European Journal of Orthodontics 34 (2012) 768–772
doi:10.1093/ejo/cjr096
Advance Access publication 9 August 2011

The 5-year-olds' Index: determining the optimal format for rating dental arch relationships in unilateral cleft lip and palate

O. Chawla*, S. A. Deacon**, N. E. Atack*, A. J. Ireland* and J. R. Sandy*
*Department of Child Dental Health, Bristol Dental Hospital and **South West Cleft Unit, Frenchay Hospital, Bristol, UK

Correspondence to: Jonathan Sandy, Department of Child Dental Health, Bristol Dental Hospital, Lower Maudlin Street, Bristol BS1 2LY, UK. E-mail: Jonathan.Sandy@bristol.ac.uk

SUMMARY The aim of this investigation was to determine the optimal format for presenting the 5-year-olds' Index reference models for the assessment of dental arch relationships in unilateral cleft lip and palate (UCLP). The 5-year-olds' Index reference models were presented in four different formats: plaster models, coloured acrylic models, and two digital formats, which included three-dimensional digital models (3D), and black and white photographs. These formats were used to rate 45 plaster models of patients born with UCLP by a total of seven examiners comprising experienced and inexperienced examiners. Scoring was undertaken 1 week apart for each format with the patient models reassigned on each occasion to reduce the effect of memory bias. For intra-examiner agreement, the patient models were scored 3 weeks later under similar conditions by the same examiners. The reliability of using the different formats was determined using the plaster reference models as the 'gold' standard for comparison. Analysis of the results using weighted kappa (*k*) statistics showed the magnitude of agreement for all the formats of the 5-year-olds' Index were good to very good between examiners. The more experienced examiners were consistently more reliable in their scoring using the different formats. This study clearly demonstrated that 3D digital models of the 5-year-olds' Index could prove to be a promising alternative to physical models of the 5-year-olds' Index.

Introduction

There are a number of assessment tools that provide a basis on which surgeons can judge their surgical results and relate any changes in technique or timing to outcome. One such method is evaluating and comparing dental arch relationships. Early methods of assessing dental arch relationships in unilateral cleft lip and palate (UCLP), described the prevalence and type of crossbite in the deciduous dentition (Pruzansky and Aduss, 1964; Huddart and Bodenham, 1972). However, scoring with such systems may fail to take into account the severity of the malocclusion as a whole and has the potential for under estimating the discrepancy between the arches. Indices such as the Goslon yardstick categorize dental arch relationships in terms of anterior–posterior, vertical, and transverse relationships in the late mixed and/or early permanent dentition (Mars *et al.*, 1987). More recently, a similar scoring system has been developed to predict surgical outcome even earlier (Atack *et al.*, 1997). Here, the format of the original Goslon yardstick was retained and led to the development of the 5-year-olds' Index for assessing outcome in children with UCLP at the age of 5 years. The 5-year-olds' Index uses reference models to grade patient study models into one of five categories (1–5), with 1 representing an excellent outcome in terms of features such as positive overjet, the absence of crossbites, and good upper arch form. By contrast, 5 represents a very poor surgical outcome with all the teeth including the anterior and posterior teeth in crossbite, together with poor maxillary arch form. A number of studies have successfully reported surgical outcome using the 5-year-olds' Index. This index underpinned the UK Clinical Standards Advisory Group (CSAG) study on cleft care in this age group (Murray, 1998) and has also shown good inter- and intra-observer agreement in local studies (Johnson *et al.*, 2000; Clark *et al.*, 2007) and inter-centre studies (Atack *et al.*, 1998; Flinn *et al.*, 2006).

In current format, the 5-year-olds' Index is presented as hand-held plaster models which are transported nationally and internationally for categorizing models of 5-year-old children born with UCLP. This format carries an obvious potential risk of model damage. A more convenient approach is to substitute the physical models for digital images which can then easily be saved on a data storage device or sent via the web to the point where scoring is to be carried out. Nollet *et al.* (2004) investigated the reliability of rating dental arch relationships using photographs of study models for both the patient models and the Goslon reference set. They found that there were no significant differences between the rating of dental arch relationships using the dental models and photographs of dental models. Currently, there are no reports in the literature investigating the reliability of presenting the 5-year-olds' Index in a 3D digital format.

The aim of this research was to investigate the reliability of using four different formats of the 5-year-olds' Index: 1. Plaster models, 2. coloured acrylic models, 3. black and white photographs, and 4. 3D digital models in a single study and to determine the optimal format for presenting the reference models in the assessment of dental arch relationships in UCLP.

Materials and methods

Study models were retrieved from the archive at Bristol Dental Hospital, Bristol, UK. These were study models that were collected from 57 centres across the UK as part of the government commissioned CSAG study in 1998. All study models were obtained from non-syndromic patients around 5 years of age with:

- A complete unilateral bony cleft with soft tissue bands of less than 5 mm
- No history of previous orthodontic intervention
- Primary lip and palate closure had been achieved

The 5-year-olds' index formats

Currently, the 5-year-olds' Index reference models are available in two formats: plaster and acrylic models (Figure 1).

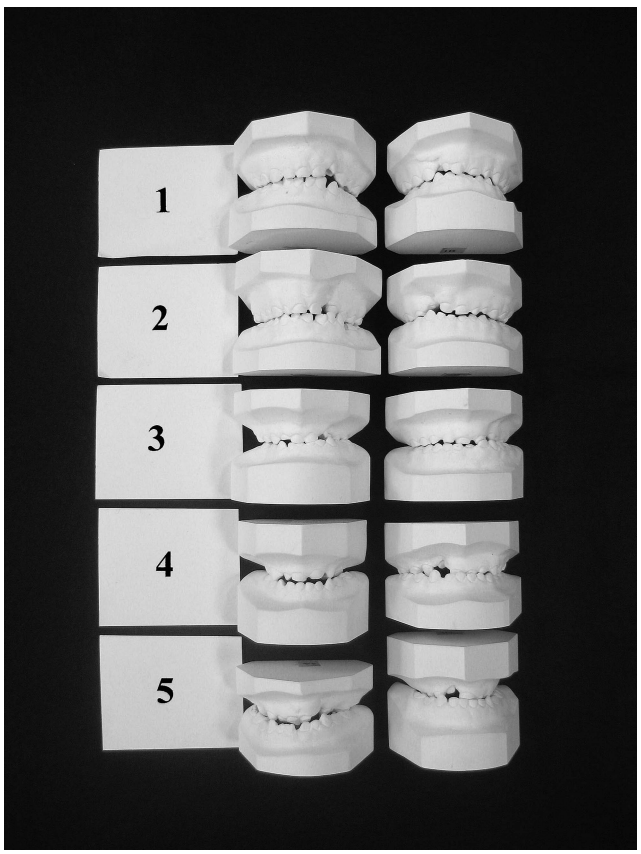

Figure 1 Presentation of the 5-year-olds' Index as plaster models.

For the photographic format, digital photographs of the plaster reference models were obtained using a Fujifilm S3 Pro digital camera (Fuji photo film Corporation, Ltd) and an AK Micro Nikon 105mm/1:2.8D lens (Nikon Corporation) at the medical illustration department, Royal United Hospitals Trust, Bath, UK. The camera was mounted on a De Vere 504 C/F camera stand with the lens object distance set at 55 cm. All study models were placed on a black velvet cloth and positioned on a light box. Once in the correct position, an illuminating lamp shade was placed over the set of models for even light distribution and to reduce the effects of shadowing. For each set of reference models, a frontal view, right and left buccal views, and upper and lower occlusal views were taken. For the buccal views, a bag of sand was positioned under the black velvet cloth to help stabilize the models. The five views for each of the reference models were then placed on a single power point slide (Microsoft, corporation) indicating the reference model number in the top left hand corner. These were then printed on photographic paper (Premium, Hewlett Packard), laminated, and mounted on a white board with a stand, displaying the 5-year-olds' Index from 1 to 5.

3D digital images of the reference models were obtained using the R640 3Shape Desktop scanner (Copenhagen, Denmark). Reference models were scanned by Joseph Burke laboratory (Limerick, Ireland) and made available for downloading and viewing using the 3Shape viewing software (Figure 2). Images of the 3D digital models were made available on dual computer screens allowing the examiners to visualize all 10 reference models at all times.

Examiners

Scoring was carried out by seven examiners comprising five Orthodontic trainees (A, B, C, D, and E) and two Consultant Orthdontists (F and G).

Examiner G was also involved in undertaking orthodontic treatment in patients with clefts at Frenchay Hospital, Bristol, UK. Prior to commencing the study, all examiners attended a calibration course in the use of the 5-year-olds' index and demonstrated a good to a very good level of agreement (0.76–0.90) using the weighted Kappa (k) value.

Each examiner independently scored the 45 patient study models using the four different formats of the reference set. The scoring took place 1 week apart for each format and for each separate scoring session, the order of the patient models was reassigned using randomly generated numbers from a number table, in order to minimize the effect of memory bias on the results. Rating was repeated 3 weeks later under similar conditions for calculation of intra-examiner agreement. For the second block rating, the order in which the reference formats were presented to the examiners was altered from the first scoring exercise.

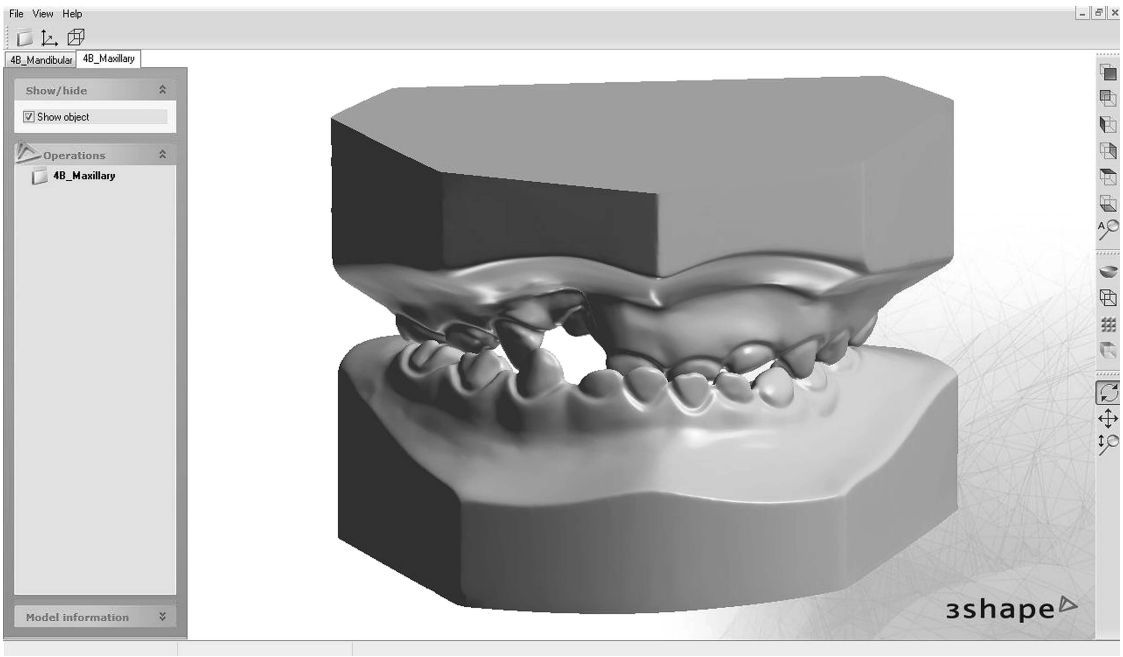

Figure 2 Three-dimensional reproduction of reference set 4b.

Results

The level of agreement was determined using Cohen's weighted kappa (*k*) statistic, which demonstrates the strength of agreement within the results (Altman, 1991). A kappa value of 1 indicates perfect agreement, whereas a kappa of 0 indicates agreement due to chance. In general, intra-examiner agreement using the four formats of the 5-year-olds' Index was moderate to very good for the seven examiners, with kappa coefficients ranging from 0.59 to 0.88 (Table 1). The plaster reference models demonstrated the greatest level of agreement for all the examiners, ranging from 0.73 to 0.86.

To determine the reliability of using the different formats, the 5-year-olds' Index scores using the plaster reference models served as the 'gold' standard for comparison with the other three formats: acrylic models, photographs, and 3D digital models (Table 2). Plaster reference models were chosen as the control as these are currently the most commonly used medium for presenting the 5-year-olds' Index reference set. All formats demonstrated a high level of agreement. However, there was a general trend across the examiners whereby lower *k* values were obtained using acrylic models in comparison to the digital formats.

Discussion

The 5-year-olds' Index is one of the most commonly used tools to assess surgical outcome in UCLP at the age of 5 years. To date, the literature suggests that it is the earliest that dental arch relationships can be assessed. The index

Table 1 Intra-examiner agreement for acrylic models, three-dimensional (3D) digital models, photographs of models, and plaster models for each examiner.

Examiner	Weighted kappa values			
	Acrylic models	3D models	Photographs	Plaster
A	0.67	0.75	0.81	0.73
B	0.61	0.67	0.74	0.83
C	0.78	0.84	0.68	0.74
D	0.67	0.77	0.75	0.84
E	0.65	0.77	0.59	0.79
F	0.78	0.88	0.70	0.84
G	0.81	0.76	0.72	0.86

Table 2 The level of agreement for acrylic models, three-dimensional (3D) digital models, and photographs using plaster reference models as the 'gold' standard.

Examiner	Weighted kappa values		
	Acrylic models	3D models	Photographs
A	0.68	0.78	0.84
B	0.69	0.91	0.76
C	0.71	0.76	0.70
D	0.79	0.82	0.82
E	0.75	0.67	0.68
F	0.83	0.85	0.90
G	0.87	0.85	0.87

itself was originally presented as 10 plaster models. As these can easily be damaged, they were also later reproduced as 10 coloured acrylic models and these are occasionally used as an alternative to the plaster models. Current practice is to transport reference models to the site of scoring. However, this can pose a number of challenges, including the risk of damage during transportation, bulkiness of models, and the need for physical storage space. Nollet *et al.* (2004), using the Goslon yardstick, investigated the reliability of rating dental arch relationships using photographs of the patient and reference models instead of plaster models. Photographs were shown to be a reliable alternative, which was further supported by a similar study using the modified Huddart and Bodenham rating system (Ali *et al.*, 2006). Photographs have the potential to eliminate the need for transportation of conventional models to a central source for scoring. They can be easily standardized and have the advantage of electronic transmission. 3D digital models also offer the promising advantage of electronic transmission with the additional benefit of manipulation of the models to allow visualization from different angles. The accuracy and reproducibility of measurements of digital models versus plaster models have been the subject of a number of investigations (Santoro *et al.*, 2003; Zilberman *et al.*, 2003; Quimby *et al.*, 2004; Stevens *et al.*, 2006). These studies compare measurements on 3D digital models and plaster models using dedicated computer software and digital callipers, respectively. In general, they all report that certain measurements on digital models can be less reliable than measurements from plaster models, but the differences are considered to be small and clinically insignificant. More recently, a comprehensive assessment of intra- and inter-examiner agreement in the measurements between plaster and 3D digital models has been investigated. Measurements of tooth dimensions and arch relationships demonstrate a high level of agreement between the plaster and 3D digital models, thus supporting the clinical acceptability of 3D digital models (Bootvong *et al.*, 2010). The clinical acceptability of digital models has been further supported by a systematic review which concluded that digital models offer a high degree of validity when compared to measurements on plaster models (Fleming *et al.*, 2010)

The 5-year-olds' Index is a subjective scoring system, using features such as the degree of overjet and presence or absence of crossbites, so a precise measurement is not necessary. In the present study, the reliability of using digital reference models compared to the 'gold' standard plaster models proved to be a good alternative, with *k* values ranging from 0.67 to 0.91. It is possible that reliability would improve if the 3D digital models of the 5-year-olds' Index were used to score patient models in the same format.

In this study, it is also interesting to note that when the different formats were compared with the plaster models scores, kappa values for the less experienced clinicians were in general lower than the more experienced examiners (examiners F and G). Despite all the examiners attending a calibration course prior to undertaking the study, it would appear that familiarity in clefts and experience of treating patients with clefts may influence the reliability of the scores. Atack *et al.* (1997) reported that consultants who were most experienced in the treatment of clefts appeared to score study models in lower groups than those who do not undertake cleft care. This, however, was not specifically investigated in this study.

One of the major arguments against the use of 3D digital models is that they are expensive and not universally accepted. However, this is changing with new electronic media rapidly evolving and becoming a more significant part of everyday clinical practice. Digital photographs and digital radiographs are already in regular use and patient records are becoming widely available in electronic format. It may only be a matter of time before 3D digital models are routinely available in hospital and dental practices. In general, a set of 3D digital models require 5 Mb of storage space. Thus, the 10 sets of the digital reference models can be easily stored on a CD-ROM or a USB flash drive for transportation.

When acrylic reference models were used in this study, *k* values were in general lower than those observed using the other formats. A possible explanation is that having colour reference models to score plaster patient models may interfere with the subjective assessment of the features used to rate dental arch relationships. Using this medium, there was a general tendency to overestimate the severity of the discrepancy between the arches.

During this study, every attempt was made to reduce the effect of recall bias on the results. Patient models were reassigned for each separate scoring. In addition, the order in which the reference formats were presented for the second block rating was changed compared to the first. This meant that the effect of any examiner fatigue on the results from repeated scoring may have been reduced, if not eliminated.

Conclusions

This study benefits from being a single study that compares the reliability of four different formats of presenting the 5-year-olds' Index. Both digital photographs and 3D digital models were found to be reliable alternatives to plaster models of the 5-year-olds' Index. At present, 3D digital models are not in widespread use but they probably offer the most promising alternative to hand-held models. With further refinements in software and technological advancements, digital models could become accepted as the norm for rating dental arch relationships in cleft lip and palate.

Acknowledgements

The authors wish to thank Joseph Burke at Joseph Burke Orthodontic Laboratory, Limerick, Ireland, for the 3D reproduction of the reference models. Simon Tutty, Medical

Illustration department, Bath, UK, for his assistance in taking photographs of the study models. Also, Sam Leary for her kind help with the statistics and Christopher Barker, Eng Ching Mah, Neil McCusker, Stephen Boyd, for all their help in scoring.

References

Ali S A, Mossey P, Gillgrass T 2006 A study model based photographic method for assessment of surgical treatment outcome in unilateral cleft lip and palate patients. European Journal of Orthodontics 28: 366–372

Altman D G 1991 Practical statistics for medical research. Chapman and Hall, London

Atack N, Hathorn I, Mars M, Sandy J 1997 Study models of 5 year old children as predictors of surgical outcome in unilateral cleft lip and palate. European Journal of Orthodontics 19: 165–170

Atack N E, Hathorn I, Dowell T, Sandy J, Semb G, Leach A 1998 Early detection of differences in surgical outcome for cleft lip and palate. British Journal of Orthodontics 25: 181–185

Bootvong K *et al.* 2010 Virtual model analysis as an alternative approach to plaster model analysis: reliability and validity. European Journal of Orthodontics 32: 589–595

Clark S A, Atack N E, Ewings P, Hathorn I S, Mercer N S 2007 Early surgical outcomes in 5-year-old patients with repaired unilateral cleft lip and palate. Cleft Palate-Craniofacial Journal 44: 235–238

Fleming P S, Marinho V, Johal A 2010 Orthodontic measurements on digital study models compared with plaster models: a systematic review. Orthodontics and Craniofacial Research 14: 1–16

Flinn W, Long R E, Garattini G, Semb G 2006 A multicenter outcomes assessment of five-year-old patients with unilateral cleft lip and palate. Cleft Palate-Craniofacial Journal 43: 253–258

Huddart A G, Bodenham R S 1972 The evaluation of arch form and occlusion in unilateral cleft palate subjects. Cleft Palate Journal 9: 194–209

Johnson N, Williams A C, Singer S, Southall P, Atack N, Sandy J R 2000 Dentoalveolar relations in children born with a unilateral cleft lip and palate (UCLP) in Western Australia. Cleft Palate-Craniofacial Journal 37: 12–16

Mars M, Plint D A, Houston W J, Bergland O, Semb G 1987 The Goslon Yardstick: a new system of assessing dental arch relationships in children with unilateral clefts of the lip and palate. Cleft Palate Journal 24: 314–322

Murray J 1998 Clinical Standards Advisory Group: cleft lip and/or palate. The Stationery Office, London

Nollet P J *et al.* 2004 Photographs of study casts: an alternative medium for rating dental arch relationships in unilateral cleft lip and palate. Cleft Palate-Craniofacial Journal 41: 646–650

Pruzansky S, Aduss H 1964 Arch form and the deciduous occlusion in complete unilateral clefts. Cleft Palate Journal 30: 411–418

Quimby M L, Vig K W, Rashid R G, Firestone A R 2004 The accuracy and reliability of measurements made on computer-based digital models. Angle Orthodontist 74: 298–303

Santoro M, Galkin S, Teredesai M, Nicolay O F, Cangialosi T J 2003 Comparison of measurements made on digital and plaster models. American Journal of Orthodontics and Dentofacial Orthopedics 124: 101–105

Stevens D R, Flores-Mir C, Nebbe B, Raboud D W, Heo G, Major P W 2006 Validity, reliability, and reproducibility of plaster vs digital study models: comparison of peer assessment rating and Bolton analysis and their constituent measurements. American Journal of Orthodontics and Dentofacial Orthopedics 129: 794–803

Zilberman O, Huggare J A, Parikakis K A 2003 Evaluation of the validity of tooth size and arch width measurements using conventional and three-dimensional virtual orthodontic models. Angle Orthodontist 73: 301–306