

Validity of two occlusal indices for determining orthodontic treatment needs of patients treated in a public university in Belo Horizonte, Minas Gerais State, Brazil

Validade de dois índices oclusais na determinação de necessidade de tratamento ortodôntico em pacientes tratados em uma universidade pública em Belo Horizonte, Minas Gerais, Brasil

Rodrigo Norremose Costa ¹
 Mauro Henrique Nogueira Guimarães de Abreu ¹
 Claudia Silami de Magalhães ¹
 Allyson Nogueira Moreira ¹

Abstract

The aim of the present study was to validate the dental aesthetic index (DAI) and index of complexity, outcome and need (ICON) based on the opinions of a panel of Brazilian orthodontists. A comparison of these two orthodontic treatment need indices was carried out based on the consensus of a panel of 20 experienced orthodontists. A set of 108 study casts representing the full spectrum of malocclusions was selected. A calibrated examiner scored the casts for both indices. The orthodontists individually rated the casts regarding the degree of orthodontic treatment need. The panel's mean rating of the need for treatment was used as the gold standard for evaluating the validity of the indices. The accuracy of the indices, as reflected in the area under receiver-operating characteristic curves, was high: DAI = 81.83% (95%CI: 71.21-92.44); ICON = 88.75% (95%CI: 78.57-98.92). Although the accuracy of the ICON was higher than that of the DAI, both indices are recommended for determining orthodontic treatment need in Brazil.

Malocclusion; ROC Curve; Validation Studies

Introduction

There is a high prevalence of malocclusion in the Brazilian population, with national values of 58.14% among children aged twelve years, varying from 53.7% to 64.1% in different regions ¹. In the city of Belo Horizonte, Minas Gerais State, there is a 61.9% prevalence among children aged 10 to 14 ². Considering the fact that malocclusion is a public health problem due to its prevalence and its impact on quality of life ^{2,3} as well as the limited treatment resources and the difficulty dentists have in diagnosing it and correctly assessing its severity, the use of a valid instrument for objectively evaluating criteria for recommending and prioritizing orthodontic treatment is necessary.

Validation and reliability tests have been carried out on a limited number of indices by a large group of orthodontists and have demonstrated that some indices precisely reflect the opinion of specialists, who, in turn, have demonstrated high degrees of inter-examiner agreement ^{4,5,6,7,8,9,10}. It is possible that the convergence of opinions among American specialists is due to the standardization of the educational process, the clinical curriculum of which is directed toward accreditation standards and criteria defined by the American Association of Orthodontists, which has led to a common notion of what constitutes the ideal occlusion ⁷. Validation studies on the dental aesthetic index (DAI) have been carried

¹ Faculdade de Odontologia, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.

Correspondence

M. H. N. G. Abreu
 Departamento de Odontologia Social e Preventiva, Faculdade de Odontologia, Universidade Federal de Minas Gerais.
 Rua Arthur Ferrari 55, apto. 401, Belo Horizonte, MG 30430-700, Brasil.
 maurohenriqueabreu@ig.com.br

out using different strategies for the obtainment of the gold standard in the USA and Australia^{8,11,12}. Regarding the index of complexity, outcome and need (ICON), validation studies have been carried out in the USA and Holland^{9,10,13}. Thus, the validation of these indices has only been carried out in three countries.

The use of a panel of orthodontists from a specific geographic region makes it difficult to make generalizations about the results. There is evidence that the country in which a specialist works has an effect on his/her assessment of orthodontic treatment needs¹⁴. Thus, the validity of an index depends on the panel of orthodontists that serve as the gold standard. Thus, validation studies on the DAI and ICON are needed, especially in developing countries.

The DAI is an occlusal screening index recommended by the World Health Organization (WHO) and has been used in an epidemiological survey carried out by the Brazilian Ministry of Health as well as in local studies in Brazil^{1,2,3}. The ICON offers the advantage of simplicity of use, as it has few characteristics to be measured and integrates both the aesthetic component and aspects related to the morbidity of malocclusion in the same equation⁴. However, the ICON has not been used in epidemiological studies in Brazil.

The aim of the present investigation was to validate the DAI and ICON based on the opinion of a panel of Brazilian orthodontists.

Material and methods

The DAI is a quantitative index that relates the perception of patients regarding dental aesthetics to the objective physical measurements of characteristics associated to malocclusions. To obtain the DAI score, ten occlusal characteristics are assessed based on socially defined occlusal standards of dental aesthetics. The score of the index ranges from 13 to approximately 80 and is categorized by cutoff points⁸. The ICON is a quantitative index used to determine the need for orthodontic treatment. To obtain the ICON score, each characteristic is measured and categorized based on a protocol that defines the score for each item. The values attributed to the five characteristics (four occlusal components and one aesthetic component) are multiplied by their respective coefficients (weights) and summed. The total score is a single value situated on an interval scale ranging from seven to 128 points⁴.

The validation of the indices of orthodontic treatment need was carried out in two main stages: the obtainment of the gold standard and the measurement of the indices.

The sample size was calculated to achieve a power of 75% and type II error of 25%, $\alpha = 0.05$ and $\delta^2 = 0.01$. Based on the kappa agreement of 0.81 ($\sigma^2 = 0.15$) obtained from the ICON validation study, it was determined that 108 cast models were needed^{9,15}.

The sample was selected from a set of 445 pairs of study casts in the archives of the Orthodontics Department of the Federal University of Minas Gerais. Each of the 455 casts was scored with the *Standardized Rating Scale of Dental Attractiveness* (SCAN) by an orthodontist with theoretical and practical knowledge of the use of this scale¹⁶. Each of the ten categories of the SCAN scale contained varieties of dentition types as well as a wide variety of malocclusions. Thus, the sample was representative of a wide range of types and severities of malocclusion.

The pairs of casts were distributed among the categories of the SCAN scale, representing ten different degrees of severity for the different dentition types and malocclusions. Mixed dentition accounted for 14.8% of the casts and permanent dentition accounted for 85.2%. The number of casts varied between severity categories such that there was a greater number of casts in the intermediate categories, as cases classified at the extremities of the treatment need scale (maximal need or no need) would not test the discriminatory capacity of the indices. For example, severity categories 5 and 6 each had 13 pairs of casts, whereas categories 1 and 10 each had six pairs. A randomized numeration was determined on an electronic spreadsheet and all pairs of casts were numerically recoded and labeled (Table 1).

The opinion of a panel of orthodontists was the reference for the gold standard regarding the recommendation for orthodontic treatment for the validation of the DAI and ICON indices. The panel was made up of 20 orthodontists from the city of Belo Horizonte, who fulfilled the following criteria: (a) master's or doctoral degree registered at the Brazilian Ministry of Education; (b) professor affiliated to a university located in Belo Horizonte; (c) a minimum of five years experience in private orthodontic practice in Belo Horizonte.

Following official contact with the respective universities, the specialists who fulfilled the inclusion criteria were contacted by telephone and asked to participate voluntarily in the study. The study and its objectives were explained and an account of the form of participation of the collaborators was given. All participants signed terms of informed consent. The study was conducted with professors who met the above-mentioned inclusion criteria. No professor refused to participate in the study. Emphasis was given regarding the need for dedication and precision in the

Table 1

Frequency of malocclusion classification in relation to classification using the *Standardized Rating Scale of Dental Attractiveness* (SCAN scale) from a set of 108 casts.

Classification of malocclusions	Classification using SCAN scale										Total (%)
	1	2	3	4	5	6	7	8	9	10	
Dentition											
Mixed	0	0	3	1	2	0	5	3	1	1	16 (14.8)
Permanent	6	11	9	11	11	13	7	9	10	5	92 (85.2)
Total	6	11	12	12	13	13	12	12	11	6	108 (100.0)

recording of the opinions. The participants were to appear on the dates defined for the initial and final evaluations of the study casts.

All 108 pairs of casts were assessed for the obtainment of the opinion of the panel regarding the need for orthodontic treatment. The evaluations were scheduled on the days when the participants were at the universities. Between one and four sessions were held, depending on the day of the week and time available in the professors' chronograms of activities. The casts were arranged in order following the previously randomly attributed numeration in rooms offered by the universities in order to facilitate access for the participants.

The evaluation forms contained the number corresponding to the case, the coded identification of the orthodontist, necessary instructions and a numerical scale to be marked. The specialists were instructed to rate the occlusion in each pair of casts on a seven-item Likert scale, on which 1 represented "no or minimal recommendation for treatment" and 7 represented "very highly recommended for treatment". The participants worked at their own pace and within their own time constraints^{7,8,9}. Each study model received a single score on a seven-item Likert scale, obtained from the mean values given by all specialists.

Each specialist was then asked to use the same seven-point scale to define the value at or above which he/she thought treatment would be recommended. The "recommended treatment point" (RTP) was obtained for each of the casts. The mean of the 20 RTP values was used as the single cutoff point, thereby reflecting the opinion of the specialists in a dichotomous manner for each model. Casts for which the mean opinion was lower than the mean RTP were classified as "treatment not recommended", whereas those equal to or higher than the RTP were classified as "treatment recommended"⁸.

To assess intra-examiner and inter-examiner agreement, a subset of forty pairs of casts was randomly selected⁸ for re-evaluation by the 20 specialists. The examinations were repeated with a minimal interval of three weeks between sessions for each index. Inter-examiner agreement in 20 randomly defined combinations of pairs of specialists and intra-examiner agreement for the 20 specialists were determined using the quadratic weighted kappa statistic.

The results of the agreement coefficients for the seven-item Likert scale (obtained from the kappa) were accompanied by the lower bound of the 95% confidence interval (95%CI). Given its equivalence in the evaluation of levels of agreement for the sensitivity and specificity model and agreement coefficients, the scale proposed by Fleiss and modified by Cicchetti was used for the assessment of the level of agreement^{17,18,19}.

A calibrated orthodontist (R.N.C.; intraclass correlation coefficients for DAI ≥ 0.978 and ICON ≥ 0.931) measured each study cast. The determination of the ICON and DAI was made on the set of 108 pairs of casts six weeks following the last calibration session for each index. The measurements on each cast were recorded on specific charts. The values of the indices attributed to the casts by the researcher, blindly, were compared to the gold standard. The cutoff points were 31 for the DAI and 43 for the ICON. The DAI was measured based on an adaptation for mixed dentition, as proposed by the WHO, which considers variations stemming from the development of the occlusion in the components absent visible teeth, spacing of anterior segments and anterior-posterior molar relation.

Specificity, sensitivity and accuracy were evaluated. Positive and negative predictive values were calculated from the sensitivity and specificity values of the indices and the prevalence of the need for orthodontic treatment in the local population².

The accuracy of each index was assessed through analysis of the receiver-operating characteristic (ROC) curve and the performances were compared through the determination of the statistical significance of the difference between the areas of the two ROC curves using the non-parametric method proposed by DeLong et al.²⁰

All mean values were tabled on an electronic spreadsheet for the purpose of statistical analysis³. The data entered were verified using the original charts with the aid of a digital text reader program (Free Natural Reader. Natural Soft.Ltd., Richmond, Canada). The calculations were carried out either on the spreadsheet itself or on a statistical program (Stata/SE for Windows, version 8.0. Satat Corp., College Station, USA). The present study and term of informed consent received approval from the Ethics Committee of the Federal University of Minas Gerais (process n°. ETIC 068/06). The study casts used in this investigation pertain to the Department of Pediatric Dentistry and Orthodontics and their use was authorized.

Results

There was variability in the opinions of the specialists regarding the 108 pairs of casts rated on a seven-point scale. The mean was 4.795 (± 0.76), with a minimum of 3.25, maximum of 6.44 and median value of 5.13. The coefficient of variation was 15.85%. The RTP used to dichotomize the opinion regarding treatment need determined by the 20 specialists was 2.90. The determination of this cutoff point classified 86.11% of the cases as "with need" and 13.89% as "without need" for treatment.

Table 2 displays the inter-examiner agreement of the specialists. The level of agreement ranged from poor to good. The level of intra-examiner agreement ranged from poor to excellent. The lowest level of intra-examiner agreement was obtained by examiners 2 and 10. Only one specialist achieved an excellent level of agreement²¹ (examiner 11) (Table 3).

The measurement of the 108 pairs of casts based on the DAI and ICON criteria resulted in a mean value of 34.033 (-8.422) and 57.565 (-23.509), respectively. For the DAI, this value was below the central value of the index scale, which

Table 2

Sampling of inter-examiner agreement by pairs. Belo Horizonte, Minas Gerais State, Brazil.

Pairs of examiners	Kw ² (lower limit of 95%CI)
1 × 2	0.631 (0.520)
2 × 4	0.565 (0.450)
3 × 16	0.588 (0.463)
4 × 14	0.502 (0.338)
5 × 20	0.299 (0.212)
6 × 12	0.651 (0.537)
7 × 11	0.744 (0.662)
8 × 18	0.499 (0.396)
9 × 15	0.256 (0.144)
10 × 17	0.644 (0.553)
11 × 3	0.698 (0.602)
12 × 1	0.577 (0.460)
13 × 6	0.662 (0.555)
14 × 19	0.528 (0.379)
15 × 13	0.751 (0.663)
16 × 9	0.423 (0.250)
17 × 5	0.351 (0.235)
18 × 10	0.692 (0.606)
19 × 8	0.574 (0.462)
20 × 7	0.548 (0.466)

Note: measurements made by sampling.

Table 3

Intra-examiner agreement for values of the opinion of the specialists. Belo Horizonte, Minas Gerais State, Brazil.

Examiner	Kw ² (lower limit of 95%CI)
1	0.826 (0.724)
2	0.497 (0.217)
3	0.808 (0.692)
4	0.745 (0.529)
5	0.711 (0.459)
6	0.774 (0.640)
7	0.846 (0.734)
8	0.582 (0.406)
9	0.645 (0.483)
10	0.558 (0.370)
11	0.859 (0.746)
12	0.835 (0.710)
13	0.822 (0.690)
14	0.710 (0.537)
15	0.697 (0.531)
16	0.859 (0.744)
17	0.758 (0.619)
18	0.820 (0.678)
19	0.604 (0.421)
20	0.769 (0.639)

Note: measurements made by sampling.

ranges from 13 to approximately 80. Using the cutoff point of 31, originally established by the authors of the DAI, 65 pairs of casts (60.19%) were classified as having treatment need, compared to the 93 pairs of casts (86.11%) classified by the specialists as having treatment need. The DAI correctly classified 70.37% of the pairs of casts. For the ICON, this value ranges from 16 to 111 on a scale the score of which ranges from approximately seven to 128. Using the cutoff point of 43 originally established by the authors of the ICON, the index classified 70 (64.81%) pairs of casts as having treatment need, whereas the panel classified 93 (86.11%) pairs of casts in this same category. The ICON correctly classified 76.85% of the pairs of casts. The sensitivity, specificity, positive predictive values and negative predictive values for DAI and ICON are presented in Table 4. Predictive values were calculated considering the prevalence of treatment need in the city of Belo Horizonte (26.4%)².

The likelihood ratios for the DAI and ICON were 5.08 and 11.12, respectively. The area under the ROC curve was 81.83% (95%CI: 71.21-92.44) for DAI and 88.75% for ICON. The comparison of accuracy measures (area under the ROC curve)

of each index revealed that the ICON appears to have a better discriminatory power regarding the need for treatment than the DAI, as the area under the ROC was significantly greater ($p = 0.0435$) (Figure 1).

A cutoff point of 30.4 led to an increase in the sensitivity of the DAI to a reasonable level (72%; 95%CI: 61.6-80.6), while specificity remained good, with the same value of 86.7% (95%CI: 58.4-97.7). A cutoff point of 40 led to an increase in the sensitivity of the ICON to 82.8% (95%CI: 73.3-89.6), although the specificity diminished to 86.7% (95%CI: 58.4-97.7).

Discussion

The assessment of specialists for the definition of treatment need among a set of orthodontic casts has demonstrated variable results in the literature¹³. A similar finding was identified in the present study. The set of casts was selected in such a way that the cases were distributed throughout the treatment need scale, with a greater frequency of cases of intermediate need and a lower proportion of cases at the extremes of the scale. This

Table 4

Frequency of orthodontic treatment need comparing diagnosis performed by panel opinion (gold standard) and dental aesthetic index (DAI) and index of complexity, outcome and need (ICON).

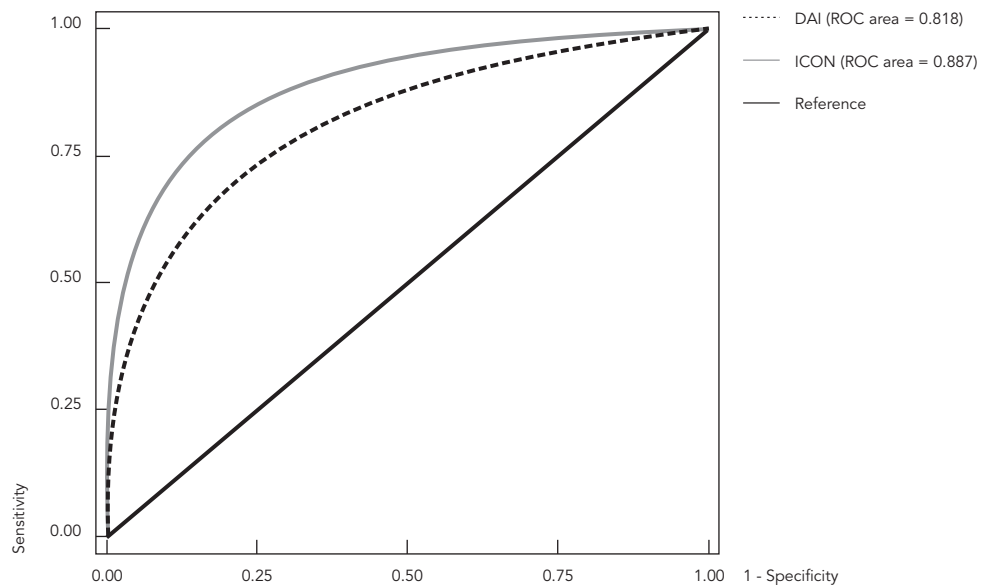
Index	Gold standard		Total (%)
	With need	Without need	
DAI *			
With need	69	1	70 (64.81)
Without need	24	14	38 (35.19)
Total (%)	93 (86.11)	15 (13.89)	108 (100.00)
ICON **			
With need	63	2	65 (60.19)
Without need	30	13	43 (39.81)
Total (%)	93 (86.11)	15 (13.89)	108 (100.00)

* DAI (95%CI): accuracy = 81.83 (71.21-92.44); sensitivity = 67.74 (57.10-76.90); specificity = 86.67 (58.40-97.70); positive predictive value = 64.57 (32.99-92.30); negative predictive value = 88.21 (79.15-92.18); Likelihood ratio of a positive test = 5.08;

** ICON (95%CI): accuracy = 88.75 (78.57-98.92); sensitivity = 74.19 (63.90-82.50); specificity = 93.33 (66.00-99.70); positive predictive value = 79.96 (40.27-99.00); negative predictive value = 90.98 (83.60-94.08); Likelihood ratio of a positive test = 11.12.

Figure 1

ROC curve for the dental aesthetic index (DAI) and index of complexity, outcome and need (ICON). Belo Horizonte, Minas Gerais State, Brazil.



distribution enabled a better evaluation of the discriminatory power of the indices in the range of greater decision-making difficulty. The mean obtained from the panel of specialists was higher than the central value of the opinion scale (four). This may be caused by the use of a sample that resulted in a large number of casts with treatment needs in order to represent a wide variety of types of malocclusion. The variability in the opinion of the specialists regarding the need for orthodontic treatment may be considered a reflection of the difficulty in defining objective signs of malocclusion. As treatment need is a measure of deviation from an arbitrary ideal, there should be a consensus regarding this ideal^{5,7,22,23}.

RTP values 3.25 and 3.53 have been reported in previous studies using the same methodology^{7,8,9}. The studies by Beglin et al.⁸ and Firestone et al.⁹, who respectively assessed 156 and 170 study casts, found more balanced classification percentages. In both studies, the panel classified 64% of the cases in the treatment need category. For the validation of the DAI and ICON, respectively, these two studies obtained better reliability values for the panel of specialists. Beglin et al.⁸ found an inter-examiner agreement of $k_w = 0.81$ (lower bound of the 95%CI: $k_w \geq 0.81$) and intra-examiner agreement of $k_w = 0.92$ ($k_w \geq 0.90$). Firestone et al.⁹ found an inter-examiner agreement of $k_w = 0.81$ ($k_w \geq 0.81$) and an intra-examiner agreement of $k_w = 0.92$ ($k_w \geq 0.90$).

In the present investigation, the removal of the most discordant specialists who determined extreme values altered the variability values, while the mean opinion and RTP values were not altered to the point of affecting the number of cases classified as having treatment need, which remained 93 (data not shown). Therefore, the original panel of 20 specialists was maintained as the gold standard of opinion regarding the need for orthodontic treatment.

Agreement regarding the opinion of the specialists was similar to that reported by Richmond et al.²², who found intra-examiner agreement values ranging from 0.54 to 0.97 for aesthetic need and 0.12 and 0.89 for dental health need. Similar levels of agreement were described by Richmond & Daniels¹⁴, who report values for the lower bound of 95%CI of the kappa coefficient for intra-examiner agreement of 0.58 and 0.55 for need based on aesthetics and health, respectively.

Younis et al.⁷ suggest that the lack of agreement among specialists whose training in orthodontics occurred outside the USA may be due to a lack of common curricula and accreditation standards. In the present study, the DAI classified fewer cases with a treatment need than the panel of specialists. It is possible that the opinion of the

panel formed by these Brazilians to treat more cases demonstrates a geographic difference in decision making. However, this is an issue that requires further investigation.

Considering the 26.4% prevalence rate in the city of Belo Horizonte, a good screening index for orthodontic treatment need would be that with a high positive predictive value, which would offer greater specificity and a lower number of false positives. False positives can lead to emotional and financial burdens, whereas false negatives generally do not result in a serious problem, as malocclusion does not offer a serious health risk. Thus, a higher specificity value should be prioritized and the sensitivity value should be balanced in order not to excessively increase the number of false negatives.

Comparing the values related to the validity of the DAI and ICON using their original cutoff points, the ICON better reflected the opinion of the panel of specialists. A possible cause for the lower validity of the DAI is its failure to consider characteristics related to the impact of malocclusion on oral health^{24,25}. Thus, cases considered by the panel as needing treatment due to factors not only related to aesthetics but to the concept of ideal functional occlusion may not have been classified by the DAI. As the ICON considers both aesthetic and functional characteristics, it has a greater chance of reflecting the opinion of the panel of specialists.

Another aspect related to the construction of the DAI that may limit the generalization of its use is how the cutoff points were obtained and the recommendation of its validation in both developed and developing countries. The cutoff points were obtained through the correlation between the frequency distribution of the numeric value of the DAI and the distribution of categories of malocclusion severity according to the report from the American National Center for Health Statistics. The degrees of treatment need were deduced from a comparison of the accumulated percentage among the samples^{8,26,27}. The recommendation of the generalization of its validity was based on a correlation between the opinion of students in four countries regarding the aesthetic impact of malocclusion, as assessed in 25 photographs²⁸.

The use of a panel of orthodontists from a specific geographical region means it is not possible to make generalizations about the results. There is evidence that the country in which a specialist works has an effect on his/her evaluation regarding the need for orthodontic treatment²³. Thus, the validity of an index depends upon the panel of specialists that serve as the gold standard. The quality of this validation depends on the mini-

mization of errors of different origins. The low degree of reliability of the validating panel makes the gold standard very subjective. As long as there are no more objective, evidence-based standards for the decision of orthodontic treatment, the opinion of specialists continues to be the best reference. Considering the variance in the objective opinion of the panel, indices may be useful educational tools for improving the agreement among professionals regarding the need for orthodontic treatment.

Although the ICON achieved significantly greater accuracy than the DAI, both may be considered valid for use as screening measures of the normative need for orthodontic treatment in epidemiological studies. In situations involving the use of these indices on a real population, a number of issues should be raised regarding their validity: whether the values measured by the index are altered during occlusal development; whether the validity of the measurements determined in study casts corresponds to those in actual patients; and whether the decision regarding the need for treatment corresponds to an actual normative need or may be related to the perception of the patient or his/her culture.

According to Summers²³, the development of the occlusion may alter the values measured by an index. A substantial change in occlusion was observed in the critical period between 12 and 16 years of age, which reduced the value measured by the DAI in this period by an average of 5.1 points. Such alterations depend on what components are measured by the index. In the case of the DAI, for instance, the components that most contributed to this temporal alteration were overjet, crowding, anterior gaps, midline diastema and inter-molar relation. These components are very mutable characteristics in the development of the occlusion. As each index uses measures of different occlusal characteristics in its calculation, the impact of changes during the development of the occlusion on the final value may not be the same for different indices²⁹. In the present study, the DAI was measured based on an adaptation for mixed dentition, as proposed by the WHO. In order to minimize the effect of occlusal development on the reduction of the DAI score, some physiological characteristics of the mixed dentition (especially in the “ugly duckling phase”) should be discarded or modified in the evaluation: midline diastema, anterior maxilla irregularity, anterior mandible irregularity and anterior open bite. Further studies are needed to determine the magnitude of these specific temporal alterations for each index and establish the most adequate correction for each case: adaptations to the index structure

or weights of specific categories or even a correction in the final value, depending on the phase of the dentition. In the present study, 16 cases (14.8%) with mixed dentition were used. Considering the temporal effect, the score for these cases should be reduced, which would potentially reduce the number of cases classified by the DAI as needing treatment.

The reliability and validity of evaluations made with study casts in comparison to clinical situations may be questioned. These aspects require further study, given the importance of the comparison of retrospective data (which are generally obtained from study casts) and current, clinically obtained data. Objective measurements from study casts are believed to offer high reliability in comparison to those obtained in clinical exams. Substantial agreement has been identified when comparing the dental health component of the index of orthodontic treatment need (IOTN) between study casts and clinical examination. Substantial agreement has also been found when comparing the aesthetic component of the IOTN (SCAN scale) between study casts and clinical examination³⁰. Although studies of this nature have not been carried out for the DAI and ICON, Beglin et al.⁸ state that the majority of occlusal indices may be administered on study casts, as general diagnostic tests do not require the presence of the patient, since the source of the data (blood, tissue or image) has been acquired.

The adjustment of the cutoff points affected the sensitivity and specificity of both indices. However, based on the confidence intervals, the sensitivity and specificity were similar to the original values for these indices, which signify that the DAI and ICON are applicable to the Brazilian population studied.

It should be pointed out that there was a higher prevalence of orthodontic treatment need in the present study than that considered for the Brazilian population as a whole, which hinders the generalization of the results of this study^{1,2}. Validation studies are generally carried out at orthodontic clinics^{8,9}. It is likely that validation studies involving representative samples of the general population would achieve different results from those of the present study. However, an effort was made to minimize this aspect by selecting study casts from all categories of the SCAN scale. Moreover, indices tend to be administered by clinicians in epidemiological surveys rather than specialists. It is therefore important to carry out validation studies involving general dentists.

The determination of malocclusion in public health allows the establishment of public poli-

cies directed at oral health and the organization of public dental services that offer orthodontic treatment. The use of occlusal indices for the determination of orthodontic treatment need generally has limitations inherent to the procedure – from the determination of the gold standard to the incorporation of the patient's perception regarding treatment need. As well as appearing structurally different, the DAI and ICON have significant differences regarding their validity. Due to its greater accuracy and the fact that it addresses important health-related components along with aesthetic aspects, the ICON could replace the DAI as an index for orthodontic treatment need. Other characteristics, such as simplicity, temporal reliability and duration of the clinical exam, should be considered when choosing indexes for oral surveys.

Conclusion

Although the accuracy of the ICON was greater than that of the DAI, both indices are recommended for determining orthodontic treatment needs in Brazil.

Resumo

O objetivo do presente estudo foi validar os índices dental aesthetic index (DAI) e index of complexity, outcome and need (ICON) a partir da opinião de um painel de ortodontistas brasileiros. A comparação desses dois índices de necessidade de tratamento ortodôntico foi feita baseada no consenso de um painel de 20 experientes ortodontistas. Um conjunto de 108 modelos de estudo representando uma grande variedade de tipos de malocclusão foi selecionado. Um examinador calibrado mensurou os modelos para ambos os índices. Os ortodontistas, de forma individual, avaliaram a necessidade de tratamento ortodôntico de cada modelo. O valor médio do painel de ortodontistas em relação à necessidade de tratamento foi utilizado como padrão-ouro para validação dos índices. Os valores de acurácia dos índices, medidos pela curva de característica de operação do receptor, foram altos: DAI = 81,83% (IC95%: 71,21-92,44); ICON = 88,75% (IC95%: 78,57-98,92). Embora a acurácia do ICON tenha sido mais alta do que a do DAI, ambos os índices são recomendados para determinar a necessidade de tratamento ortodôntico no Brasil.

Malocclusão; Curva ROC; Estudos de Validação

Contributors

R. N. Costa was responsible for the study design, statistical analysis and interpretation of data, as well as organizing and writing the paper. M. H. N. G. Abreu was responsible for analysis and interpretation of the data and drafted the paper. C. S. Magalhães was responsible for the study design, supervision, and data collection, and also assisted with the analysis, interpretation of data and with writing the paper. A. N. Moreira was responsible for the study design, supervision, and data collection, and also assisted with the analysis, interpretation of data and writing the paper.

Acknowledgments

To the Graduate Studies Coordinating Board (CAPES) for financial support.

References

1. Ministério da Saúde. Projeto SB 2003: condições de saúde bucal da população brasileira 2002-2003. Resultados principais. Brasília: Ministério da Saúde; 2004.
2. Marques LS, Barbosa CC, Ramos-Jorge ML, Pordeus IA, Paiva SM. Prevalência da maloclusão e necessidade de tratamento ortodôntico em escolares de 10 a 14 anos de idade em Belo Horizonte, Minas Gerais, Brasil: enfoque psicossocial. *Cad Saúde Pública* 2005; 21:1099-106.
3. Peres KG, Traebert ESA, Marcenes W. Diferenças entre auto percepção e critérios normativos na identificação de oclusopatias. *Rev Saúde Pública* 2002; 36:230-6.
4. Daniels C, Richmond S. The development of the index of complexity, outcome and need (ICON). *Br J Orthod* 2000; 27:149-62.
5. Richmond S, Shaw WC, O'Brien KD, Buchanan IB, Jones R, Stephens SD, et al. The development of the PAR index (peer assessment rating): reliability and validity. *Eur J Orthod* 1992; 14:125-39.
6. Deguzman L, Bahiraei D, Vig KW, Vig PS, Weyant RJ, O'Brien K. The validation of the Peer Assessment Rating index for malocclusion severity and treatment difficulty. *Am J Orthod Dentofacial Orthop* 1995; 107:172-6.
7. Younis JW, Vig KWL, Rinchuse DJ, Weyant RJ. A validation study of three indexes of orthodontic treatment need in the United States. *Community Dent Oral Epidemiol* 1997; 25:358-62.
8. Beglin FM, Firestone AR, Vig KWL, Beck FM, Kuthy RA, Wade D. A comparison of the reliability and validity of 3 occlusal indexes of orthodontic treatment need. *Am J Orthod Dentofacial Orthop* 2001; 120:240-6.
9. Firestone AR, Beck FM, Beglin FM, Vig KWL. Validity of the index of complexity, outcome and need (ICON) in determining orthodontic treatment need. *Angle Orthod* 2002; 72:15-20.
10. Lowerse TJ, Aartman IHA, Kramer GJC, Prah Andersen B. The reliability and validity of the index of complexity, outcome and need for determining treatment need in Dutch orthodontic practice. *Eur J Orthod* 2006; 28:58-64.
11. Keay PA, Freer TJ, Basford KE. Orthodontic treatment need and the dental aesthetic index. *Aust Orthod J* 1993; 13:4-7.
12. Cons NC, Kohout FJ, Jakobsen J. Predicting handicapping malocclusion using the Dental Aesthetic Index (DAI). *Int Dent J* 1993; 43:128-32.
13. Cooke M. Assessment of the validity of the index of complexity, outcome and need (ICON), and handicapping labio-lingual deviations with California modifications (HLD CalMod). *Am J Orthod Dentofacial Orthoped* 2006; 130:806.
14. Richmond S, Daniels CP. International comparisons of professional assessments in orthodontics: Part 1: treatment need. *Am J Orthod Dentofacial Orthop* 1998; 113:180-5.
15. Montgomery DC, Runger GC. Applied statistics and probability for engineers. 2nd Ed. New York: John Wiley; 1999.
16. Evans R, Shaw W. Preliminary evaluation of an illustrated scale for rating dental attractiveness. *Eur J Orthod* 1987; 9:314-8.
17. Fleiss JL. Statistical methods for rates and proportions. 2nd Ed. New York: John Wiley & Sons; 1981.
18. Cicchetti DV. Methodological commentary: the precision of reliability and validity estimates revisited: distinguishing between clinical and statistical significance of sample size requirements. *J Clin Exp Neuropsychol* 2001; 23:695-700.
19. Cicchetti DV, Volkmar F, Sparrow SS, Cohen D, Fermanian J, Rourke BP. Assessing the reliability of clinical scales when the data have both nominal and ordinal features: proposed guidelines for neuropsychological assessments. *J Clin Exp Neuropsychol* 1992; 14:673-86.
20. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a non-parametric approach. *Biometrics* 1988; 44:837-45.
21. Landis JR, Koch CG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33:159-74.
22. Richmond S, O'Brien KD, Roberts CT, Andrews M. Dentists variation in the determination of orthodontic treatment need. *Br J Orthod* 1994; 21:65-8.
23. Summers CJ. The occlusal index: a system for identifying and scoring occlusal disorders. *Am J Orthod Dentofacial Orthop* 1971; 59:552-67.
24. Danyluk K, Lavelle C, Hassard T. Potential application of the dental aesthetic index to prioritize the orthodontic service needs in a public funded dental program. *Am J Orthod Dentofacial Orthop* 1999; 116:279-86.
25. Otuemy OD, Noar JH. Variability in recording and grading the need for orthodontic treatment using the handicapping malocclusion assessment record, occlusal index and dental aesthetic index. *Community Dent Oral Epidemiol* 1996; 24:222-4.
26. Jenny J, Cons NC, Kohout FJ, Jakobsen J. Predicting handicapping malocclusions using the dental aesthetic index (DAI). *Int Dent J* 1993; 43:128-32.
27. Jenny J, Cons NC. Establishing malocclusion severity levels on the dental aesthetic index (DAI) scale. *Aust Dent J* 1996; 41:43-6.
28. Cons NC, Jenny J, Kohout FJ, Songpaisan Y, Jotikastira D. Utility of the dental aesthetic index in industrialized and developing countries. *J Public Health Dent* 1989; 49:163-6.
29. Tarvit DJ, Freer TJ. Assessing malocclusion: the time factor. *Br J Orthod* 1998; 25:31-4.
30. Buchanan IB, Downing A, Stirrups DR. A comparison of the index of orthodontic treatment need applied clinically and to diagnostic records. *Br J Orthod* 1994; 21:185-8.

Submitted on 24/Mar/2010

Final version resubmitted on 27/Aug/2010

Approved on 29/Nov/2010