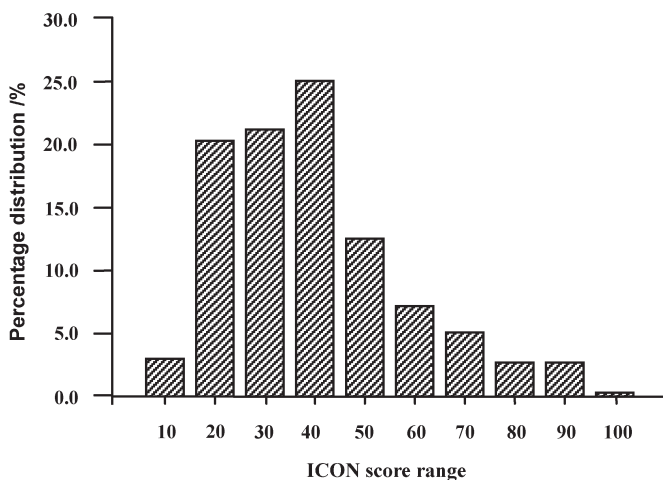


Figure 1 The distribution of the 335 models' ICON scores. ICON, Index of Complexity, Outcome and Need.

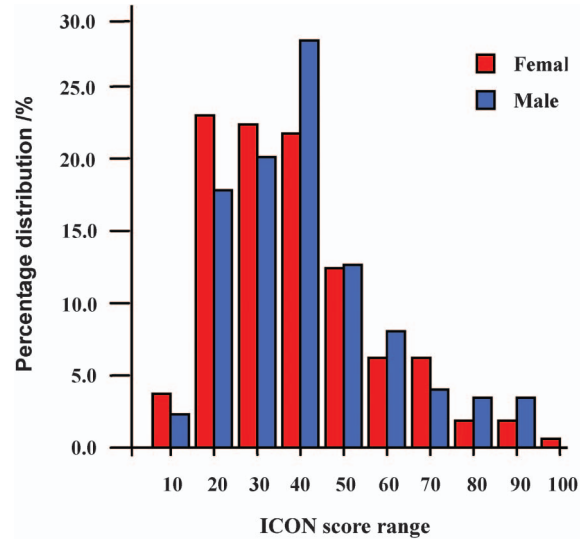


Figure 2 The distribution of the models' ICON scores relating to gender. The data were made up of 174 boys and 161 girls. ICON, Index of Complexity, Outcome and Need.

respectively. The prevalence of orthodontic treatment need according to the recommended cutoff point of 43 was 26.9%. The cross tabulation of the ICON treatment need categories according the gold standard is shown in Table 2. The sensitivity and specificity of the ICON scores were 0.45 and 0.98, respectively. The overall agreement between the ICON and the gold standard treatment need categories (the kappa coefficient) was 0.38. This agreement was fair according to the Altman classification.²²

The area under the ROC curve was 0.91, suggesting a high validity²¹ of the ICON. Thus, the index reflects the decisions of the gold-standard experts of orthodontists to a high degree (Figure 3). The best compromise between sensitivity and specificity in Chengdu, considering our gold standard, may be found at a cutoff point of 29. Different cutoff points with their sensitivity and specificity are displayed in Table 3. Based on this cutoff point, the sensitivity was 0.86, and the specificity was 0.83. The prevalence of orthodontic need according to an adjusted cutoff point of 29 was 58.2%. The kappa value also improved to 0.71, which is higher than the value of international cutoff point (43). The cross-tabulation between ICON and the gold standard with the new cutoff point of 29 is shown in Table 2.

Table 2 Distribution of the categorized scores of the ICON

| Gold standard | ICON | | Total |
|-------------------|------|-----|-------|
| | + | - | |
| Cutoff print = 43 | | | |
| + | 87 | 108 | 195 |
| - | 3 | 137 | 140 |
| Total | 90 | 245 | 335 |
| Cutoff print = 29 | | | |
| + | 171 | 24 | 195 |
| - | 24 | 116 | 140 |
| Total | 195 | 140 | 335 |

ICON, Index of Complexity, Outcome and Need.

Over the categorized (with international and adjusted cutoff point) clinical sense for treatment need (the gold standard) ('+' means treatment category; '-' means no treatment category).

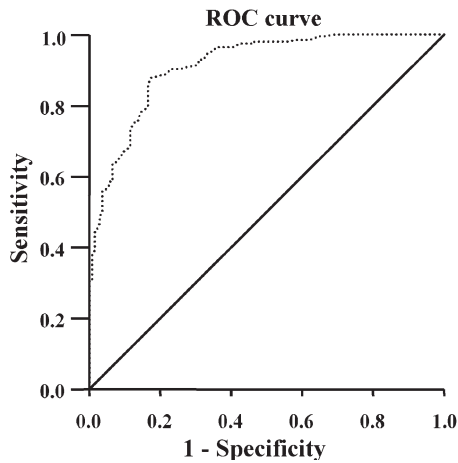


Figure 3 Sensitivity and specificity at different cutoff points for the ICON score (335 models). ICON, Index of Complexity, Outcome and Need.

Under the optimal cutoff point, the number of false negatives (i.e., cases recommended for no orthodontic treatment by the ICON index but judged as needing treatment by the gold standard) decreased from 108 to 24. The number of false positives (i.e., cases recommended to receive orthodontic treatment by the ICON index but not by the expert panel) increased from 3 to 24.

Table 3 Different cutoff points with their sensitivity and specificity

| Cutoff | Sensitivity | Specificity |
|--------|-------------|-------------|
| 20.0 | 0.98 | 0.53 |
| 21.0 | 0.97 | 0.54 |
| 22.0 | 0.96 | 0.63 |
| 23.0 | 0.96 | 0.64 |
| 24.0 | 0.94 | 0.67 |
| 25.0 | 0.92 | 0.69 |
| 26.0 | 0.91 | 0.73 |
| 27.0 | 0.90 | 0.77 |
| 28.0 | 0.89 | 0.79 |
| 29.0 | 0.86 | 0.83 |
| 30.0 | 0.79 | 0.83 |
| 31.0 | 0.75 | 0.87 |
| 32.0 | 0.71 | 0.89 |
| 33.0 | 0.68 | 0.89 |
| 34.0 | 0.66 | 0.91 |
| 35.0 | 0.64 | 0.92 |
| 36.0 | 0.59 | 0.94 |
| 37.0 | 0.56 | 0.95 |
| 38.0 | 0.54 | 0.96 |
| 39.0 | 0.52 | 0.96 |
| 40.0 | 0.49 | 0.97 |
| 41.0 | 0.48 | 0.97 |
| 42.0 | 0.47 | 0.97 |
| 43.0 | 0.45 | 0.98 |
| 44.0 | 0.43 | 0.99 |
| 45.0 | 0.40 | 0.99 |
| 46.0 | 0.38 | 0.99 |
| 47.0 | 0.36 | 0.99 |
| 48.0 | 0.33 | 0.99 |
| 49.0 | 0.31 | 0.99 |
| 50.0 | 0.30 | 1.00 |
| 51.0 | 0.29 | 1.00 |
| 52.0 | 0.28 | 1.00 |

Table 4 Statistical analysis of the gender difference test of the ICON scores by using the rank-sum test

| Gender | N | Mean rank | Sum of ranks | Asymptotic significance (two-tailed) |
|--------|-----|-----------|--------------|--------------------------------------|
| Girls | 161 | 159.78 | 25 724.50 | 0.135 |
| Boys | 174 | 175.61 | 30 555.50 | |
| Total | 335 | | | |

ICON, Index of Complexity, Outcome and Need.

Table 5 Comparison of the diagnostic performance characteristics of the ICON at the international and adjusted ICON score cutoff point for determining orthodontic treatment need when applied to the 335 casts

| ICON score cutoff | Sensitivity | Specificity | Kappa |
|---------------------|-------------|-------------|-------|
| >43 (international) | 0.45 | 0.98 | 0.38 |
| >29 (adjusted) | 0.83 | 0.88 | 0.71 |

ICON, Index of Complexity, Outcome and Need.

Gender differences. The result of normality test revealed that the ICON scores did not obey the Gaussian distribution ($P < 0.05$). Therefore, we used the rank-sum test for gender differences when using the ICON. Table 4 demonstrates that gender differences were not statistically significant in our study samples. Table 5 displays the comparison of the diagnostic performance of ICON at the international (43) and adjusted (29) ICON score cutoff points when applied to the 335 casts.

Analysis results of the 50 casts with the gold standard of 25 experts ICON scores. The ICON scores ranged from 7.0 to 89.0, and the mean ICON score of the casts was 30.

Reliability of the gold standard. The kappa value for the 25 experts was 0.90. The mean indicated treatment point (mean \pm SD) was 4.15 ± 1.26 . Casts with mean scores equal to or greater than 4.15 were classified in the 'treatment' category. The remaining casts, with scores less than 4.15 points, were assigned to the 'no treatment' category.

Reliability of the ICON. The kappa value of the ICON score of the three ICON raters was 0.81.

Validity of the ICON. The Spearman's correlation coefficient between the ICON score and the gold ITP score was 0.84. The area under the ROC curve (0.90) suggested a high validity of the ICON (Figure 4).

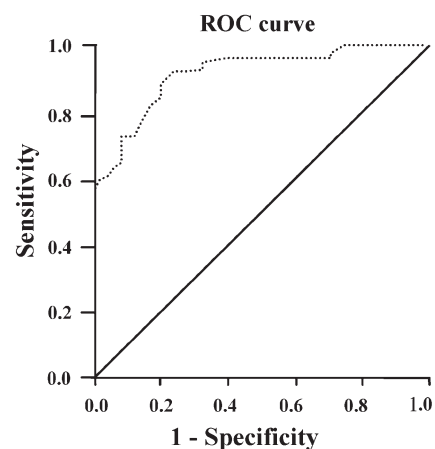


Figure 4 Sensitivity and specificity at different cutoff points for the ICON score (50 models). ICON, Index of Complexity, Outcome and Need.

The best compromise between sensitivity and specificity in Chengdu, compared with our gold standard, may be found at a cutoff point of 29. At this cutoff point, the sensitivity and specificity were 0.86 and 0.79, respectively.

DISCUSSION

The methods for evaluating malocclusion include orthodontic screening exams, dental cast analysis, cephalometric records, etc. There may be a lack of agreement among investigators concerned with assessing occlusion by orthodontic screening exams. However, the occlusal indices based mainly on analysis of dental casts would serve for epidemiological purposes more objectively.

Many investigators employed a subset of casts to test the reliability and validity of ICON.^{1,12–13,23–29} Most scholars concluded that ICON could be a substitute for IOTN, Peer Assessment Rating and Dental Aesthetic Index as a good index for the prediction and evaluation of orthodontic treatment need and treatment outcome.^{12,23,25,28–31} Some scholars believed that the relationship between ICON and Peer Assessment Rating needs to be studied further.³²

There was a relatively high agreement between the scores of the three associate professors using ICON to assess orthodontic treatment need for both inter- and intra-examiner agreement. This may be attributed to that all of the professors have received similar orthodontic trainings in China and worked in the same orthodontic treatment centre, and that the ICON offers a very objective way to determine it. With regard to validity, the relationship between ICON scores and ITP could be described as a relatively strong association, as the Spearman's correlation coefficient³³ was 0.83. However, when using the international cutoff point of 43, the sensitivity, specificity and kappa value were 0.29, 0.98 and 0.38, respectively.

A diagnostic test is described by the ROC curve. By changing the cutoff point, one can change the number of true-positive diagnoses and true-negative diagnoses (i.e., to change the sensitivity and specificity).

Based on the present study, the cutoff point of ICON for the 12–13 year-olds in Chengdu should be adjusted to 29. However, the cutoff point setting relies on many factors including the opinion of experts regarding the need for treatment, available resources and the selection of a reliable gold standard. We invited two experts at first. Then we invited other 23 experts from domestic and board (the premier two experts were not included) to reevaluate the randomly selected 50 casts of this study. They all used the ITP. The scores of the 50 casts valued by the primer two experts, together with the 23 experts' scores, were collected and statistically analysed. We found that kappa value for the 25 experts and the mean indicated treatment point did have a few changes, but the best compromise between sensitivity and specificity in Chengdu, compared with the gold standard, may be found at a cutoff point of 29, where there were no changes. The ROC curve was adopted to determine the cutoff point. The best compromise between sensitivity and specificity in Chengdu, compared with the gold standard, were obtained with a cutoff point of 29.

With this new cutoff point, the sensitivity and the specificity were 0.88 and 0.83, respectively. The kappa value was 0.71, which was better than the previous kappa value for the international cutoff point. Lowering the cutoff to 29 would result in approximately 60% of individuals needing treatment, while the prevalence of malocclusion in children and adolescents in China is 67.8%.³⁴ However, this prevalence was only 26.9% with the international cutoff point of 43; thus, the cutoff point of 29 may be in line with Chinese conditions.

One possible explanation for the lower recommended cutoff point is cultural and ethnic differences. The photos used by ICON were of Caucasian children. The aesthetic component comprises the most important part of the ICON system.³⁵ According to a recent study (Borzabadi-Farahani 2010), with a cutoff of 43, nearly half of Iranian schoolchildren need orthodontic treatment; however, judging by the aesthetic component (IOTN), the value was 17.9%. Aesthetic assessment is to some extent impacted by culture and education, and the final ICON score with the adjusted cutoff point may therefore deviate from the international score. Some investigators have therefore recommended that, when ICON is used to determine treatment need locally, the cutoff point should be adjusted to optimal levels.¹

In one study,³⁶ the author found that the mean ICON value in the age group of 12–13 years was 35.8 and was the highest (39.9) in the age group of 18 years. Majority of the individuals in the age group of 12–13 years had a complexity grade of easy and mild and only a small proportion of individuals had difficult complexity grade. In the age group of 18 years, the majority of individuals had a mild (46.9%) complexity grade and easy and moderate complexity grade was observed 1.8 and 2.1 times less frequently, respectively. Only a very small proportion of individuals in this age group had a very difficult complexity grade. Overall, we conclude that the orthodontic treatment complexity grade tended to increase with age. In our study, the ages of the included 335 participants ranged from 12 to 13 years, which may lead to a relatively lower ICON score according to Dr Ilga Urtane's study.³⁶ A possible limitation of the present study is the use of a local panel of ICON examiners. A larger number of ICON raters would be necessary if we want to obtain a more accurate and a countrywide opinion.

The lower cutoff point may be due to the high prevalence of bimaxillary protrusion in our region. This condition has a relatively high prevalence in certain provinces of China.³⁷ The molar relationship is usually class I, with the incisor overjet and overbite being normal. The ICON scores of cases with bimaxillary protrusion are often lower; however, facial aesthetics are not well perceived. Some patients with this malocclusion have severe jaw deformities, which can be solved only by orthognathic surgery. The treatment complexity and need are not comparable with the ICON score in this scenario. It would be more appropriate to consider bimaxillary protrusion separately when applying ICON.

The treatment needs for boys and girls were the same, although their self-perceived needs may be different, with girls feeling more need for treatment than boys.³⁸

The distribution of the ICON scores in our study was not normal and did not obey the Gaussian distribution. A larger number of samples may lead to data that are in line with the Gaussian distribution.

ICON is recommended to be used with late mixed dentition and permanent dentition.⁶ We chose students in grade seven from 16 randomly selected middle schools in Chengdu with permanent dentition for the present study. For future studies in China, there is a need for a sample with wider age ranges to make the study findings applicable to a broader population. Multicentre clinical studies, especially studies including both northern and southern Chinese, are necessary if we want to make ICON a suitable index for assessing the orthodontic need in a larger geographic area of China.

We hope that the ICON system can be adopted in China based on this study, as this may help the provision of orthodontic treatment by guiding the limited Chinese health resources to those who need it most.

CONCLUSIONS

It was feasible to obtain good agreement with ICON among trained orthodontists. However, the international cutoff point (43) had poor sensitivity and specificity compared to the experts rating in southern China. Adjusting the cutoff point to 29 enhanced the sensitivity and specificity of ICONs in determining treatment need. Other studies are warranted to support or refute a lowering of ICON's cutoff point for treatment need in China. We may draw the following conclusions: when used to evaluate the treatment need of 12–13 year-olds in southern China, the international ICON cutoff value did not correspond well with Chinese orthodontists' opinions of the need for orthodontics. In this study, a lower cutoff value (29) had greater sensitivity and specificity with respect to expert orthodontists' perception of treatment need.

ACKNOWLEDGEMENTS

We thank all the children and teachers who agreed to take part in this study. We thank Professor Yi-jin Ren from the Department of Orthodontics of the University Medical Center Groningen, The Netherlands; Professor Yin Ding from the Department of Orthodontics, School of Stomatology, the Fourth Military Medical University; Professor Si-wei Yang from the Department of Orthodontics, Luzhou Medical College; Professors Song-jiao Luo and Mei-ying Zhao, Associate Professors Xi Du and Li Zhou and other teachers from the Department of Orthodontics of West China School of Stomatology, Sichuan University. The study would not have been possible without their participation.

- 1 Louwerse TJ, Aartman IH, Kramer GJ *et al*. The reliability and validity of the Index of Complexity, Outcome and Need for determining treatment need in Dutch orthodontic practice. *Eur J Orthod* 2006; **28**(1): 58–64.
- 2 Brook PH, Shaw WC. The development of an index of orthodontic treatment priority. *Eur J Orthod* 1989; **11**(3): 309–320.
- 3 Draker HL. Handicapping labio-lingual deviations: a proposed index for public health purposes. *Am J Orthod* 1960; **46**(4): 295–305.
- 4 Cons NC, Jenny J, Kohout FJ. *DAI: The Dental Aesthetic Index*. Iowa City: University of Iowa, 1986.
- 5 Richmond S, Shaw WC, O'Brien KD *et al*. The development of the PAR Index (Peer Assessment Rating): reliability and validity. *Eur J Orthod* 1992; **14**(2): 125–139.
- 6 Daniels C, Richmond S. The development of the index of complexity, outcome and need (ICON). *J Orthod* 2000; **27**(2): 149–162.
- 7 Borzabadi-Farahani A, Eslamipour F. The relationship between the ICON index and the dental and aesthetic components of the IOTN index. *World J Orthod* 2010; **11**(1): 43–48.
- 8 Borzabadi-Farahani A, Eslamipour F. Comparison between the index of complexity outcome and need and the index of orthodontic treatment need. Proceeding of the 86th Congress of the European Orthodontic Society; 15–19 June 2010. Portorož, Slovenia. London: European Orthodontic Society, 2010.
- 9 Younis JW, Vig KW, Rinchuse DJ *et al*. A validation study of three indexes of orthodontic treatment need in the United States. *Community Dent Oral Epidemiol* 1997; **25**(5): 358–362.
- 10 Beglin FM, Firestone AR, Vig KW *et al*. A comparison of the reliability and validity of 3 occlusal indexes of orthodontic treatment need. *Am J Orthod Dentofacial Orthop* 2001; **120**(3): 240–246.
- 11 Richmond S. *Evaluating Effective Orthodontic Care*. Cardiff: First Numeric Ltd, 2005.
- 12 Ngom PI, Diagne F, Dieye F *et al*. Orthodontic treatment need and demand in Senegalese school children aged 12–13 years. An appraisal using IOTN and ICON. *Angle Orthod* 2007; **77**(2): 323–330.
- 13 Firestone AR, Beck FM, Beglin FM *et al*. Validity of the Index of Complexity, Outcome, and Need (ICON) in determining orthodontic treatment need. *Angle Orthod* 2002; **72**(1): 15–20.

- 14 Ngom PI, Diagne F, Aidara-Tamba AW *et al*. Relationship between orthodontic anomalies and masticatory function in adult subjects. *Am J Orthod Dentofacial Orthop* 2007; **131**(2): 216–222.
- 15 Royston P. Approximating the Shapiro–Wilk W-test for non-normality. *Stat Comput* 1991; **2**(3): 117–119.
- 16 Lam FC, Longnecker MT. A modified Wilcoxon rank sum test for paired data. *Biometrika* 1983; **70**(2): 510–513.
- 17 Kerosuo H, Kerosuo E, Niemi M, *et al*. The need for treatment and satisfaction with dental appearance among young Finnish adults with and without a history of orthodontic treatment. *J Orofac Orthop* 2000; **61**(5): 330–340.
- 18 Metz CE. Basic principles of ROC analysis. *Sem Nucl Med* 1978; **8**(4): 283–298.
- 19 Beck JR, Shultz EK. The use of receiver operating characteristic (ROC) curves in test performance evaluation. *Arch Pathol Lab Med* 1986; **110**(1): 13–20.
- 20 Swets JA, Getty DJ, Pickett RM *et al*. Enhancing and evaluating diagnostic accuracy. *Med Decis Making* 1991; **11**(1): 9–17.
- 21 Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; **143**(1): 29–36.
- 22 Altman DG. *Practical Statistics for Medical Research*. London: Chapman & Hall, 1991.
- 23 Koochek AR, Yeh MS, Rolfe B *et al*. The relationship between Index of Complexity, Outcome and Need, and patients' perceptions of malocclusion: a study in general dental practice. *Br Dent J* 2001; **191**(6): 325–329.
- 24 Cooke M, Gerbert B, Gansky S *et al*. Assessment of the validity of HLD (CalMod) in identifying orthodontic treatment need. *Community Dent Oral Epidemiol* 2010; **38**(1): 50–57.
- 25 Fox NA, 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**(4): 225–230.
- 26 Weerakone S, Dhopatkar A. Clinical outcome monitoring program (COMP): a new application for use in orthodontic audits and research. *Am J Orthod Dentofacial Orthop* 2003; **123**(5): 503–510; discussion 510–511.
- 27 Savastano NJ, Firestone AR, Beck FM *et al*. Validation of the complexity and treatment outcome components of the index of complexity, outcome, and need (ICON). *Am J Orthod Dentofacial Orthop* 2003; **124**(3): 244–248.
- 28 Mirabelli JT, Huang GJ, Siu CH *et al*. The effectiveness of phase I orthodontic treatment in a Medicaid population. *Am J Orthod Dentofacial Orthop* 2005; **127**(5): 592–598.
- 29 Onyeaso CO, Begole EA. Relationship between index of complexity, outcome and need, dental aesthetic index, peer assessment rating index, and American Board of Orthodontics objective grading system. *Am J Orthod Dentofacial Orthop* 2007; **131**(2): 248–252.
- 30 Templeton KM, Powell R, Moore MB *et al*. Are the Peer Assessment Rating Index and the Index of Treatment Complexity, Outcome, and Need suitable measures for orthognathic outcomes? *Eur J Orthod* 2006; **28**(5): 462–466.
- 31 Onyeaso CO. Orthodontic treatment complexity and need in a group of Nigerian patients: the relationship between the Dental Aesthetic Index (DAI) and the Index of Complexity, Outcome, and Need (ICON). *J Contemp Dent Pract* 2007; **8**(3): 37–44.
- 32 Brown R, Richmond S. An update on the analysis of agreement for orthodontic indices. *Eur J Orthod* 2005; **27**(3): 286–291.
- 33 Kerosuo H, Al Enezi S, Kerosuo E *et al*. Association between normative and self-perceived orthodontic treatment need among Arab high school students. *Am J Orthod Dentofacial Orthop* 2004; **125**(3): 373–378.
- 34 Fu M, Zhang D, Wang B *et al*. The prevalence of malocclusion in China—an investigation of 25392 children. *Chin J Stomat* 2002; **37**(5): 371–373.
- 35 Ren Y, Boxum C, Sandham A. Patients' perceptions treatment need, and complexity of orthodontic re-treatment. *Eur J Orthod* 2009; **31**(2): 189–195.
- 36 Urtane I, Pugaca J, Liepa A *et al*. The severity of malocclusion and need for orthodontic treatment in correspondence with the age. *Stomatologija* 2006; **8**(2): 35–38.
- 37 Zhang J, Qiao M. Cephalometric analysis of the patients with bimaxillary protrusion in South China. *Hua Xi Kou Qiang Yi Xue Za Zhi* 2001; **19**(1): 32–34. Chinese.
- 38 Wheeler TT, McGorray SP, Yurkiewicz L *et al*. Orthodontic treatment demand and need in third and fourth grade schoolchildren. *Am J Orthod Dentofacial Orthop* 1994; **106**(1): 22–33.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivative Works 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0>