# A study model based photographic method for assessment of surgical treatment outcome in unilateral cleft lip and palate patients

Shaymaa Abdulreda Ali\*, Peter Mossey\*\* and Toby Gillgrass\*

\*Edinburgh Postgraduate Dental Institute and \*\*Dundee Dental School, Scotland

SUMMARY The aim of this study was to test the reliability of using digital photographs of study models as an alternative to the use of plaster study models in the assessment of surgical treatment outcome in 5-year-old children with unilateral cleft lip and palate (UCLP).

Fifty-six dental study models available from the Managed Clinical Network for Cleft Services in Scotland (CLEFTSiS) database of patients aged 5-years with non-syndromic UCLP were employed. An experienced examiner scored the plaster study models using the modified Huddart/Bodenham system. A set of digital photographs stored in the CLEFTSiS electronic patient record database of five different views of the same study models were scored by three examiners to allow calculation of interexaminer reliability. The same examiners repeated the scoring 1 month later under similar conditions to determine intraexaminer reliability and minimize the influence of memory bias on the results.

The mean kappa ( $\kappa$ ) value for the application of the modified Huddart/Bodenham system on photographs of 5-year-old UCLP study models was 0.65 ± 0.05. The mean  $\kappa$  value for the measurement of overjet on the digital photographs was 0.68 ± 0.07.

Using the interpretation suggested by Altman, good agreement for both scoring systems was found. Therefore, digital photographs of study models are a reliable alternative to measuring treatment outcome using study models of 5-year-old children with UCLP.

## Introduction

Treatment of all types of cleft lip and palate (CLP) is multidisciplinary, although there are a wide range of treatment protocols. Quality of outcome for surgical repair of CLP varies considerably. This is related to particular surgical techniques, timing and sequence of surgery, and skill of the surgeons (Semb and Shaw, 1996). There is no general agreement on which surgical protocol produces the best results. Multiple outcome measures may be used, but success by one measure may not necessarily be judged as successful using another criterion. An example for this is related to controversy over timing of hard palate repair (Dahl *et al.*, 1982; Normando *et al.*, 1992). Delayed closure of the hard palate may reduce growth disturbance but may also have a deleterious effect on speech (Witzel *et al.*, 1984).

The evaluation of treatment outcome of unilateral cleft lip and palate (UCLP) reveals several problems that make research in the field difficult (Semb and Shaw, 1996). In retrospective studies there is often difficulty in obtaining adequate information due to insufficient detail in the hospital records. For prospective studies, the time lapse between surgery and measurement of outcome extends the duration of the research.

Uncertainty about the reproducibility of outcome measurements and their validity frequently diminishes the

credibility of research findings and reduces the extent to which they can be generalized (Shaw *et al.*, 1992). The relatively low incidence of clefts, the variety of cleft subtypes, and the decentralized nature of care make only a few centres able to accumulate sufficient numbers for hypothesis testing.

Assessment of treatment outcome using study models of CLP patients is one of the most commonly used methods. The GOSLON ranking system (Mars *et al.*, 1987) can be used to examine surgical outcome in the late mixed and early permanent dentitions using study models. A similar scoring system was developed to measure outcome in 5-year-old children (Atack *et al.*, 1997). Among the limitations of the GOSLON and 5-year systems is the element of subjectivity used in making the assessment and as such calibration courses are required for those who want to use the index.

The Huddart and Bodenham system was originally applied to the study models of UCLP patients in the primary dentition (Huddart and Bodenham, 1972). The system uses the frequency of crossbites of the dental occlusion to evaluate maxillary arch constriction. The maxillary arch is divided into two buccal segments and a labial segment. In the labial segment, the lateral incisors are not assessed, as they are frequently missing or unreliable in their position. Each maxillary tooth is scored according to its relationship with the corresponding tooth in the mandible. Individual scores are then totalled for each set of models. A negative score represents maxillary arch constriction.

There are some advantages to applying assessment methods on digital photographs of dental study models, such as the reduced need for study model storage, images become easily retrievable, electronic transmission becomes feasible therefore providing a tool for treatment planning or intercentre audit, and photographic methodology can be standardized.

The advantage of using photographs of study models rather than similar views of intraoral photographs is the reproducibility of the photography methodology. Standardization of intraoral photographs is difficult due to limited accessibility and visibility as well as technical problems. Study models can be placed in any position and all areas of interest can be visualized. The major drawback of course is that impressions are required to obtain study models.

Nollet *et al.* (2004) investigated the reliability of using photographs of study casts as an alternative to casts for rating dental arch relationships. Records of children with UCLP (n = 49) at the age of 9 years were included and scored using the GOSLON ranking system. The results showed no significant differences between the rating of dental casts and photographs of dental casts. However, printed photographs were used in the study, and a subjective scoring system. For intercentre studies, the problem of accessing records will still persist when using printed photographs.

The aim of this research was to test the reliability of the use of digital photographs of study models of 5-year-old UCLP subjects as a treatment outcome assessment tool. The methodology to be used involves an objective assessment method, scoring on a computer screen, and using customized computer software.

#### Materials and methods

Fifty-six dental study models available from the Managed Clinical Network for Cleft Services in Scotland (CLEFTSiS) database of patients 5 years of age with non-syndromic UCLP were employed in this research.

A set of digital photographs stored in the CLEFTSIS electronic patient record (EPR) database of five different views of the same study models were also used. Views used for the scoring included a frontal view of the study model in occlusion, right and left buccal views in occlusion, and both upper and lower arch occlusal views.

The photographs were scored from the EPR window on a laptop located at Perth Royal Infirmary (Figure 1). To support the EPR and collaborative care system, the software program Excelicare<sup>TM</sup> (AxSys Technology, Glasgow, Scotland, UK), which includes forms, charts, and document designer toolkits tailored to multidisciplinary CLP patient management, was used. The Excelicare<sup>TM</sup> system incorporates a system of user-friendly and clinically familiar folders for the storage and

review of all clinical documents and multimedia items, including radiographic, photographic, audio, and video images.

## Examiners

Four examiners were employed for the modified Huddart/ Bodenham scoring system. Examiner R, a third-year specialist registrar in orthodontics experienced in the use of the modified Huddart/Bodenham system, provided the reference scores. Examiner A was a second-year postgraduate student in orthodontics, experienced in using the modified Huddart/Bodenham scoring system. Examiner B was a specialist registrar in orthodontics in the first year and examiner C a qualified dental hygienist.

For overjet scorings, four examiners were also used, three of whom, examiners A, B, and C were also employed for the modified Huddart/Bodenham system. Examiner A provided the reference overjet measurements from the plaster study models. Overjet measurement on trimmed study models was deemed to be an objective score, against which the photographic scores could be measured. Examiners B, C, and D measured overjet on the photographs on two different occasions 1 month apart. The additional examiner, D, was a consultant orthodontist at Edinburgh Postgraduate Dental Institute.

# The modified Huddart/Bodenham system

The study models and photographs were set out in a quiet office and scored using the modified Huddart/Bodenham scoring system. The examiners were given a reference guide which described the scoring protocol (Figure 2) and gave details of modifications to the scoring system. These modifications were as follows:

- 1. If a central incisor is missing, the other central incisor score is used.
- 2. Where primary canines are missing, the score is determined by the midpoint of the alveolar ridge.
- 3. If a primary molar is absent, then a score is allocated equivalent to the adjacent molar, if it exists. Where both molars are absent, the score is determined by the midpoint of the alveolar ridge.
- 4. At 5 years of age, the first permanent molars are not erupted and not scored; therefore, the maximum range of scores is -24 to +8.

Examiner R scored the 56 study models once only. The photographic scores were compared with these reference scores for each examiner to determine the reliability of the photographs.

Examiners A, B, and C used the scoring system on the first study models. Twenty-five study models were selected randomly and scored by the three examiners on two different occasions 1 month apart. They then scored the photographs of all 56 study models using the modified Huddart/Bodenham scoring system on two occasions 1 month apart.



Figure 1 An example of the window of Managed Clinical Network for Cleft Services in Scotland electronic patient record on which photographs were scored.



**Figure 2** Scoring of the buccolingual dental relationship (from Bongaarts C A, Kuijpers-Jagtman A M, van 't Hof M A, Prahl-Andersen B 1984 The effect of infant orthopedics on the occlusion of the deciduous dentition in children with complete unilateral cleft lip and palate (Dutchleft). Cleft Palate-Craniofacial Journal 41: 633–641, with kind permission of the Alliance Communications Group).

The same three examiners scored the digital photographs of the dental study models and repeated the scoring 1 month later under similar conditions to allow calculation of interand intraexaminer reliability.

## Measurement of overjet

The same 56 study models were scored for overjet. Examiner A provided the reference overjet measurements from the actual study models. Twenty-eight of the photographs of the models were randomly selected to obtain overjet measurements. Examiners A, B, and D measured the overjet on the photographs on two different occasions 1 month apart.

Overjet was measured on the actual study models using a ruler. When overjet was different on both incisors, the following rules were followed:

- 1. In the case of a positive overjet, the most positive overjet was recorded.
- 2. In the case of a negative overjet, the most negative overjet was recorded.
- 3. If the incisors were missing, the measurement was made from the midpoint of the alveolar ridge.

The overjet on the photographs was obtained by measuring the distance from the base of the model to the incisor of interest on both upper and lower models; the difference in



Figure 3 Measurement on upper and lower models using the electronic patient record software for overjet measurement.



Figure 4 Assembled photography stand.

the distance is the overjet. Measurements were obtained by subtraction of the lower arch from the upper arch digital measurement, and the scale on the photographs was also measured to obtain the actual distance (Figure 3). The length of the incorporated scale on both upper and lower arch views was measured to correct for magnification. Because the measurement using the software ruler was accurate to three decimal places, the measurement was rounded to the nearest 0.5 mm.

The same rules were applied when scoring overjet on the photographs in addition to the following rules: The type of overjet (positive, negative, or edge-to-edge) was checked from the antero-posterior view. If edge-to-edge, the distance to the base of each model from both incisal edges was measured. If the overjet was positive, the distances to the base of the model from the incisal edge of the tooth of interest on the upper model and from the most anterior point on the labial surface of the incisor of interest on the lower model were measured. If the overjet was negative, the distances to the base of the model from the incisal edge of the tooth of interest on the lower model and from the most anterior point on the labial surface of the incisor of interest on the upper model were measured.

## Methodology of photography

*Resources.* The study models were photographed using a Nikon Cooplix990 digital camera. This 3.34-megapixel CCD offers true (non-interpolated) image resolution of 2048  $\times$  1536. The built-in  $\times$ 3 zoom lens provides 38–115 mm coverage (35 mm equivalent) for clear, sharp images with  $\times$ 4 stepless digital zoom. It also features a 1.8-inch low-temperature polysilicon LCD screen for easy composition and playback. Daylight balance copy-stand flash was also used for photography. A box filled with sand (antero-posterior and buccal views) and a box filled with lentils (occlusal views) were used to position the study models on a black velvet cloth.

#### Method for positioning of the study models

For the occlusal views of the maxillary and mandibular models, the copy stand and lighting systems were set up at pre-determined positions (Figure 4). The camera was then mounted onto the stand and levelled. The lentil box was placed on a raised baseboard and covered with a black velvet cloth, and the model was placed along with a 1-cm scale (Figure 5).

For the antero-posterior and buccal views, the casts were placed onto the velvet and located in correct occlusion in the sand. The antero-posterior view was taken first with the anterior teeth facing the camera's lens (Figure 6A). The right and left buccal views were then taken (Figure 6B,C). A scale of 1 cm was placed in each view. A record was kept for each patient; a log of the cast number and the order of photography were also recorded.

## Statistical analysis

Intra- and interexaminer reliability was calculated using Cohen's weighted kappa ( $\kappa$ ) statistic and the degree of agreement was interpreted as described by Altman (1991). Weighted  $\kappa$  values less than 0.20 indicate poor level of agreement and between 0.21 and 0.40 a fair level of agreement, while those between 0.41 and 0.60 indicate



Figure 5 Occlusal view of the upper and lower dental arches.

moderate level of agreement, 0.61 and 0.80 good agreement, and 0.81 and 1.00 very good agreement.

Statistical calculations were carried out using Analyse-It for Microsoft Excel (General and Clinical Laboratory Statistics version 1.68, Analyse-It Software Ltd, Leeds, UK). Interexaminer reliability was calculated for the use of the system on the actual study models. Intraexaminer reliability was calculated for the three examiners when using the scoring system on study models by comparing their scores on two different occasions as described above. Interexaminer reliability was determined for individual teeth (incisors, canines, and molars) scorings.



**Figure 6** Study models in occlusion: (A) antero-posterior, (B) right buccal, and (C) left buccal views.

Table 1 Weighted kappa ( $\kappa$ ) values showing the level of agreement with the reference scorings when using the modified Huddart/ Bodenham system on photographs.

	Examiner A	Examiner B	Examiner C
Examiner R	0.63 (1st occasion)	0.57 (1st occasion)	0.58 (1st occasion)
Examiner R	0.60 (2nd occasion)	0.66 (2nd occasion)	0.67 (2nd occasion)

**Table 2** Weighted kappa ( $\kappa$ ) values showing level of agreement with reference scorings when overjet values were derived from digital photographs.

	Examiner A	Examiner B	Examiner C
Overjet	0.69 (1st occasion)	0.64 (1st occasion)	0.65 (1st occasion)
Overjet	0.64 (2nd occasion)	0.59 (2nd occasion)	0.63 (2nd occasion)

## Results

#### Huddart/Bodenham scoring

The digital photographs of the 56 sets of study models scored by the three examiners on the CLEFTSiS EPR window were compared with the reference scores. The mean score for this sample was -6.94 (range -21 to +3). Interexaminer reliability was calculated by comparing the scores of the three examiners with the reference scores (Table 1).

Twenty-eight study models were randomly selected and scored for overjet by the reference scorer. The digital photographs of the same 28 models were scored for overjet by the other three examiners on the CLEFTSIS EPR window. Interexaminer reliability was calculated by comparing the scores of the three examiners with the reference scores. Weighted  $\kappa$  values are shown in Table 2.

Table 3 shows the  $\kappa$  values for inter- and intraexaminer reliability using the modified Huddart/Bodenham scoring system on the digital photographs.

#### Overjet measurement

The inter- and intraexaminer reliability was calculated to further evaluate the reliability of overjet on the digital photographs. Table 4 shows the  $\kappa$  values for inter- and intraexaminer reliability.

## Discussion

Mossey *et al.* (2003) modified the Huddart and Bodenham system to allow scoring of the mixed and permanent dentition. This system was easy and reliable to use without the need for a calibration course, and also provided a sensitive and objective assessment of maxillary arch constriction when applied to the study models of patients

with UCLP. The authors compared and contrasted the new scoring system with the GOSLON and 5-year-old indices and reported four main advantages of this system over the other two systems:

*Objectivity combined with relative simplicity.* No clinical experience is required and therefore non-professional auxiliary staff, such as laboratory technicians, can accurately score the models. This simplifies training of the assessors, as no calibration course is required. An overall numerical score can readily calculated.

*Versatility.* The modified Huddart/Bodenham scoring system can be applied to any age, which is important as the recommendation for the appropriate age to obtain study casts varies and the GOSLON and 5-year-old indices should not be used on models other than the 10- or 5-year-old age group, respectively.

*Sensitivity.* The scale is a continuous scale of severity of arch constriction rather than a categorical scale and therefore provides a greater degree of sensitivity. As a continuous numerical scale, it is also quantifiable which means that it lends itself well to statistical analysis.

*Digital recording.* The measurements used in the modified Huddart/Bodenham scoring system lend themselves to calculations based on the assessment of digital images. This would facilitate measurement and analysis of the data and allow easy intercentre comparisons to be made.

In this electronic age there is a general move towards digital records. In orthodontics this includes digital photography, radiography, and study models. Photographic assessment of treatment outcome has been developed (Asher-McDade *et al.*, 1991) and is in current use (Johnson and Sandy, 2003), but is limited to the soft tissues of patients with CLP from extraoral photographs.

The mean  $\kappa$  value for the application of modified Huddart/ Bodenham system on photographs of 5-year-old UCLP study models was 0.65 (±0.05) and for the measurement of overjet on digital photographs 0.68 (±0.07). Using the interpretation suggested by Altman (1991), the results show good agreement.

All three examiners who scored the photographs using the modified Huddart/Bodenham system used the system on the actual study models first. They all showed good to very good reliability with  $\kappa$  values ranging from 0.79 to 0.87 (Table 5). These values confirm the reliability and reproducibility of the scoring system itself.

**Table 3** Weighted kappa ( $\kappa$ ) values for inter- and intraexaminer (italicized values) reliability when using the modified Huddart/ Bodenham system on digital photographs.

	Examiner A	Examiner B	Examiner C
Examiner A Examiner B Examiner C	0.71	0.64 0.69	0.71 0.66 <i>0.66</i>

**Table 4** Weighted kappa ( $\kappa$ ) values for inter- and intraexaminer (italicized values) reliability when overjet values were derived from digital photographs.

	Examiner A	Examiner B	Examiner D
Examiner A	0.80	0.84	0.75
Examiner B		0.65	0.69
Examiner D			0.64

**Table 5** Weighted kappa ( $\kappa$ ) values for inter- and intraexaminer (italicized values) reliability when using the modified Huddart/ Bodenham system on the plaster study models.

	Examiner A	Examiner B	Examiner C
Examiner R Examiner A	0.79 (1st occasion) 0.83	0.85 (1st occasion)	0.81 (1st occasion)
Examiner B Examiner C		0.87	0.83

When digital photographs were scored, the scores were compared with the reference scores of the actual study models by the reference scorer. The photographs proved to be a reliable tool for assessment as the  $\kappa$  values showed good to moderate reliability and ranged from 0.57 to 0.67.

The scoring system also showed good reliability when used on photographs as the interexaminer  $\kappa$  values ranged from 0.64 to 0.71. These values compare favourably with those reported by Atack *et al.* (1997; 0.49–076). However, the intraexaminer reliability  $\kappa$  values reported by Atack *et al.* (1997) were better (0.73–0.96) than those found in this study (0.66–0.71). Nevertheless, they still show good intraexaminer reliability.

#### Conclusions

Digital photographs of dental casts of 5-year-olds with UCLP proved to be a reliable tool for assessment of dental arch relationships using the modified Huddart/Bodenham system. Overjet measurement derived from digital photographs was also reliable. These findings therefore show that it is feasible to carry out surgical outcome assessment on two-dimensional digital images of casts on an EPR.

#### Address for correspondence

Peter A. Mossey Orthodontic Department Dundee Dental Hospital and School 1 Park Place Dundee DD1 4HR Scotland UK E-mail: p.a.mossey@dundee.ac.uk

#### Acknowledgement

We would like to thank Dr C. Tothill, Dr J. Asquith, and Mrs M. Davie for their kind help as examiners. We are also grateful to Mr B. Lawson for his advice with the photography methodology and for his time spent taking extra views solely for this project. We would also very much like to thank Mrs T. McDonald for all her kind help by providing her office and laptop to do the scorings.

#### References

- Altman D G 1991 Practical statistics for medical research. Chapman and Hall, London
- Asher-McDade C, Roberts C, Shaw W C, Gallager C 1991 Development of a method for rating nasolabial appearance in patients with clefts of the lip and palate. Cleft Palate-Craniofacial Journal 28: 385–390; discussion 390–391
- 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
- Dahl E, Kreiborg S, Jensen B L, Fogh-Andersen P 1982 Comparison of craniofacial morphology in infants with incomplete cleft lip and infants with isolated cleft palate. Cleft Palate Journal 19: 258–266
- 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, Sandy J 2003 An aesthetic index for evaluation of cleft repair. European Journal of Orthodontics 25: 243–249
- Mars M, Plint D A, Houston W J B, 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
- Mossey P A, Clark J D, Gray D 2003 Preliminary investigation of a modified Huddart/Bodenham scoring system for assessment of maxillary arch constriction in unilateral cleft lip and palate subjects. European Journal of Orthodontics 25: 251–257
- 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
- Normando A D, Da Silva Filho O G, Capelozza Filho L 1992 Influence of surgery on maxillary growth in cleft lip and/or palate patients. Journal of Craniomaxillofacial Surgery 20: 111–118
- Semb G, Shaw W C 1996 Facial growth in orofacial clefting disorders. In: Turvey TA, Vig K WL, Fonseca R J (eds) Facial clefts and craniosynostosis: principles and management. W B Saunders, Philadelphia, pp. 28–56
- Shaw W C *et al.* 1992 A six-center international study of treatment outcome in patients with clefts of the lip and palate: part 5. General discussion and conclusions. Cleft Palate-Craniofacial Journal 29: 413–418
- Witzel M A, Salyer K E, Ross R B 1984 Delayed hard palate closure: the philosophy revisited. Cleft Palate Journal 21: 263–269

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.