Intra- and inter-examiner reliability of intraoral malocclusion

assessment

Maja Ovsenik*, F. Farčnik* and I. Verdenik**

Departments of *Orthodontics and **Obstetrics and Gynaecology Research Unit, Medical Faculty, University of Ljubljana, Slovenia

SUMMARY Malocclusion assessment methods are based on measurements of study casts, which requires that impressions be taken. In addition to being costly and time consuming, this process can be unpleasant for children. Therefore, the aim of this study was to evaluate intra- and inter-examiner reliability of intraoral score measurements to determine malocclusion severity in the permanent dentition. The research was a part of a longitudinal study from which a cohort of 92 children (39 boys, 53 girls), with a mean age of 14.8 years (standard deviation = 0.18), were randomly selected and classified into severity grades based on total malocclusion score. Subsequently, 12 children were randomly selected for a reliability study to assess intra-examiner reliability of malocclusion trait measurements. Nine subjects gave informed consent to participate in the study. Quantitative registrations of space and occlusal anomalies were performed intraorally by five examiners, on two occasions with a 1-month interval between the two measurements. Intra- and inter-examiner reliability was determined using intraclass correlation coefficients (ICCs).

Overall classification into severity grades, based on total malocclusion score, showed almost perfect intra-examiner reliability for all examiners (ICC = 0.97-0.99); inter-examiner ICC was almost perfect (0.97). Near perfect intra-examiner reliability was determined for eight occlusal trait measurements (ICC = 0.89-1.0); substantial reliability for midline deviation (ICC = 0.68), overbite (ICC = 0.78), but large variability for space condition assessment (ICC = 0.42-0.52). Inter-examiner reliability was almost perfect for the eight traits (ICC = 0.81-1.0); substantial reliability for midline deviation (ICC = 0.65), and axial tooth inclination (ICC = 0.75), but large variability for space condition assessment (ICC = 0.13-0.26).

Intra- and inter-examiner malocclusion assessment, recorded and measured intraorally to determine malocclusion severity scores in 14-year old children, is reliable. 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.

Introduction

Different countries have adopted varying methods of funding orthodontic care for children. In countries that embrace the principle of publicly funded orthodontic care for all children with high objective need, reliable population data are required to evaluate the effectiveness of orthodontic services (Burden *et al.*, 2001).

Malocclusion assessment methods differ not only in the choice of the recorded morphological or functional criteria used but also in the mode of their evaluation. It can be undertaken on study casts (Summers, 1971; Eismann, 1974, 1980; Farčnik *et al.*, 1985, 1988; Brook and Shaw, 1989; Richmond *et al.*, 1992), clinically (Baume *et al.*, 1974; Cons *et al.*, 1986; Brook and Shaw, 1989; Richmond *et al.*, 1992; Ovsenik *et al.*, 2004), or both (Grainger, 1967; Brook and Shaw, 1989; Richmond *et al.*, 1992; Uğur *et al.*, 1998; Daniels and Richmond, 2000).

Indices of orthodontic treatment need to have the potential both for acquiring descriptive data on the distribution of treatment need in populations (epidemiological use) and for establishing priorities for treatment (administrative use; Helm, 1970; Espeland *et al.*, 1992; Tang and Wei, 1993; Burden *et al.*, 2001; Liepa *et al.*, 2003). Generalized use of an index by individual members of the speciality would depend, in part, on the reliability of the various descriptors of malocclusion. Health care providers are often unaware of the imperfect reliability of the methods and data of clinical practice. In clinical orthodontics, the assessment of malocclusion remains problematic (Keeling *et al.*, 1996; Ovsenik *et al.*, 2004). Malocclusion indices (Occlusal Index: Summers, 1971; Index of Orthodontic Treatment Need: Brook and Shaw, 1989; Index of Complexity, Outcome, and Need: Daniels and Richmond, 2000) have been shown to have acceptable reliability (Brook and Shaw, 1989; Buchanan *et al.*, 1993; Burden *et al.*, 2001; Fox *et al.*, 2002; Johansson and Follin, 2005). However, the reliability of the individual measurements that are scored has only been reported by Keeling *et al.* (1996) and Ovsenik *et al.* (2004).

The indices have proved to be complicated and time consuming in daily use (Solow, 1995; Ovsenik *et al.*, 2004). In some countries, a system has been developed in which various types of malocclusion are graded into four categories, depending on the type and severity of deviation (Brook and Shaw, 1989; Espeland *et al.*, 1992; Solow, 1995) or especially modified for use in oral health surveys (Burden *et al.*, 2001).

In Slovenia, the Eismann index is used as a method for epidemiological studies, in training specialists and undergraduate students, in research, as well as by public health care organizations to help determine the level of a third-party payment. As orthodontic treatment in Slovenia is publicly funded for all children and adolescents up to the age of 18 years, regardless of malocclusion severity, the problem of treatment priority arises.

In order to assess malocclusion in the early dental development period, the method of Eismann (1974) was modified for the primary and mixed dentition (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 two methods to obtain study casts are often quite unpleasant, especially for children, while the procedure itself can be costly, complicated, and time consuming in daily use (Solow, 1995; Ovsenik *et al.*, 2004).

It has been established that malocclusion assessment in the early mixed dentition period, based on intraoral measurements, is as reliable as that carried out on study casts and is the method of choice in epidemiological studies, in screening, and in clinical orthodontic assessment (Ovsenik *et al.*, 2004). Application of the proposed method in clinical orthodontics is preferred as it requires less clinical time when compared with assessment based on study cast measurements. The reliability of intraoral malocclusion assessment in the period of the permanent dentition according to the Eismann (1974) method 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, to determine malocclusion severity in the permanent dentition period, and to identify intraorally determined measurements of malocclusion that have poor intra- and inter-examiner reliability.

Subjects and methods

The research, approved by the Ethics Committee, University of Ljubljana, Medical Faculty, Division for Dentistry (Ovsenik, 2003), was a part of a longitudinal study (Farčnik *et al.*, 1986) in Slovenia on a sample of 530 three-year-old children. A cohort of 92 children (39 boys, 53 girls), with a mean age of 14.8 years (standard deviation = 0.18), were randomly selected and classified into severity grades based on total malocclusion score. Subsequently, 12 children were selected at random for a reliability study to assess intra- and inter-examiner reliability of malocclusion trait measurements by five examiners trained in the use of the index. Nine subjects gave informed consent to participate in the study. The measurements were performed on all five occasions by all five examiners.

Examiners

Examiner A: An orthodontic specialist with 1 year's clinical experience who had completed training in malocclusion

assessment based on intraoral measurements of morphological malocclusion traits.

Examiner B: An orthodontic specialist with 1 year's clinical experience who had completed training in malocclusion assessment based on morphological malocclusion traits measurements intraorally and on study casts.

Examiner C: An orthodontic specialist with 5 years' clinical experience who had completed training in malocclusion assessment based on intraoral morphological trait measurements.

Examiner D: An orthodontic specialist with 30 years' clinical experience who had completed training in malocclusion assessment based on morphological trait measurements intraorally and on study casts.

Examiner E: A postgraduate orthodontic student with 1 year's clinical experience who had completed training in morphological trait measurements on study casts.

Each examiner used a head-held light, rulers, gloves, and completed a scoring form for each child (Figure 1), which also included demographic and malocclusion variables.

Training of the examiners

Prior to the start of recordings and measurements, the five examiners met over a period of 2 months to determine the data to be collected and to practise using the index both on study casts and intraorally. The examinations were performed independently in dental chairs in five dental practices, by all five examiners.

Recordings and measurements

For each set of measurements, the registrations were carried out according to the Eismann (1974) scoring form (Figure 1). For measurements of linear dimensions, a metric ruler (Zürcher modell, 042-751; Dentaurum, Ispringen, Germany), accurate to 1/10 mm, was used. Angles were measured with a protractor (Eismann, 1974) to determine the rotation of the incisors and the axial inclination of the teeth.

Intra-arch assessment involved measurement of anterior and posterior crowding and spacing, rotation of the incisors, and axial inclination of the teeth. For inter-arch measurements, overbite, anterior and posterior open bite, overjet, reverse overjet, anterior and posterior crossbite, and buccal segment relationships were recorded.

All morphological signs, measured intraorally and expressed in millimetres and degrees, were weighted and scored against the Eismann (1974) evaluation table for each subject (Figure 2). The weighted sum of the recorded occlusal traits thus represented the total malocclusion index severity score. The overall malocclusion scores were categorized according to Eismann (1974) in terms of mild (1–15), moderate (16–40), severe (41–65), and very severe (over 66) malocclusion.

					birth data.									-					
name:								birth date:				no.:	no.:			F			
examiner: examiner			exan	nination:				date of exam.:				age:							
ANTER	IOR A	ND PO	OSTER	IOR CF	ROWE	ING													
S=					S= S							S=							
		I-I=			D=		I-I=			D=			I-I=		D=				
١.		S=			S=		S=			S=			S=		S=				
		I-I=			I-I=		I-I=			I-I=			I-I=		I-I=		ant. crowd.		
	Т			Τ							Ι								
		S=			S=		S=			S=			S=		S=				
III.		I-I=			I-I=		I-I=			I-I=			I-I=		I-I=				
	S		S=	=				S=	S=					S=		post. crowd.			
		I-I=	-		D=			I-I=		[)=		I-	I=	D)=			
ANTER	IOR A	ND PO	OSTER	IOR SP	ACIN	G													
	SI	=			S	L=					SL=				Г	ant.si	oacing		
п.	F											1	Т	1					
	+	+	+			+						+	+	+	1	post, s	pacing		
				SL=					SL=				-						
VESTIB	ULAR	ERUP	TION	OF CA	NINE,	ROTA	TION	AND	ROSS	BITE						1			
V.								rotation											
VI.	13		23	vestibu	ılar er	uptio	n of o	canine	3	2	2	1	1	2	3				
43 33							3	- 2	2	1	1	2	3	-	crossbite				
OPEN B	BITE A	ND A	XIALT	OOTH	INCL	NATIO	NC												
VII.														Ι	7	oper	n bite		
IX.														<	<				
															axial tooth		on		
Χ.		-												<					
OVERB	ITE.O	VERJE			INF D	EVIA	FION												
									dev	viatior	betv	veen	the mi	dline		7			
VII.		overbite			overjet				of the dental arch and the raphe						deviation between the midlines				
				palatina mediana of the upper a							ind lower	Jun							
XI.		mm mm						mm					mm						
XIV.																			
BUCCAL SEGMENT RELATIONSHIP																			
VIII	7	6	5	4	3					3	4	5	6	7	bucca	l segme	nt relatio	nship	
XIII.	7	6	5	4	3					3	4	5	6	7					
TDANC		- DOC	TEDIO	DOCC		N.I													
xv.	7	6	5	4	3		-	-	-	3	4	5	6	7	transve	erse pos	lerior occ	lusion	
	/	0	S	4	3		1			3	4	5	6	/					

Figure 1 The Eismann scoring form used by the examiners (reproduced with kind permission).

Statistical analysis

The intraclass correlation coefficient (ICC) was used to evaluate the intra- and inter-examiner reliability of 14 occlusal trait measurements, recorded and measured intraorally. ICC 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 ICC interpretation scale was used: poor to fair (below 0.4), moderate (0.41–0.60), excellent (0.61–0.80), and almost perfect (0.81–1).

For the analysis, the Statistical Package for Social Sciences Windows, version 12 (SPSS Inc., Chicago, Illinois, USA) was used.

Results

Classification of malocclusion scores into grades of severity

The classification of malocclusion scores into four severity grades is shown in Figure 3. Most of the children were classified as grade 1 or 2 (score 1-40), with only one child classified as grade 4 (score = 99).

Intra-examiner repeatability (ICC) for the 14 morphological signs and five examiners

Almost perfect intra-examiner reliability was determined for the eight occlusal trait measurements (ICC = 0.89-1.0), excellent reliability for midline deviation (ICC = 0.68) and

I. Space analysis	7. Axial tooth inclination	
1. Anterior crowding	Incisors	Reverse overjet
0-2 mm 0	10-25° 2	0-1 mm 4
2-3 mm 1	25-40° 4	1-2 mm 8
3-4 mm 1	40-55° 6	2-3 mm 12
4-5 mm 3	55-70° 8	3-4 mm 14
5-6 mm 5	posterior teeth	4-5 mm 15
6-7 mm 7	10-25° 1	> 5 mm 16
7-8 mm 9	25-40° 2	o mini To
> 8 mm 11	40-55° 3	12 Anterior crossbite incl
Intense growding in the region	55 70° 4	conines
of one teeth	55-70 4	first noin 8
	II. Manubala sinal tusita af	nist pair 8
3-4 mm 2	II. Morphological traits of	next pair 5
4-5 mm 4	malocclusion in the vertical	
	plane	13. Anteroposterior occlusion of
2. Anterior spacing	8. Overbite	posterior teeth: Occlusion of
0-2 mm 0	0-4 mm 0	single cusps in the case of 1-2
2-4 mm 2	4-5 mm 1	pairs of opposite teeth
4-6 mm 4	5-6 mm 1	1
6-8 mm 6	6-7 mm 3	
8-10 mm 8	7-8 mm 8	Occlusion of single cusps in the
> 10 mm 10	8-9 mm 11	case of 3 and more pairs of
	9-10 mm 13	opposite teeth
3. Posterior crowding	10-11 mm 16	3
0-1 mm 0	11-12 mm 18	5
1-2 mm 1	12-13 mm 20	IV Mornhological traits of
2 3 mm 2	12-15 mm 20	malocelusion in the transverse
2-5 mm 2	0 Antonion on on hits	maiocerusion in the transverse
5-4 mm 2	9. Anterior open one	plane
4-5 mm 5	incl.canines	14. Midline's deviation of the
5-6 mm 7	0-1 mm 3	upper jaw
> 6 mm 9	1-2 mm 6	0-1 mm 0
	2-3 mm 9	1-2 mm 1
 Posterior spacing 	3-4 mm 12	2-3 mm 2
0-2 mm 0	4-5 mm 14	> 3 mm 3
2-4 mm 1	> 5 mm 15	
4-6 mm 2		Deviation between the midlines
> 6 mm 3	10. Posterior open bite	of the upper and lower jaw
	0-1 mm 1	0-1 mm 0
5. Vestibular eruption of canine	1-2 mm 3	1-2 mm 1
0-2 mm 0	2-3 mm 5	2-3 mm 2
2-4 mm 4	3-4 mm 6	$\geq 3 \text{ mm}$ $= 3$
4-6 mm 7	4-5 mm 7	- 5 11111 - 5
>6 mm 10		15 Transverse posterior
	III Morphological traits of	15. Hansverse posterior
6 Detetion of incisons	malacelusion in the entrus	account of single cusps in the
o. Kotation of incisors	malocclusion in the antero-	case of 1-2 premotars
0-15° 0	posterior plane	
15-30° 2	11. Overjet	occlusion of single cusps in the
30-45° 3	0-4 mm 0	case of premolars and molars
45-60° 5	4-5 mm 2	2
60-75° 6	5-6 mm 5	crossbite per pair of opposite
75-90° 8	6-7 mm 8	teeth
90-105° 9	7-8 mm 12	1
105-120° 7	8-9 mm 16	buccal or oral nonocclusion per
120-135° 6	9-10 mm 20	pair of opposite teeth
135-150° 4	> 10 mm 20	A
150-165° 3	- 10 mm 2 4	4
165 1900 1		
100-100 1		

Figure 2 The Eismann scoring table (reproduced with kind permission).

overbite (ICC = 0.78), but in the assessment of dental arch crowding (ICC = 0.40-0.80) and spacing (ICC = 0.25-0.82) there was large variability among all five examiners (Table 1).

The highest intra-examiner agreement of space condition assessment intraorally was found for examiner D (ICC = 070). Despite the fact that there was intra-examiner variability in scoring space conditions, almost perfect intra-examiner reliability for total malocclusion assessment and for classification into grades of severity was determined (ICC = 0.97-0.99).

Inter-examiner repeatability (ICC) for the 14 morphological signs among the five examiners

The inter-examiner reliability (repeatability) among the morphological scores, measured intraorally among the five examiners, was assessed by the ICC test. The results are presented in Table 2.

Perfect agreement of intraoral malocclusion trait measurements among the examiners was found for eight



Figure 3 Classification of the eight malocclusion trait scores into grades of severity for the cohort of 92 children.

occlusal traits (ICC = 0.8-1.00) and excellent agreement for measuring axial tooth inclination (ICC = 0.60) and midline deviation (ICC = 0.65). Large variability among the five examiners, however, was found for intraoral assessment of dental arch spacing and crowding (ICC = 0.13-0.26). Despite poor inter-examiner reliability for space condition assessment, the agreement of examiners in assessment of total malocclusion severity score (ICC = 0.97) and classification into severity grades (ICC = 0.94) was almost perfect.

Discussion

There is considerable international interest in guidelines for the screening of children for orthodontic treatment (Solow, 1995). New malocclusion indices and indices of treatment need are being developed and tested in many countries (Brook and Shaw, 1989; Espeland *et al.*, 1992; Burden *et al.*, 2001; Tausche *et al.*, 2004; Johansson and Follin, 2005), and their deficiencies are well recognized.

Malocclusion assessment methods for screening and epidemiological studies were designed either for study cast measurements (Summers, 1974; Helm, 1977; Ghafari *et al.*, 1989) or for clinical use (Baume *et al.*, 1974; Cons *et al.*, 1986; Brook and Shaw, 1989; Richmond *et al.*, 1992; Uğur *et al.*, 1998; Daniels and Richmond, 2000; Burden *et al.*, 2001; Ovsenik *et al.*, 2004). The method proposed by Baume *et al.* (1974) was designed for observations and measurements to be made directly in the mouth. Although there are certain advantages and conveniences in undertaking measurements on casts, obtaining these may not be possible under many field conditions (children, costs, time) and thus, for consistency the assessments are limited to direct observations (Baume *et al.*, 1974; Ovsenik *et al.*, 2004).

In clinical orthodontics, malocclusion assessment remains problematic. Index scores have been shown to have acceptable reliability (Brook and Shaw, 1989; Richmond *et al.*, 1992). However, the reliability of the individual measurement that compute these scores has only been reported by Keeling *et al.* (1996) and Ovsenik *et al.* (2004). As studies assessing the reliability of scoring components of malocclusion have most frequently been performed on diagnostic records, reports on the reliability of assessing an individual malocclusion trait during clinical examination of subjects are scarce.

In a study by Ovsenik *et al.* (2004) it was determined that the total malocclusion severity score, composed of 10 morphological signs as in the method of Eisamann modified by Farčnik *et al.* (1985), was no different when recorded

 Table 1
 Intra-examiner repeatability (intraclass correlation coefficient) for 14 morphological signs and five examiners.

Malocclusion trait	Examiner							
	A	В	С	D	E 0.47	0.51		
Anterior crowding	0.40	*	0.48	0.70				
Anterior spacing	0.38	0.82	*	*	0.37	0.52		
Posterior crowding	0.80	0.41	0.26	0.41	0.53	0.48		
Posterior spacing	0.25	0.82	0.40	*	0.25	0.43		
Vestibular eruption of canine	0.88	1.00	1.00	1.00	0.84	0.95		
Axial tooth inclination	0.93	0.96	0.92	0.92	0.86	0.92		
Rotation of incisors	1.00	0.93	0.98	0.99	0.99	0.98		
Overbite	0.84	0.80	0.95	0.99	0.33	0.78		
Anterior and posterior open bite	1.00	1.00	1.00	1.00	0.95	0.99		
Anterior crossbite	1.00	1.00	1.00	1.00	1.00	1.00		
Overjet	0.76	0.87	0.92	1.00	0.90	0.89		
Buccal segment relationships	0.91	0.78	1.00	0.97	0.99	0.93		
Midline deviation	0.43	0.73	0.80	0.80	0.67	0.69		
Transverse posterior occlusion	0.74	0.94	0.97	0.98	0.99	0.92		
Malocclusion severity score	0.98	0.97	0.99	0.99	1.00	0.99		
Classification into severity grades	0.96	0.91	0.97	0.97	1.00	0.96		

*As the teeth in the dental arch were rotated and inclined in some subjects and different cut-off points were used by examiners C and D, no malocclusion trait was recorded.

Table 2Inter-examiner repeatability [intraclass correlationcoefficient (ICC)]for 14 morphological signs among the fiveexaminers.

Morphological trait of malocclusion	ICC
Anterior crowding	0.22
Anterior spacing	0.13
Posterior crowding	0.26
Posterior spacing	0.18
Vestibular eruption of canine	0.95
Rotation of incisors and canines	0.87
Axial tooth inclination	0.76
Overbite	0.98
Anterior and posterior open bite	0.99
Anterior crossbite	1.00
Overjet	0.81
Buccal segment relationships	0.85
Midline deviation	0.65
Transverse buccal occlusion	0.91
Total malocclusion severity score	0.98
Classification into grades of severity	0.94

intraorally or on study casts. Some occlusal traits were scored intraorally with high values, and some lower, but in most cases the measurements were scored equally for both, and thus malocclusion assessment between the two methods did not differ significantly. An equal percentage of children were classified into each grade of malocclusion severity, regardless of the recorded method. Thus, the modified method for malocclusion assessment in the early dental development period in the mixed dentition can be used as an epidemiological tool for screening and in the identification of children who can most benefit from orthodontic treatment.

Malocclusion indices were designed to interpret malocclusion severity objectively in terms of treatment priority. Eismann (1977) suggested four grades of severity, into which the present sample was classified. Most of the children were classified as grade 1 or 2, which is in agreement with other studies on orthodontic treatment need obtained using other malocclusion assessment methods (Brook and Shaw, 1989; Holmes, 1992; Uğur *et al.*, 1998; Burden *et al.*, 2001; Tausche *et al.*, 2004). Any index should be valid and reliable for one examiner and among examiners (Richmond *et al.*, 1992). The results show that an almost equal percentage of children were classified into the four severity grades by a single examiner and among the examiners, with the corresponding ICC as high as 0.96 and 0.94, indicating almost perfect intra-and inter-examiner agreement.

Almost perfect intra-examiner agreement was determined for all examiners, with the highest results for Examiner D with the longest clinical experience in orthodontics and also in the use of the method; on the other hand, excellent agreement was determined for examiner E, who had the least clinical orthodontic experience and in the use of the method (Table 1). Intra-examiner agreement was almost perfect for eight occlusal trait measurements, and substantial for two occlusal traits. The reliability results were higher for all the trait measurements due to the fact that registrations were performed in a practice setting on a dental chair, with good lighting and no time limitation, compared with the results achieved by Keeling *et al.* (1996) and also because all the examiners had had training in the use of the method intraorally and on the study casts, with written instructions in the use of the method.

Variability of the five examiners was, however, found for space condition assessment in the dental arches (Table 1). The poor to fair reliability for the measurement of crowding may be related to dental arch crowding and spacing. In the present study, crowding was recorded according to the method of Björk et al. (1964). In crowding of either incisors or in the buccal segments, the teeth are positioned lingually or buccally and are rotated or inclined, and thus different cut-off points can be determined, which results in poor reliability both for a single examiner and among the examiners. In the methods used by other researchers (Grainger, 1967; Brook and Shaw, 1989; Ghafari et al., 1989; Uğur et al., 1998; Tang and Wei, 1996), space condition assessment records the potential tooth displacement and is, as such, unreliable. Crowding has been determined to be the most common anomaly (Helm, 1970, 1977; Eismann, 1974; Farčnik et al., 1985, 1988; Brook and Shaw, 1989; Ghafari et al., 1989; Thilander et al., 2001; Ovsenik et al., 2004) and thus well defined cut-off points in the assessment of space conditions in the dental arch would certainly improve its reliability for a single examiner and among examiners. Poor reliability among orthodontists in scoring crowding revealed that clinical orthodontic definitions for crowding assessment are imprecise. Therefore, methods to improve diagnostic terms, altering measurement scales, more rigid definitions, and a more systematic approach in training orthodontic specialists should be considered (Keeling et al., 1996).

The four examiners, for whom almost perfect agreement in measurements of malocclusion traits was found, represent the population of experts (orthodontic specialists), who would use the proposed method of malocclusion assessment intraorally in everyday routine clinical work. The data demonstrate that the method is reliable and worthwhile to be used widely in the field of clinical orthodontics. The results for Examiner E showed excellent intra- and inter-examiner agreement, while those for the orthodontic specialists showed that increased clinical experience additionally improved the intra- and inter-examiner reliability. Therefore, the reliability among dentists with special training in orthodontics should be further evaluated. As all the traits were easy to record, with suitable training and calibration it may also be possible for less highly trained personnel to apply the index (Brook and Shaw, 1989; Keeling et al., 1996; Burden et al., 2001).

Conclusions

The results obtained from determining the intra- and inter-examiner reliability of intraoral measurements that compute the total malocclusion severity score support the following conclusions.

- Almost perfect to good intra- and inter-examiner reliability was determined for almost all occlusal trait measurements. Despite poor intra- and inter-examiner reliability for space condition assessment, the malocclusion severity grade, defined by total malocclusion severity score, showed almost perfect agreement for one examiner and among the examiners.
- Malocclusion assessment in a clinical orthodontic setting based on intraoral measurements is reliable for one examiner and among the examiners. 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.

Address for correspondence

Maja Ovsenik Department of Orthodontics, Medical Faculty University of Ljubljana Zaloska 2 1000 Ljubljana Slovenia E-mail: maja.ovsenik@dom.si

References

- Baume L J *et al* 1974 A method for the measurement of occlusal characteristics. Commission on Classification and Statistics for Oral Conditions of the FDI (COCSTOC). International Dental Journal 24: 90–97
- Björk A, Krebs A, Solow B 1964 A method for epidemiological registration of malocclusion. Acta Odontologica Scandinavica 22: 27–41
- Brook P H, Shaw W C 1989 The development of an index of orthodontic treatment priority. European Journal of Orthodontics 11: 309–320
- Buchanan I B, Shaw W C, Richmond S, O'Brien K D, Andrews M 1993 A comparison of the reliability and validity of the PAR Index and Summers' Occlusal Index. European Journal of Orthodontics 15: 27–31
- Burden D J, Pine C M, Burniside G 2001 Modified IOTN: an orthodontic treatment need index for use in oral health surveys. Community Dentistry and Oral Epidemiology 29: 220–225
- Cons N C, Jenny J, Kohout F J 1986 DAI: Dental Aesthetic Index. College of Dentistry University of Iowa, Iowa City
- Daniels C, Richmond S 2000 The development of the Index of Complexity, Outcome and Need (ICON). Journal of Orthodontics 27: 149–162
- Eismann D 1974 A method of evaluating the efficiency of orthodontic treatment. Transactions of the European Orthodontic Society pp. 223–232
- 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
- Eismann D 1980 Reliable assessment of morphological changes resulting from orthodontic treatment. European Journal of Orthodontics 2: 19–25
- Espeland L V, Ivarsson K, Stenvik A 1992 A new Norwegian index of orthodontic treatment need related to orthodontic concern among 11-year-olds and their parents. Community Dentistry and Oral Epidemiology 20: 274–279
- Farčnik F, Korpar M, Premik M, Zorec R 1985 Numerical evaluation of malocclusion in study models of the mixed dentition. Zobozdravstveni Vestnik 40: 169–176
- Farčnik F, Korpar M, Premik M, Zorec R 1986 Morphological and functional occlusal traits significant in the assessment of malocclusion

to determine the severity score in deciduous dentition. Research Project for the Research Community of Slovenia No. C3-0560-329-86. URP, Stomatology University Dental Clinic, Ljubljana, pp. 1–17

- Farčnik F, Korpar M, Premik M, Zorec R 1988 An attempt at numerically evaluating dysgnathias in the deciduous dentition. Stomatologie DDR 38: 386–391
- Fox N A, Daniels C, Gilgrass T 2002 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). British Dental Journal 193: 225–230
- Ghafari J, Locke S A, Bentley J M 1989 Longitudinal evaluation of the Treatment Priority Index (TPI). American Journal of Orthodontics and Dentofacial Orthopedics 96: 382–389
- Grainger R M 1967 Orthodontic treatment priority index. Public Health Service Publication No. 1000, Series 2, No. 25 US Government Printing Office, Washington DC
- Helm S 1970 Prevalence of malocclusion in relation to development of the dentition. An epidemiological study of Danish school children. Acta Odontologica Scandinavica Supplementum 58: 1
- Helm S 1977 Intra-examiner reliability of epidemiologic registrations of malocclusion. Acta Odontologica Scandinavica 35: 161–165
- Holmes A 1992 The prevalence of orthodontic treatment need. British Journal of Orthodontics 19: 177–182
- Johansson A M, Follin M E 2005 Evaluation of the aesthetic component of the Index of Orthodontic Treatment Need by Swedish orthodontists. European Journal of Orthodontics 27: 160–166
- Keeling S D, Wheeler T T, Wheeler T T, King G J 1996 Imprecision in orthodontic diagnosis: reliability of clinical measures of malocclusion. Angle Orthodontist 66: 381–391
- Korpar M et al 1994 Changes in the orofacial system between the 3rd and the 9th years of age. In: Farčnik F (ed.) Preventive and interceptive orthodontics. Book of Proceedings. Slovenian Orthodontic Society, Rantovi dnevi, Ljubljana, pp. 41–47
- Landis J R, Koch G G 1977 The measurement of observer agreement for categorical data. Biometrics 33: 159–174
- Liepa A, Urtane I, Richmond S, Dunstan F 2003 Orthodontic treatment need in Latvia. European Journal of Orthodontics 25: 279–284
- Ovsenik M 2003 Reliability of the intraoral assessment of malocclusion in fourteen-year-old children. Thesis, University of Ljubljana, Slovenia
- Ovsenik M, Farčnik F, Verdenik I 2004 Comparison of intra-oral and study cast measurements in the assessment of malocclusion. European Journal of Orthodontics 26: 273–277
- Richmond S *et al* 1992 The development of the PAR Index (Peer Assessment Rating): reliability and validity. European Journal of Orthodontics 14: 125–139
- Solow B 1995 Guest editorial: orthodontic screening and third party financing. European Journal of Orthodontics 17: 79–83
- Summers C J 1971 The Occlusal Index. American Journal of Orthodontics 59: 552–567
- Tang E L K, Wei S H Y 1993 Recording and measuring malocclusion: a review of the literature. American Journal of Orthodontics and Dentofacial Orthopedics 103: 344–351
- Tausche E, Luck O, Harzer W 2004 Prevalence of malocclusions in the early mixed dentition and orthodontic treatment need. European Journal of Orthodontics 26: 237–244
- Thilander B, Pena L, Infante C, Parada S S, Mayorga C 2001 Prevalence of malocclusion and orthodontic treatment need in children and adolescents in Bogota, Colombia. An epidemiological study related to different stages of dental development. European Journal of Orthodontics 23: 157–176
- Uğur T, Ciğer S, Aksoy A, Telli A 1998 An epidemiological survey using the Treatment Priority Index. European Journal of Orthodontics 20: 189–193

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