# Assessment of arch constriction in patients with bilateral cleft lip and palate and isolated cleft palate: a pilot study

Catherine Tothill and Peter A. Mossey
Orthodontic Department, Dundee Hospital and School, Scotland, UK

SUMMARY The objective of this study was to test the use of the modified Huddart/Bodenham scoring system on patients with a bilateral cleft lip and palate (BCLP) or an isolated cleft palate (CP). Study models of 5- and 10-year-old patients with BCLP ($n = 19$) and isolated CP ($n = 55$) from the Cleft Services in Scotland database were scored using the system by four examiners on two separate occasions to test inter- and intraexaminer reliability. The weighted κ statistic was used to evaluate reliability.

The interexaminer reliability ranged from 0.48 to 0.74 (moderate to good agreement) for the BCLP group and 0.77 to 0.86 (good to very good agreement) for the isolated CP group. The number of study models available for the BCLP group was small due to the low incidence of this cleft group, but the scoring system was shown to be reliable when used in cleft types other than unilateral cleft lip and palate.

The modified Huddart/Bodenham scoring system provides a useful tool suitable for assessing arch constriction in all types of orofacial clefting, which is objective, sensitive, and versatile.

## Introduction

Investigations into surgical outcome such as the Eurocleft study (Shaw *et al.*, 1992) and the Clinical Standards and Audit Group (CSAG) report (Sandy *et al.*, 1998) indicate that there is a need for improvement in cleft care in the United Kingdom. In order to improve surgical outcome, it is essential to have a tool to assess results.

One of the major problems in research into treatment effectiveness in orofacial clefting is the consequences surgery has on an array of outcomes, including speech, external facial appearance, dental relationships, craniofacial growth, ear, nose, and throat status and socio/psychological status (Bearn *et al.*, 2001). Clinical research tends to focus on each of these outcomes individually. Indices have been developed to measure treatment effectiveness relating to different aspects of anatomical form and function in the parts affected by the clefting process, often reflecting the particular interests of individual disciplines.

Many studies have used dental arch relationship to evaluate the outcome of cleft lip and/or cleft palate (CP) surgery. Poor surgical outcome tends to result in constriction or collapse of the maxilla (Semb, 1991), and therefore success or failure can be related to the dental arch relationships and the frequency with which crossbites occur. Study models provide an inexpensive, easy and minimally invasive record of the dental arch relationship. They are the most universal method of recording surgical outcome and therefore the most readily available of all records.

Several classifications have previously been described to assess dental arch relationships and therefore surgical outcome (Pruzansky and Aduss, 1964; Matthews *et al.*, 1970; Huddart and Bodenham, 1972; Mars *et al.*, 1987; Atack *et al.*, 1997). At present, no validated measurement tools exist for cleft types other than unilateral cleft lip and palate (UCLP). Mossey *et al.* (2003), however, suggested that the Huddart and Bodenham system is sufficiently versatile to be applied to models of any cleft type.

This system measures lateral and antero-posterior arch constriction and modified forms of this index have been used previously to look at arch dimensions of both CP and bilateral cleft lip and palate (BCLP). Heidbuchel and Kuijpers-Jagtman (1997) used the system in an investigation into maxillary and mandibular dental arch dimensions in BLCP patients aged 3–17 years. Nyström and Ranta (1994) also used the system when measuring the effect of timing and method of closure of CP on development of the dental arches in patients from 3 to 6 years of age. In both these investigations, other methods were also used to measure arch constriction, and the system has not been validated for use in either of these cleft types.

The aim of this study was to validate the modified Huddart and Bodenham scoring system for use in patients with isolated CP or BCLP which would provide a unique tool for measuring lateral and antero-posterior arch constriction and therefore one aspect of surgical outcome in all types of orofacial clefting.

## Materials

Study models were taken from the Cleft Services in Scotland (CLEFTSiS) record archive. CLEFTSiS is the National Managed Clinical Network for Cleft Lip and Palate Services in Scotland, and contains an archive of records of patients with orofacial clefting in Scotland. All dental study models available from the archive of patients aged 5 and 10 years with BCLP and CP were used in this investigation. The BCLP group consisted of eight study models of 5-year-old

patients and 11 study models of 10-year-old patients. The CP group comprised 29 study models of 5-year-old patients and 26 models of 10-year-old patients. All were of good quality and accurately trimmed.

The models were set out in a quiet office and scored by four examiners using the modified Huddart/Bodenham scoring system (Figure 1).

Four examiners scored the models to allow calculation of interexaminer reliability. Two examiners were specialist registrars in orthodontics (examiners A and B), one was a consultant orthodontist (examiner C), and one a qualified orthodontic technician (examiner D). The same examiners repeated the scoring 1 month later under similar conditions to allow calculation of intraexaminer reliability and minimize the influence of memory bias on the results. The inter- and intraexaminer reliability were calculated using the weighted κ statistic.

The κ statistic is a measure of agreement and allocates a score of zero if the agreement between examiners is no better than would be expected by chance. Perfect agreement gives a score of one. Scores can be negative if there is consistent disagreement. The weighted κ statistic takes into account 'near-misses', as although they are not as good as total agreement, they are better than total disagreement.

Interpretation of κ values between zero and one is arbitrary. However, it is conventional to accept the following guidelines, as described by Altman (1991). A κ value of less than 0.2 when used to rate strength of agreement between raters (intra- or interreliability) is poor agreement, and if greater than 0.8 is rated as very good. The intermediate range 0.21–0.4 is described as fair, 0.41–0.6 as moderate and 0.61–0.8 as good agreement.

## Results

Nineteen study models were available in the CLEFTSiS database of patients with BCLP. Eight of these were in the 5-year-old age group and 11 in the 10-year-old group. The mean score for the 5-year-old group was −6 (range −19 to +2) and for the 10-year-old group −7 (range −20 to +6). The weighted κ values for intra- and interexaminer reliability are shown in Tables 1 and 2.

Fifty-five study models were available in the CLEFTSiS database of patients with CP. Twenty-nine of these were in the 5-year-old age group and 26 in the 10-year-old age group. The average score for the 5-year-old group was −3 (range −22 to +8) and for the 10-year-old group −2 (range −24 to +4). The weighted κ values for intra- and interexaminer reliability are shown in Tables 3 and 4.

## Discussion

The CSAG report (Sandy *et al.*, 1998) investigated surgical outcome in patients with UCLP in the United Kingdom in terms of dental arch relationships, facial growth, aesthetics,
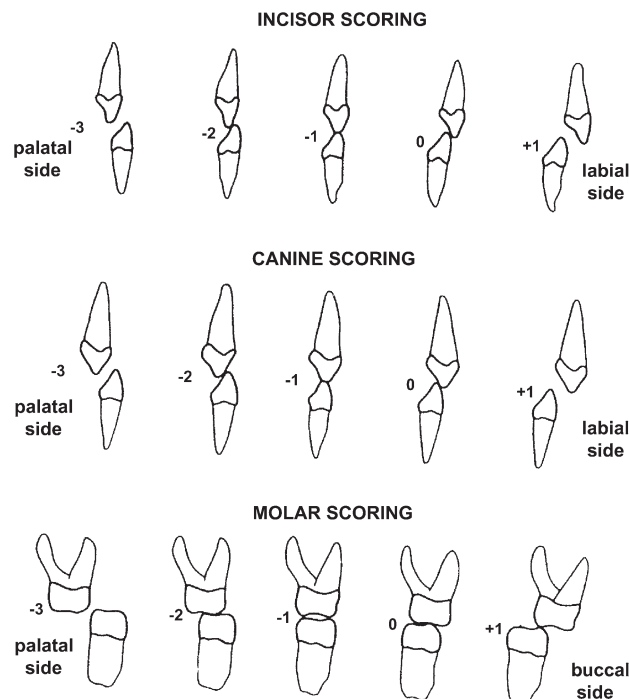


**Figure 1** The modified Huddart and Bodenham scoring system. (Reproduced with permission from Heidbuchel K L, Kuijpers-Jagtman A M 1987 Maxillary and mandibular dental-arch dimensions and occlusion in bilateral cleft lip and palate patients from 3 to 17 years of age. Cleft Palate-Craniofacial Journal, Alliance Communications Group). All maxillary teeth are scored according to their buccolingual relationship to the corresponding mandibular tooth, except for the lateral incisors, which may be missing or in an abnormal position in cleft lip and palate, subjects. The following modifications may be taken into account:

1. Premolars should be scored as for primary molars.
2. If a central incisor is missing, the other central incisor is used to score the missing incisor.
3. Where canines are unerupted, the canine score is determined by the midpoint of the maxillary alveolar ridge.
4. If a premolar is absent (for example, due to non-eruption or hypodontia), then a score is allocated equivalent to the adjacent premolar, if erupted. Where no premolars are erupted, the score is determined by the midpoint of the maxillary ridge.
5. Before the age of 6 years, the first permanent molars are not scored, even if erupted and therefore the maximum range of scores is –24 to +8. After the age of 6 years, first permanent molars are scored if present, or the midpoint of the maxillary alveolar ridge is used in a similar way as described previously. In this case, the maximum range of scores is –30 to +10.

**Table 1** Weighted κ values for 5-year-old bilateral cleft lip and palate models (intraexaminer reliability in italics).

| Examiner | A | B | C | D |
|---|---|---|---|---|
| A | *0.78* | 0.62 | 0.57 | 0.73 |
| B | | *0.52* | 0.69 | 0.74 |
| C | | | *0.31* | 0.63 |
| D | | | | *0.72* |

oral health, speech, and patient satisfaction. The results from the investigation were disappointing and recommendations were made for the centralization of cleft centres, improved

**Table 2** Weighted κ values for 10-year-old bilateral cleft lip and palate models (intraexaminer reliability in italics).

| Examiner | A | B | C | D |
|---|---|---|---|---|
| A | *0.66* | 0.68 | 0.48 | 0.69 |
| B | | *0.70* | 0.50 | 0.61 |
| C | | | *0.61* | 0.61 |
| D | | | | *0.58* |

**Table 3** Weighted κ values for 5-year-old cleft palate models (intraexaminer reliability in italics).

| Examiner | A | B | C | D |
|---|---|---|---|---|
| A | *0.82* | 0.77 | 0.71 | 0.81 |
| B | | *0.84* | 0.74 | 0.81 |
| C | | | *0.83* | 0.64 |
| D | | | | *0.80* |

**Table 4** Weighted κ values for 10-year-old cleft palate models (intraexaminer reliability in italics).

| Examiner | A | B | C | D |
|---|---|---|---|---|
| A | *0.79* | 0.74 | 0.74 | 0.66 |
| B | | *0.86* | 0.71 | 0.72 |
| C | | | *0.80* | 0.73 |
| D | | | | *0.77* |

record keeping, and organized training for staff and intercentre audit. If standards of care are to be improved in orofacial clefting, it is essential to have a means of evaluating surgical outcome, such as assessment of arch constriction.

Scoring systems to assess treatment outcome in UCLP such as the GOSLON and 5-year indices already exist and research has been undertaken to assess the effect of surgery on arch form in this cleft type. However, no validated method of assessing arch constriction exists at present for other cleft types, such as BCLP and CP. BCLP is the most severe of cleft types and is associated with the most disruption of normal growth in operated subjects. CP is the most common cleft type in Scotland with a mean incidence of 45 births per year. Primary and secondary clefts differ embryologically, and both genetic (Fraser, 1970) and epidemiological studies (Mossey and Little, 2002) suggest a separate aetiology of this cleft type. Investigation of these cleft types could provide valuable information regarding cleft care. It is essential therefore, that a method of assessing surgical outcome is also available for cleft types other than UCLP if standards are to be improved in cleft care as a whole.

The modified Huddart/Bodenham scoring system provides a simple, yet sensitive, method of assessing surgical outcome in the form of arch constriction. Validation of this tool for use in BCLP and CP will give a method of assessment of arch constriction in these cleft types and allow comparison of surgical outcome of all forms of orofacial clefting.

*Study model numbers*

Fifty-five study models were available of patients with CP in the CLEFTSiS database, while only 19 were available of patients with BCLP. This reflects the higher incidence of CP than BCLP in Scotland.

Obtaining sufficient numbers of patients for CLP research can be a significant problem due to the low incidence, the considerable variety of cleft types, and the common decentralized nature of care (Roberts *et al.*, 1991). Under-reporting can also be a problem, compounded particularly in isolated CP by the possibility that the cleft is not always diagnosed or reported at birth. It is essential that all clinicians involved in cleft care ensure that information is obtained for every child and that this is available so that comparisons can be made both within the United Kingdom and with external centres, as recommended by the CSAG report (Sandy *et al.*, 1998). In the Scottish Managed Clinical Network (CLEFTSiS), this is no longer optional, but is part of the clinician's duty of care.

With the formation of CLEFTSiS and the record registry in Perth, efforts are being made to improve record collection for cleft patients in Scotland. A record and audit protocol has been laid down corresponding to recommendations from the World Health Organisation (2002) report. Record collection is impossible for some patients, for example, due to broken appointments, refusal by the patient to participate, or the family moving away from the area. There is evidence, however, to suggest that record collection within the network is improving. The development of an electronic patient record system also means that the records are more readily available.

*Arch constriction*

In this study, the mean modified Huddart/Bodenham score for the BCLP group was −6 at 5 years of age and −7 at 10 years of age. This indicates constriction of the maxillary dental arch relative to the mandibular dental arch. The mean modified Huddart/Bodenham score for the CP group was −3 and −2, respectively. This also indicates constriction of the dental arch relative to the mandibular arch, but to a lesser extent.

Maxillary arch constriction was therefore present in both the CP and the BCLP groups, which is in agreement with previous studies (Nyström and Ranta, 1994; Heidbuchel and Kuijpers-Jagtman, 1997). Arch constriction may be related to palatal closure, which often includes incisions along the dental arches. The scars produced may induce inward deflection of the dentoalveolar processes, resulting in anterior and transverse crossbites (Semb and Shaw, 1996).

*Inter- and intraexaminer reliability*

In the BCLP group, κ for intraexaminer reliability ranged from 0.31 to 0.78 and, according to the interpretation

described by Altman (1991) was therefore fair to good. κ for interexaminer reliability ranged from 0.48 to 0.74 and was moderate to good. The CP group showed better reliability, with κ ranging from 0.77 to 0.86 (good to very good) for intraexaminer reliability and 0.64 to 0.81 (good to very good) for interexaminer reliability.

This study therefore found the modified Huddart/ Bodenham scoring system to be reliable for both the BCLP and the CP groups, although reliability was better in the CP group. There are possible reasons for the differences found in reliability of the scoring system between the two cleft types.

### Nature of the malocclusion

BCLP is the most severe of the common orofacial cleft subtypes, and is associated with the most disruption of normal growth. The maxillary hypoplasia and collapse of buccal segments found in BCLP result in a severe form of malocclusion. This means that it can be very difficult to articulate the study models, even if they have been trimmed correctly, as the teeth simply do not fit together well. The modified Huddart/Bodenham scoring system relies on examining the study models in occlusion, and if the models are not articulated in the same position on each occasion, this will affect the reliability of the score. It is important to note that the modified Huddart/Bodenham scoring system could theoretically be used directly intraorally without relying on the intermediary of study models. In patients with BCLP, the reliability may be improved by using the score directly in the patient's oral cavity as nature provides the best articulator. This also provides an option in circumstances where there are difficulties, or a lack of patient co-operation in attempting to record impressions for study models.

### Sample numbers

The number of study models available for patients with BCLP was much smaller than for those with CP. This was to be expected as the incidence of BCLP is much less. This meant that the sample size for the BCLP group was relatively small and therefore one inaccurate scoring would have a greater effect on the reliability of the whole group.

The examiners reported difficulty in scoring one particular set of study models. This was in the BCLP 5-year-old group and the severity of the malocclusion made it difficult to articulate the models accurately. The examiners gave widely ranging scores on both occasions for this set of models, and due to the small numbers this had an effect on the weighted κ statistic for the whole group. This group had the lowest weighted κ values for intra- and interexaminer reliability, and the values are quite different if that set of study models is excluded from the investigation (Table 5).

In the reduced sample, the intraexaminer reliability ranged from 0.53 to 0.83 (moderate to good) and the interexaminer reliability from 0.75 to 0.85 (good to very good).

**Table 5** Weighted κ values for 5-year-old bilateral cleft lip and palate models with one set of study models excluded (intraexaminer reliability in italics).

| Examiner | A | B | C | D |
|----------|------|------|------|------|
| A | *0.70* | 0.78 | 0.78 | 0.85 |
| B | | *0.67* | 0.75 | 0.76 |
| C | | | *0.53* | 0.77 |
| D | | | | *0.83* |

The small sample size may also have led to a greater fluctuation of the weighted κ score, and ideally this study should be repeated with a larger sample size to gain a more accurate picture of reliability of the scoring system.

### Advantages of the modified Huddart/Bodenham scoring system

Objectivity combined with relative simplicity. The modified Huddart/Bodenham scoring system is an objective continuous numerical scale. Since the scoring system uses a cumulative score which is derived from six or eight separate categorical assessments, the effect of random operator error is minimized.

The scale is also simple to use and requires no calibration, with no need for reference models. The results of this and a previous study (Mossey *et al.*, 2003) indicate that no clinical experience is required to reliably use the modified Huddart/ Bodenham system. Examiner D, a dental technician, achieved high intra- and interrater reliability scores, which compared favourably with the other examiners. The ease with which the scoring system can be learnt would simplify the training of assessors, and improve intercentre collaboration studies. The use of auxiliary staff to score the models would be more time efficient for the clinical staff involved in cleft care.

*Versatility*. The modified Huddart/Bodenham scoring system can be used on the study models of a patient of any age, which is useful in instances where the timing of recording of study models may vary from the recommended age (www.eurocran.org).

This study confirms that the scoring system may be used in types of clefting other than UCLP. No other measure of arch constriction exists at present for CP and BCLP. This is important so that outcome assessment is possible for all types of orofacial clefting. The scoring system will also allow comparison of arch constriction between the different cleft types and may provide further understanding of the specific causes of growth restriction.

The modified Huddart/Bodenham scoring system has an additional advantage in that it lends itself to the possibility of direct intraoral use without relying on the intermediary of study models. This would provide an option in circumstances where there are difficulties in attempting to record impressions for study models.

*Sensitivity*. The scoring system is a continuous scale of severity of arch constriction and is more sensitive than a categorical scale such as the 5-year-old or GOSLON indices. As a continuous numerical scale, it is also quantifiable, thus lending itself to parametric statistical analysis. In this study, the Huddart/Bodenham score was further modified to make it more sensitive. The original system scored the incisors from −3 to +1 and the canines and buccal teeth from −2 to 0. In this study, all teeth were given a score from −3 to +1, according to the degree of crossbite, to allow a more sensitive measurement of crossbites in all quadrants of the mouth.

*Digital recording*. It is possible to apply the modified Huddart/Bodenham scoring system to scanned digital images. This will allow the scoring system to be used in conjunction with the electronic patient record.

## Conclusions

The objective of this study was to test the use of the modified Huddart/Bodenham scoring system on patients with BCLP and isolated CP.

The results indicate satisfactory intra- and interexaminer reliability in the use of the modified Huddart/Bodenham scoring system in these cleft types. Taken in conjunction with other work in the field, the conclusions are as follows:

1. This scoring system can reliably be used to assess arch constriction, and therefore surgical outcome, in all types of orofacial clefting.
2. It provides an objective measure of arch constriction.
3. It is simple to use and requires no calibration.
4. It is reliable, even in the absence of clinical experience, and is therefore suitable for use by non-professional auxiliary staff.
5. The scoring system may be used to measure relative maxillary arch constriction of a patient of any age after complete eruption of the primary dentition.

## Address for correspondence

Professor 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

## References

Altman D G 1991 Practical statistics for medical research. Chapman & 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

Bearn D *et al*. 2001 Cleft lip and palate care in the United Kingdom—the Clinical Standards Advisory Group (CSAG) Study. Part 4: outcome comparisons, training, and conclusions. Cleft Palate-Craniofacial Journal 38: 38–43

Fraser F C 1970 The genetics of cleft lip and cleft palate. American Journal of Human Genetics 22: 336–352

Heidbuchel K L, Kuijpers-Jagtman A M 1997 Maxillary and mandibular dental-arch dimensions and occlusion in bilateral cleft lip and palate patients from 3 to 17 years of age. Cleft Palate-Craniofacial Journal 34: 21–26

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

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

Matthews D, Broomhead I, Grossmann W, Goldin H 1970 Early and late bone grafting in cases of cleft lip and palate. British Journal of Plastic Surgery 23: 115–129

Mossey P A, Little J 2002 Epidemiology of oral clefts: an international perspective. In: Wyszynski D F (ed). Cleft lip and palate. From origin to treatment. Oxford University Press, Oxford, pp. 127–158.

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

Nyström M, Ranta R 1994 Effect of timing and method of closure of isolated cleft palate on development of dental arches from 3 to 6 years of age. European Journal of Orthodontics 16: 377–383

Pruzansky S, Aduss H 1964 Arch form and the deciduous occlusion in complete unilateral clefts. Cleft Palate Journal 4: 411–418

Roberts C T, Semb G, Shaw W C 1991 Strategies for the advancement of surgical methods in cleft lip and palate. Cleft Palate-Craniofacial Journal 28: 141–149

Sandy J *et al*. 1998 The Clinical Standards Advisory Group (CSAG) cleft lip and palate study. British Journal of Orthodontics 25: 21–30

Semb G 1991 A study of facial growth in patients with unilateral cleft lip and palate treated by the Oslo CLP team. Cleft Palate-Craniofacial Journal 28: 1–21

Semb G, Shaw W C 1996 Facial growth in orofacial clefting disorders. In: Turvey T A, Vig K W, Fonseca R J (eds). Facial clefts and craniostenosis. Principles and management. W B Saunders, Philadelphia, pp. 28–56

Shaw W C *et al*. 1992 Intercentre clinical audit for cleft lip and palate—a preliminary European investigation. In: Jackson I T, Sommerlad B C (eds). Recent advances in plastic surgery: Churchill Livingstone, Edinburgh

World Health Organisation 2002 Human genetics programme: management of noncommunicable diseases. International collaborative research on craniofacial anomalies. Global strategies towards reducing the health care burden of craniofacial anomalies WHO, Geneva