Assessment of malocclusion in the permanent dentition: reliability of intraoral measurements

Maja Ovsenik
Department of Orthodontics, Medical Faculty, University of Ljubljana, Slovenia

SUMMARY Malocclusion assessment methods are based on registrations and measurements made on study casts, which requires that impressions be taken. In addition to being costly and time consuming, this process can be unpleasant for children and adolescents. Therefore, the aim of this study was to evaluate the reliability of intraoral measurements that compute a malocclusion index score to determine malocclusion severity in permanent dentition.

The research was a part of a longitudinal study of 530 3-year-old children. In Slovenia at 14 years of age [mean = 14.8 years, standard deviation (SD) = 0.2], a cohort of 92 children (39 boys and 53 girls) were selected at random in a cross-sectional study. Quantitative registrations of space and occlusal anomalies were performed intraorally as well as on study casts. Kappa (κ) statistics were used to evaluate agreement between clinical and study cast malocclusion assessment. Systematic bias of measurements was tested using Wilcoxon's signed-rank test.

The results showed almost complete agreement between the two measurements for anterior crossbite, anterior open bite (AOB), transverse occlusion of the posterior teeth, and crowding (κ = 0.81–1); excellent reliability for rotation of incisors and canines, for buccal segment relationship, overjet, and axial inclination of teeth (κ = 0.61–0.80); and for the remainder of the traits the reliability was moderate: vestibular canine eruption, overbite, and midline deviation (κ = 0.41–0.60). Intraorally small, but statistically significant (P < 0.05) lower scoring of axial inclination of teeth was identified. Overall classification into severity grades, based on total malocclusion score, showed excellent agreement between the two methods (κ = 0.84), without statistically significant bias.

Malocclusion assessment, recorded and measured intraorally, is as reliable as assessment on study casts. The proposed method can be used in screening, in epidemiological studies, and in clinical orthodontic assessment.

Introduction

There is considerable international interest in guidelines for the screening of children for orthodontic treatment (Solow, 1995). New malocclusion indices and indices of treatment have been developed and tested in many countries, and their deficiencies are well recognized (Brook and Shaw, 1989; Espelander et al., 1992; Burden et al., 2001).

Methods that describe, evaluate, and classify occlusion can basically be divided into qualitative and quantitative methods (Thilander et al., 2001; Ovsenik et al., 2004) and are designed either for study cast measurements, clinical use, or both (Table 1).

Eismann (1974, 1980) developed a method for evaluating the efficiency of orthodontic treatment and treatment need in the permanent dentition, based on the determination of morphological criteria in a method analogous to that used by Björk et al. (1964). In order to assess malocclusion in the early dental development period, the Eismann method was modified for the primary and mixed dentitions (Farčnik et al., 1985, 1988) and used in Slovenia in a longitudinal study as an indicator of interceptive treatment results (Korpar et al., 1994). The modified Eismann index (Farčnik et al., 1985, 1988) has also been found to be a valid and reliable method for assessing malocclusion severity in everyday clinical work (Ovsenik and Primožič, 2007).

Both these methods (Eismann, 1974; Farčnik et al., 1985) are performed on study casts. However, preparation of study casts requires that impressions be taken, and this is often unpleasant, especially for children and adolescents. In addition, the procedure itself can be costly, and measurements have proved to be complicated and time consuming in daily use (Solow, 1995; Ovsenik et al., 2004, 2007).

In clinical orthodontics, malocclusion assessment remains problematic. Index scores have been shown to have acceptable reliability (Brook and Shaw, 1989; Richmond et al., 1992) when measured on casts. Only the study by Keeling et al. (1996) and Ovsenik et al. (2004, 2007) report the reliability of scoring components of malocclusion in the clinical setting. Obtaining casts involves clinical and laboratory procedures and is thus a costly and time-consuming method for assessing malocclusion. Conversely, performing the measurements on casts is more pleasant for the examiner, who can manually handle the cast while sitting at a table under excellent lighting and using a
measuring device (protractor, gauge) designed specifically for the purpose. Although there are certain advantages and conveniences in making measurements on casts, the obtaining of casts may not be possible under many field conditions (very young children, taking impressions, costs, and time) and thus for consistency the assessments are limited to direct observations (Ovsenik et al., 2004).

It has been established in a previous study (Ovsenik et al., 2004) that malocclusion assessment in the period of the early mixed dentition based on intraoral measurements is as reliable as assessment carried out on study casts, and is thus the method of choice to be used in malocclusion assessment in epidemiological studies, in screening, and in clinical orthodontic assessment. Application of the proposed method in clinical orthodontics is preferred, as it requires less clinical time when compared with assessments based on study cast measurements.

Comparison between intraoral and study cast measurements in the assessment of malocclusion in the permanent dentition according to the Eismann index has not yet been evaluated. Therefore, the aim of the present study was to assess the reliability of occlusal traits, recorded and measured intraorally to compute the malocclusion score, and to determine malocclusion severity grade in the permanent dentition.

### Subjects and methods

The research was a part of the longitudinal study by Farčnik et al. (1986) in Slovenia on a sample of 530 3-year-old children. A cohort of 92 children (39 boys and 53 girls), at the age of 14 years (mean = 14.8 years, SD = 0.18), was selected at random to assess the malocclusion severity score, classified into grades of severity.

Clinical examinations were performed by the author, an experienced orthodontist trained in the use of the index. During the intraoral examination, measurements of 15 morphological signs were carried out (Figure 1). Impressions were then taken of the upper and lower dentitions for study cast measurements, which were repeated after an interval of 1 month. For each set of measurements, registrations were carried out according to Eismann (1974, 1980). For measurements of linear dimensions, a metric ruler (Zürcher modell, Dentaurum 042-751, Ispringen, Germany), accurate to 1/10 mm, was used, while angles were measured with a protractor designed by Eismann (1974) for measuring the rotation of incisors and canines (Figure 2) and the axial inclination of the teeth.

Intra-arch assessment involved measurement of incisor crowding, and rotation of the incisors, and axial inclinations of the teeth. For inter-arch measurements, overbite, anterior open bite (AOB), overjet, reverse overjet, anterior crossbite, and buccal segment relationships were recorded.

All morphological signs, measured intraorally as well as on study casts and expressed in millimetres and degrees, were weighted and scored against the evaluation table for each subject (Eismann, 1974). The weighted sum of recorded occlusal traits thus represented the total malocclusion index score—where the first was measured intraorally and the second on study casts. The overall malocclusion scores were categorized according to Eismann (1974, 1980) in terms of mild (1–15), moderate (16–45), severe (46–65), and very severe (over 66).

### Statistical analysis

Kappa (κ) statistics were used to evaluate the agreement observed between intraoral and study cast individual measurements. K values equal to 0 represent agreement equivalent to that expected by chance, while 1 represents perfect agreement. In accordance with Landis and Koch (1977), the following κ interpretation scale was used: poor...
Wilcoxon’s signed-rank test was used for statistical analysis of the bias between clinical and study cast malocclusion assessment and a non-parametric test because of non-normal distribution of data. For analysis, the Statistical Package for Social Sciences, Windows version 10.1 (SPSS Inc., Chicago, Illinois, USA) was used.

Results

The results for the clinical and study cast malocclusion assessment are summarized in Table 2. \( \kappa \) statistics indicated agreement for (AOB), anterior crossbite, transverse occlusion of posterior teeth, and crowding. There was excellent agreement for rotation of incisors and canines, buccal segment relationship, overjet, and axial inclination of the teeth. Moderate agreement was found
for vestibular eruption of the canine, overbite, and midline deviations.

Systematic bias was found for axial tooth inclination, which tended to be scored slightly worse intraorally. As can be seen from Table 2, despite statistical significance between the two methods, the measurements were in most cases equal (e.g. crossbite was scored equally in all 92 cases).

The classification of malocclusion scores into four grades of severity according to intraoral and study cast assessment is shown in Table 3.

In eight patients, the intraorally recorded score grade was lower, in eight patients higher, and in the remaining 76 patients the scores were equal. $\kappa$ statistics for the agreement between the two methods yielded a value of 0.84 (excellent agreement). The analysis of bias using Wilcoxon’s signed-rank test revealed no statistically significant difference ($Z = -0.057, P = 0.95$) between the malocclusion severity grade obtained intraorally or on the study casts.

**Discussion**

Malocclusion assessment methods differ not only in the choice of the morphological or functional criteria used but also in the mode of evaluation, which can be performed on study casts (Summers, 1971; Eismann, 1974, 1977, 1980; Farčnik et al., 1985, 1988; Brook and Shaw, 1989), clinically (Baume et al., 1974; Cons et al., 1986; Brook and Shaw, 1989), or using both of these modes (Grainger, 1967; Brook and Shaw, 1989; Ghafari et al., 1989; Uğur et al., 1998; Daniels and Richmond, 2000; Ovsenik et al., 2004, 2007).

Most of the methods were developed for malocclusion assessment in the permanent dentition (Cons et al., 1986; Brook and Shaw, 1989), and only the Occlusal Index (Summers, 1971) was designed for different stages of dental development. In Slovenia, the Eismann method, modified for the mixed and primary dentitions, based on recordings

---

**Figure 2** Rotation of incisors and canines measured intraorally (a) and on study casts (b).

**Table 2** The difference and agreement between the morphological scores evaluated intraorally (a) and on the study casts (b) assessed by Wilcoxon’s signed-ranks test and kappa ($\kappa$) statistics.

<table>
<thead>
<tr>
<th>Malocclusion trait</th>
<th>a &gt; b</th>
<th>b &gt; a</th>
<th>a = b</th>
<th>Wilcoxon’s $Z$</th>
<th>$P$</th>
<th>$\kappa$</th>
<th>SE</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anterior crowding</td>
<td>8</td>
<td>5</td>
<td>79</td>
<td>-1.165</td>
<td>0.244</td>
<td>0.812</td>
<td>0.108</td>
<td>***</td>
</tr>
<tr>
<td>Anterior spacing</td>
<td>2</td>
<td>2</td>
<td>88</td>
<td>-0.378</td>
<td>0.705</td>
<td>0.183</td>
<td>0.310</td>
<td>NS</td>
</tr>
<tr>
<td>Posterior crowding</td>
<td>23</td>
<td>17</td>
<td>52</td>
<td>-0.866</td>
<td>0.386</td>
<td>0.836</td>
<td>0.259</td>
<td>***</td>
</tr>
<tr>
<td>Posterior spacing</td>
<td>12</td>
<td>4</td>
<td>76</td>
<td>-1.695</td>
<td>0.090</td>
<td>0.103</td>
<td>0.067</td>
<td>NS</td>
</tr>
<tr>
<td>Vestibular eruption of canine</td>
<td>7</td>
<td>5</td>
<td>80</td>
<td>-0.477</td>
<td>0.633</td>
<td>0.544</td>
<td>0.065</td>
<td>***</td>
</tr>
<tr>
<td>Rotation of incisors and canines</td>
<td>22</td>
<td>24</td>
<td>46</td>
<td>-0.492</td>
<td>0.623</td>
<td>0.652</td>
<td>0.062</td>
<td>***</td>
</tr>
<tr>
<td>Axial tooth inclination</td>
<td>15</td>
<td>24</td>
<td>53</td>
<td>-2.193</td>
<td>0.028*</td>
<td>0.790</td>
<td>0.057</td>
<td>***</td>
</tr>
<tr>
<td>Overbite</td>
<td>10</td>
<td>12</td>
<td>70</td>
<td>-0.844</td>
<td>0.399</td>
<td>0.486</td>
<td>0.071</td>
<td>*</td>
</tr>
<tr>
<td>Open bite</td>
<td>3</td>
<td>6</td>
<td>83</td>
<td>-0.656</td>
<td>0.512</td>
<td>0.903</td>
<td>0.054</td>
<td>*</td>
</tr>
<tr>
<td>Anterior crossbite</td>
<td>0</td>
<td>1</td>
<td>92</td>
<td>-1.000</td>
<td>0.317</td>
<td>0.984</td>
<td>0.039</td>
<td>*</td>
</tr>
<tr>
<td>Overjet</td>
<td>9</td>
<td>8</td>
<td>75</td>
<td>-0.833</td>
<td>0.405</td>
<td>0.704</td>
<td>0.044</td>
<td>*</td>
</tr>
<tr>
<td>Buccal segment relationships</td>
<td>21</td>
<td>10</td>
<td>61</td>
<td>-1.915</td>
<td>0.055</td>
<td>0.762</td>
<td>0.050</td>
<td>*</td>
</tr>
<tr>
<td>Midline deviation</td>
<td>17</td>
<td>17</td>
<td>58</td>
<td>-0.390</td>
<td>0.696</td>
<td>0.415</td>
<td>0.059</td>
<td>*</td>
</tr>
<tr>
<td>Transverse buccal occlusion</td>
<td>17</td>
<td>16</td>
<td>59</td>
<td>-0.503</td>
<td>0.615</td>
<td>0.933</td>
<td>0.056</td>
<td>*</td>
</tr>
<tr>
<td>Classification into grades of severity</td>
<td>35</td>
<td>43</td>
<td>14</td>
<td>-0.057</td>
<td>0.954</td>
<td>0.845</td>
<td>0.066</td>
<td>*</td>
</tr>
</tbody>
</table>

* $P < 0.05$; *** $P < 0.001$; SE = standard error; NS = not significant.
and measurements on study casts has proved to be a valid diagnostic tool for malocclusion assessment in the early dental developmental stages (Farčnik et al., 1985, 1988) and reliable when used intraorally (Ovsenik et al., 2004). The method has also been demonstrated to be reliable for one examiner and among examiners, and is therefore a proposed method of choice to be used in epidemiological studies, in screening, and in clinical orthodontic assessment (Ovsenik et al., 2007). In the present study, perfect agreement was found for four occlusal traits, excellent for four traits, and moderate for four traits. The results were almost the same as achieved in the mixed dentition (Ovsenik et al., 2004).

Agreement between the two methods was better in both of the studies for all traits compared with the results of Keeling et al. (1996). The reason for this could be due to the registrations performed in a practise setting on a dental chair, with good lighting and no time limitation.

Bias between the two measurement methods was found only for axial inclination of the teeth (Table 1). Axial inclination of the teeth was measured using a protractor, which was more difficult to use intraorally than on casts, thus accounting for the bias between the two measurements. The results of this study showed that the total malocclusion score composed of all the morphological sign scores, whether recorded intraorally or on study casts, showed no systematic bias between the two methods (Table 1). One occlusal trait tended to be scored worse intraorally, but in most cases, the measurements were scored equally and thus malocclusion assessment between the two methods did not differ significantly (Ovsenik et al., 2004).

Malocclusion indices were designed to interpret malocclusion severity objectively in terms of treatment priority. Eismann (1980) suggested classification into four grades of severity into which the present sample was classified. Table 2 shows that an almost equal percentage of individuals were classified into severity grades according to both methods, with the corresponding \( \kappa \) as high as 0.84, indicating almost total agreement.

As there is no universally accepted method that defines all characteristics of a malocclusion, this is a multifactorial problem (Tang and Wei, 1993; Uğur et al., 1998; Ovsenik et al., 2004). Application of the proposed method for malocclusion assessment is more favourable for children and requires less non-clinical time when compared with assessments based on study cast measurements.

Malocclusion assessment, recorded and measured intraorally to determine malocclusion severity score in 14-year-old children, was found to be reliable for intra- and inter-examiner agreement (Ovsenik et al., 2007). As all the traits are easy to record, it may also be possible, followed by suitable training and calibration for less highly trained personnel, to apply the index (Brook and Shaw, 1989; Keeling et al., 1996; Burden et al., 2001; Ovsenik et al., 2007). It is therefore proposed as the method of choice to be used not only in epidemiological studies and screening but also in clinical orthodontic assessment.

Thus, the modified method for malocclusion assessment in the permanent dentition can be used as an epidemiological tool for screening in the identification of those children who can benefit most from orthodontic treatment. The cost–benefit of the method should be evaluated further in a longitudinal study.

### Conclusions

The results obtained from studying the reliability of intraoral measurements that compute the malocclusion index in the permanent dentition lead to the following conclusions:

1. The malocclusion severity grade, defined by a total malocclusion score composed of all the morphological sign scores, showed almost perfect agreement and no bias between the intraoral and study cast measurements.

2. Malocclusion assessment in a clinical orthodontic setting based on intraoral measurements is as reliable as that carried out on study casts. It is therefore proposed as the method of choice to be used in epidemiological studies, in screening, and in clinical orthodontic assessment.

### Table 3 Classification of malocclusion scores into grades of severity.

<table>
<thead>
<tr>
<th>Intraorally</th>
<th>Study cast</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0–15 (mild)</td>
</tr>
<tr>
<td>0–15 (mild)</td>
<td>41</td>
</tr>
<tr>
<td>16–45 (moderate)</td>
<td>5</td>
</tr>
<tr>
<td>46–65 (severe)</td>
<td>2</td>
</tr>
<tr>
<td>66 and above (very severe)</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>46</td>
</tr>
</tbody>
</table>

Bold numbers represent equally determined grades of severity classified according to the intraoral and study casts measurements.
Address for correspondence  
Maja Ovsenik  
Department of Orthodontics  
Medical Faculty  
University of Ljubljana  
Hrvatski trg 6  
1000 Ljubljana  
Slovenia  
E-mail: maja.ovsenik@dom.si

References

Angle E H 1907 The malocclusion of teeth. S S White Dental Manufacturing, Philadelphia


Cons N C, Jenny J, Kohout F J 1986 DAI: Dental Aesthetic Index. College of Dentistry, University of Iowa, Iowa City


Eismann D 1977 The morphology of the dentition as one criterion in the assessment of the need for orthodontic treatment. Transactions of the European Orthodontic Society, pp. 125–129


Lundström A 1977 Need for treatment in cases of malocclusion. Transactions of the European Orthodontic Society, pp. 111–123


