

An aesthetic index for evaluation of cleft repair

Nicola Johnson* and Jonathan Sandy**

*Department of Orthodontics, University Dental Hospital, Cardiff and ** Division of Child Dental Health, University of Bristol Dental Hospital, UK

SUMMARY The aim of this study was to develop an aesthetic index for the evaluation of complete unilateral cleft lip and palate (UCLP) repair. The reproducibility and validity of this index were evaluated.

A pilot study was devised, in which photographic slides of 50 subjects with UCLP were evaluated by seven orthodontic examiners on two separate occasions. After determining levels of agreement, the original index was modified for use in a clinical study.

A field study was undertaken during an audit of cleft repair outcome in Perth, Western Australia. The subjects were graded using direct observation by two trained examiners. Standardized photographs of these subjects were taken at the same time. These were later graded by the same examiners independently on two separate occasions. Direct and indirect patient gradings could therefore be compared. A further field study involved the evaluation of the photographic slides of the original 50 subjects by 22 cleft surgeons. Reproducibility was determined by calculation of weighted kappa statistics.

The intra-examiner agreement ranged from moderate to good and inter-examiner agreement from fair to good. Comparison of direct and indirect aesthetic evaluation revealed that the agreement was moderate to good, and no significant systematic bias existed.

The new index offers advantages in simplicity and flexibility. Its reproducibility compares favourably with other aesthetic indices, and is not affected by the subject's age.

Introduction

Nasolabial appearance is arguably one of the most important measures of the success of treatment outcome for cleft repair (Roberts *et al.*, 1991). It reflects deformities in the lip and nose repair site, as well as the underlying growth and development of the maxilla. Residual deformities may range from minor irregularities, such as scarring of the philtral area and slight asymmetry of the vermillion, to more major stigmata such as wide alar implantation, shortening and flattening of the upper lip, retrusion of the upper lip, and flattening of the nose (Vegter *et al.*, 1997).

These distortions have an important influence on facial attractiveness, which may in turn have psychosocial implications. The psychosocial parameters of importance in cleft lip and palate (CLP) care are not clearly established (Morris and Bardach, 1989), but they include peer acceptance, social competence, educational achievement, self-esteem, and satisfaction. It is widely accepted that facial appearance is important in forming first impressions, and acts as a cue for social stereotyping (Tobiasen, 1987). Children consistently rate photographs of children with clefts more negatively on bipolar adjective scales than photographs of non-cleft children (Shaw, 1981; Schneiderman and Harding, 1984; Shaw *et al.*, 1985; Tobiasen, 1987). However, the relationship between psychosocial adjustment and facial morphology is complex and still not well understood, although some evidence suggests that a greater physical disfigurement is associated with greater internalizing behaviour (Richman, 1997).

The methods described for assessment of nasolabial appearance can be broadly divided into qualitative and quantitative. The latter aim to analyse objectively the extent of abnormal morphology and the degree of disproportion through facial measurements (Farkas *et al.*, 1993). This approach is clearly well suited to an asymmetric deformity, but potentially neglects the fact that the harmony of a person's face is more than the sum of the contributing parts (Vegter *et al.*, 1997). Qualitative methods are more subjective and analyse facial aesthetics and appearance impairment using scales, indices, scoring systems, and rankings. Subjective assessment of appearance and aesthetics is fraught with difficulties, but is most likely to reflect the patient's and general public's perception of facial impairment. Both methods have been employed directly for 'live' patient evaluation, and indirectly through various media such as photographs, videos, radiographs, and plaster casts.

The aims of this study were to develop a new simple and flexible aesthetic index to enable evaluation of nasolabial aesthetics with a single grade on a scale of 1 to 5, and to test its reproducibility and validity.

Methods

Development of the aesthetic index

The source of subjects for the development of this index was the photographic archive from the UK national Clinical Standards Advisory Group (CSAG), 1998

study. This archive offers significant advantages, since it comprises standardized, high quality photographic slides of 239 five year olds and 218 twelve year olds with repaired unilateral cleft lip and palate (UCLP). The archive was reviewed in order to gain an appreciation of the range of nasolabial appearance in repaired UCLP patients. Frontal and cleft side profile were selected as the most appropriate views, and representative subjects for each of the categories 1–5, with 1 representing excellent and 5 representing very poor, were identified (Figure 1).

In order to introduce an element of objectivity into an essentially subjective decision, criteria were devised for assessment of each view, based upon the characteristic stigmata of repaired UCLP. For the profile view, the antero-posterior relationships of the soft tissues of the lower face and the naso-labial profile were assessed. For the frontal view, the criteria included upper lip symmetry, scarring, and continuity of the vermillion border. For the nose, the criteria were symmetry of the nostrils and alar bases and centrality of the columella.

An index guide was constructed on A3 card, comprising photographic illustrations and descriptive criteria for each of the five categories.

Pilot study

A pilot study was designed to test the reproducibility of this index. Seven orthodontists from Bristol Dental Hospital were assembled, of whom two were consultants, two were senior registrars, and three were registrars. The index was explained briefly, and copies of the index guide were distributed. Fifty pairs of slides had been selected from the CSAG archive of 5 and 12 year olds to represent a broad range of nasolabial aesthetic outcome. These were presented to the panel in a randomized order for grading, and after a delay of 2 weeks, the same slides were re-randomized and re-presented to the panel.

Following the pilot study, the index was modified and refined to ensure that photographs which were consistently attributed the same grade were used to represent that grade (Appendix 1).

Primary field study

The primary field study was undertaken during an audit of cleft repair outcome at Princess Margaret Hospital (PMH) in Perth, Western Australia, following CSAG criteria and protocols (CSAG, 1998; Sandy *et al.*, 1998). These patients received pre-surgical orthopaedic treatment from birth until lip repair, which comprised an intra-oral acrylic plate plus extra-oral bonnet and strapping. At three months, primary lip repair, together with an anterior palate repair (Muir flap) and primary rhinoplasty were carried out. The palate was repaired

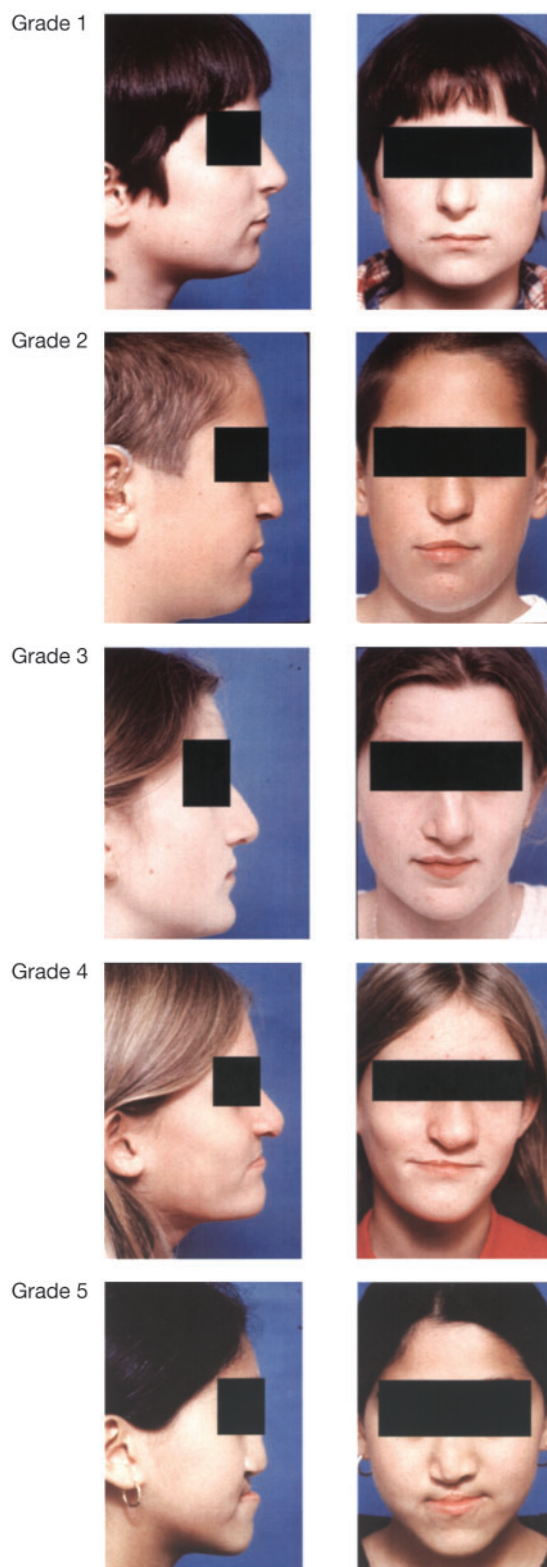


Figure 1 Aesthetic index for evaluation of cleft repair. Published with kind permission from the Clinical Standards Advisory Group.

at around 12 months using Wardill or VY push-back techniques. A single plastic surgeon undertook nearly

all cleft surgery until his retirement in 1991, after which two surgeons were involved in providing this treatment.

Two examiners used the index to independently assess all the subjects who attended the audit data collection sessions ($n = 25$). On the same occasion, standardized photographs were taken of these patients, but unfortunately only 21 were of a sufficiently high standard. These photographs were graded on return to the UK by the same two examiners, on two separate occasions in different random orders.

Secondary field study

A further field study was undertaken at a meeting of the British Association of Plastic Surgeons and involved a panel of 22 cleft surgeons. After an explanation of the index and its criteria, the same 50 pairs of slides from the original pilot study were presented for independent grading.

Statistical analysis

Reproducibility was tested by analysis of intra- and inter-observer agreement using a weighted Kappa statistic. The weighting modifies the Kappa value to reflect the relative importance of the degree of disagreement in an ordered scale, such that disagreement by one category is less important than disagreement by two categories (Roberts and Richmond, 1997).

The interpretation of the Kappa statistic follows the recommendations of Landis and Koch (1977), as shown in Table 1.

Validity of the index was tested by analysis of the agreement between the direct subject grade, and the first, second, and agreed indirect photographic grades. This was again achieved by calculation of the weighted Kappa statistic (Roberts and Richmond, 1997). A Wilcoxon signed rank test was undertaken to investigate whether any systematic bias existed between the index grade when a subject was assessed directly, and that allocated when photographs were used.

Further statistical analysis of the data from the pilot and second field studies was undertaken in order to investigate the potential influence of rater speciality upon subject grading using the aesthetic index. The median subject grades for each of the 50 subjects were

calculated separately for the first and second orthodontists' assessments from the pilot study, and for the plastic surgeons' assessment from the second field study. These medians were tested for agreement using the weighted Kappa statistic, and for systematic bias by means of a Wilcoxon signed rank test.

Results

Pilot study

Pairwise intra-observer agreement ranged from good to moderate, with a mean Kappa value of 0.57, whilst inter-observer agreement ranged from good to fair, with a mean Kappa value of 0.53.

Intra-observer agreement, when calculated separately for the 5- and 12-year-old cohort, ranged from fair to good for the 5 year olds, with a mean Kappa value of 0.58, and from moderate to good for the 12 year olds, with a mean Kappa value of 0.57. Five of the seven raters on the panel showed better intra-examiner agreement for the 5-year-old cohort than for the 12-year-old cohort, but the magnitude of the difference was variable. Inter-observer agreement for the 5-year-old cohort ranged from poor to good, and for the 12-year-old cohort from fair to very good. Both had a mean Kappa value of 0.53.

Primary field study

For the direct assessment of patients during data collection, the pairwise inter-observer agreement was good, with a weighted Kappa value of 0.71 (confidence interval, CI 0.44–0.97). For assessment of the standardized photographs, intra-examiner agreement was good to moderate, with weighted Kappa values of 0.77 (CI 0.48–1.07) and 0.44 (CI 0.17–0.72). Inter-examiner agreement was found to be fair, with weighted Kappa values of 0.34 (CI 0.08–0.60) and 0.25 (CI 0.05–0.54).

Validity of the index was tested by analysis of the agreement between subject grade by direct assessment and the first, second, and agreed photograph grades (Table 2). Agreement ranged from good to fair, and the

Table 1 Interpretation of Kappa statistic.

Kappa value	Strength of agreement
<0.20	Poor
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Good
0.81–1.00	Very good

Table 2 Intra-examiner Kappa values in validity test of aesthetic index.

	Observer 1	Observer 2
'Live patient' and 1st photographic assessment	0.50 (CI 0.22–0.77)	0.67 (CI 0.39–0.95)
'Live patient' and 2nd photographic assessment	0.50 (CI 0.23–0.77)	0.29 (CI 0.00–0.58)
'Live patient' and agreed photographic assessment	0.47 (CI 0.17–0.76)	0.61 (CI 0.31–0.91)

CI, confidence interval.

mean Kappa value was 0.51, suggesting that there is validity in the index.

A Wilcoxon signed rank test was undertaken to investigate whether any systematic bias existed between the index grade when a subject was assessed directly, and that allocated when photographs were used. The 'live' patient grade versus the agreed photographic grade was tested for each observer. No significant bias was found between direct and indirect aesthetic grading (Table 3).

Second field study

The results for the 231 pairwise inter-observer agreement events are summarized in Table 4. Agreement ranged from poor to good, with a mean Kappa value of 0.48.

The results of the comparison of orthodontists' and plastic surgeons' index grades are presented in Table 5. The Kappa values indicated good agreement. Interestingly, a significant bias was detected between grades assigned by plastic surgeons and the second assessment by the orthodontists, but not the first

Table 3 Wilcoxon signed rank test for 'live patient' grade versus the agreed photographic grade.

	Observer 1	Observer 2
Index grade 'live' < photographic	5	4
Index grade 'live' > photographic	3	2
Index grade 'live' = photographic	13	15
Z statistic	-0.71	-0.82
Significance	ns	ns

Table 4 Summary of inter-examiner agreement for the secondary field study of the aesthetic index ($n = 231$).

Weighted Kappa range	Agreement signified	Frequency	Percentage
0.00–0.20	Poor	1	<1
0.21–0.40	Fair	42	18
0.41–0.60	Moderate	176	76
0.61–0.80	Good	12	5
0.81–1.00	Very good	0	0

assessment by the orthodontists. The plastic surgeons gave the photographs numerically higher grades, and therefore worse aesthetic ratings (see Appendix 1), than the second assessment by the orthodontists.

Discussion

The development of this index was greatly assisted by the availability of a large archive of standardized photographic records. A careful study of this material, supplemented by a review of the available literature, enabled selection of representative subject photographs and identification of the pertinent criteria for each index grade. Testing of the index was also facilitated by the archive, since it was possible to select 50 pairs of high quality slides that incorporated a wide range of aesthetic outcomes.

If an index is to be a useful tool in analysis of CLP outcome it must satisfy the requirements of scientific reproducibility. Intra-examiner agreement ranged from moderate to good in both the pilot and primary field studies. Inter-examiner agreement ranged from fair to good in the pilot study, was fair in the primary field study, and ranged from poor to good in the secondary field study. Intra-examiner agreement was generally better than inter-examiner agreement.

Comparison of these agreement statistics with other published studies of CLP aesthetic indices or scales is hampered by failure to quote reproducibility statistics or by the use of different statistical methods. There are two indices suitable for comparison with this new index. First is a similar five category aesthetic scale used in behavioural research, which is reported to have very good reproducibility using intra-class correlation coefficients from mean rating scores (Richman, 1976, 1997). More thoroughly tested and reported, however, is an index developed by Asher-McDade *et al.* (1991) in which four nasolabial components (nasal form, nose symmetry, vermilion border, and nasal profile) were rated separately on five point scales. Their pilot study reported fair to good inter-examiner agreement. This index was subsequently utilized in the Eurocleft Study (Asher-McDade *et al.*, 1992), where moderate reproducibility was reported. It was further employed,

Table 5 Investigation of rater variables by comparison of orthodontic and plastic surgery index grades.

	Orthodontists' 1st assessment versus plastic surgeons	Orthodontists' 2nd assessment versus plastic surgeons
Weighted kappa	0.74	0.68
95% confidence intervals	0.19 to 1.29	-0.02 to 1.38
Orthodontists' grade > plastic surgeons	6	2
Orthodontists' grade < plastic surgeons	10	16
Orthodontists' grade = plastic surgeons	34	32
Z statistic	-1.00	-3.30
Significance	ns	significant (0.001)

in a modified form, in a Scandinavian inter-centre study in which intra-observer agreement ranged from fair to very good, but inter-observer agreement was fair (Brattström *et al.*, 1992). This index was also used in the CSAG study (1998), and panel agreement between two rating sessions was good.

The reproducibility of this new index appears to compare favourably with these previously reported indices. However, all seem to suffer from less than ideal reproducibility, and there are several possible factors that warrant further discussion. The more objective the decision-making process in ascribing an index grade, the more reproducible it is likely to be. For example, the indices used for dental arch relationship analysis demonstrate significantly better reproducibility statistics (Atack *et al.*, 1997a,b) than aesthetic indices. The assessment of facial attractiveness and nasolabial appearance, however, remains essentially a subjective opinion, which may be regarded as complex and highly individual (Asher-McDade *et al.*, 1991). Thus, in order to enhance reproducibility, this index endeavours to introduce an element of objectivity by including text criteria for each index category.

The index presented requires consideration of the nasolabial area as a whole and allocation of a score from 1 to 5. Other indices divide the nasolabial area into components, which are assessed separately and then summed for an overall evaluation (Asher-McDade *et al.*, 1991; Morrart and Shaw, 1996). The latter method is reportedly preferred by raters (Asher-McDade *et al.*, 1991), but their reproducibility statistics are not overtly better. In addition, a greater numbers of categories may make the scoring system cumbersome, and be associated with poor reproducibility (Morrart and Shaw, 1996).

It has also been suggested that replication and pooling of grades will improve reliability (Fleiss, 1986) and remove inter-examiner bias (Asher-McDade *et al.*, 1991). This is most readily achieved by using a panel of raters to generate a mean score for each subject (Asher-McDade *et al.*, 1991; Brattström *et al.*, 1992; CSAG, 1998). However, a mean value is an inappropriate measure of central tendency for categorical data, and instead a median or mode should be used.

Reproducibility is likely to be further improved by familiarization with the index and the type of material being assessed. Thus, practice rating tasks are recommended prior to commencing the grading of subjects (Howells and Shaw, 1985; Asher-McDade *et al.*, 1991). Interestingly, in the pilot study, the level of agreement was not directly related to seniority or experience of the grader.

The possibility exists that the reproducibility could be different for those subjects in the 5- and 12-year-old cohorts, since considerable changes in facial appearance occur during the years that separate these two groups. This was investigated by separate calculation of

intra- and inter-examiner Kappa statistics for the two cohorts. The intra-observer agreement ranged from fair to good for the 5 year olds, and from moderate to good for the 12 year olds, with mean Kappa values of 0.58 and 0.57, respectively. The inter-observer agreement ranged from poor to good for the 5 year olds, and from fair to very good for the 12 year olds, with a mean Kappa value of 0.53 for both. These results suggest that reproducibility is not significantly affected by subject age, and the index is equally suitable for evaluation of either age cohort. However, the question still remains whether this aesthetic index grading is valid as an early predictor of outcome. Investigation of this would require longitudinal subject data through to completion of facial growth and development, without any 'contamination' of the nasolabial appearance by secondary surgical procedures. Nevertheless, in the search for early outcome predictors, this is an area worth considering for future research.

Evaluation of the aesthetic outcome for CLP subjects is most commonly undertaken indirectly using various photographic media. The limitations of still photography are widely recognized (Asher-McDade *et al.*, 1992), since it remains a two-dimensional representation with no analysis of function or form. Standardized video recordings of the nasolabial area have been employed for appearance and function evaluation, but the reported agreement amongst plastic surgeons using this system was generally poor (Morrart and Shaw, 1996). Direct subject evaluation would overcome the reported limitations of indirect media and provide the truest assessment of a subject, yet it has not been widely considered. Previous reports of direct assessment using subjective scales are found in behavioural research (Richman, 1976, 1997; Richman *et al.*, 1985; Thomas *et al.*, 1997) and for comparisons of the outcomes of lip repair (Williams, 1968).

Some aesthetic indices developed for indirect subject assessment do not lend themselves readily to direct evaluation, especially those which utilize cropped views of the nasolabial area (Asher-McDade *et al.*, 1991; Morrart and Shaw, 1996). Examination of only the nasolabial area is advocated to remove the potential influence of background facial attractiveness upon the assessment of cleft impairment (Asher-McDade *et al.*, 1991; Tobiasen *et al.*, 1991). Both of these latter studies, however, report a strong correlation between the scores for full face and for nasolabial area, so the extent of any potential background influence seems unclear. An index that relies on masked photographs has less flexibility and uncertain benefits.

The index developed in this study has the potential for direct and indirect aesthetic evaluation. However, it cannot be assumed that an aesthetic index is suitable for both direct and indirect evaluation. Therefore, for each observer, the aesthetic grades for live subjects were compared with those for photographic slides of the

same subjects. The sample obtained for this part of the study was small, but statistical analysis revealed that the agreement was moderate to good, and no significant systematic bias existed. There are surprisingly few other studies available for comparison with these results. Direct and indirect grading has previously been undertaken as a form of validity testing for the use of photographic media in grading dental and facial attractiveness (Howells and Shaw, 1985). Only moderately high correlations were found between the dentofacial appearance of live subjects and their photographs, and statistically significant higher ratings were given to the live subjects (Howells and Shaw, 1985). However, in that study, the indirect rating was made from only one three-quarter view of the face rather than a frontal and a profile view. In addition, the subjects were non-cleft and a visual analogue scale was used. These key differences in study design may explain the variations in the results obtained.

It is interesting to consider whether raters who vary in profession might use this aesthetic index differently, and attribute significantly different grades to the same subjects. This could clearly have important implications for individual case management and outcome analysis (Eliason *et al.*, 1991). Further analysis of the data obtained in the pilot and field studies enabled comparison of orthodontists with plastic surgeons, but revealed inconsistent and therefore confusing results. The agreement was good, and for the first assessment by orthodontists no significant bias was found. However, systematic bias was found between the second assessment by orthodontists and the assessment by the plastic surgeons, with the surgeons giving worse grades. This has rarely been considered in previous reports of aesthetic indices, but a comparable study observed that the orthodontists tended to score more severely than plastic surgeons (Brattström *et al.*, 1992). Thus it remains unclear whether plastic surgeons and orthodontists use such indices similarly, but potential bias would be most readily overcome by using a mixed panel of raters.

There is some evidence to suggest that professional experience or familiarity with facial disfigurement is associated with more negative evaluation of facial appearance (Eliason *et al.*, 1991). This implies that both plastic surgeons and orthodontists would grade facial appearance more critically than non-professionals, perhaps because they react to improvements that could be accomplished by additional treatment (Eliason *et al.*, 1991). However, that study has been criticized for the vague rater instructions, the photograph selection method, and the potentially wide variation in experience, and thus lack of homogeneity of the professional group (Tobiasen, 1991). Furthermore, the data were categorical, yet the grades for each sub-group were averaged and compared using analysis of variance and

t-tests. However, if professionals are rating aesthetics differently to non-professionals, the truest measure of aesthetic impairment as perceived by society and the general public may necessitate a panel of lay people.

Conclusions

A new simple index was developed to enable ready evaluation of nasolabial aesthetics with a single grade on a scale of 1–5. The reproducibility of this index compared favourably with other aesthetic indices, and was not affected by subject age. This index offers advantages of flexibility for direct and indirect aesthetic evaluation. It remains unclear whether plastic surgeons and orthodontists use this index similarly and therefore a mixed panel of raters would be advisable.

Address for correspondence

Professor Jonathan Sandy
Division of Child Dental Health
University of Bristol Dental Hospital
Lower Maudlin Street
Bristol BS1 2LY, UK

Acknowledgements

This project was made possible by a grant from the Craniofacial Society of Great Britain for which the authors wish to express their gratitude. We are also very grateful to everyone in Bristol and PMH who participated in this study, and to Roy Powell for his statistical assistance.

References

- Asher-McDade C, Roberts C, Shaw W C, Gallagher C 1991 Development of a method for rating nasolabial appearance in patients with clefts of the lip and palate. *Cleft Palate-Craniofacial Journal* 28: 385–391
- Asher-McDade C *et al.* 1992 A six-centre international study of treatment outcome in patients with clefts of the lip and palate: Part 4. Assessment of nasolabial appearance. *Cleft Palate-Craniofacial Journal* 29: 409–411
- Atack N, Hathorn I, Mars M, Sandy J R 1997a Study models of five year old children as predictors of surgical outcome in unilateral cleft lip and palate. *European Journal of Orthodontics* 19: 165–170
- Atack N, Hathorn I S, Semb G, Dowell T, Sandy J R 1997b A new index for assessing surgical outcome in unilateral cleft lip and palate subjects aged five: reproducibility and validity. *Cleft Palate-Craniofacial Journal* 34: 242–246
- Brattström V, McWilliam J, Larson O, Semb G 1992 Craniofacial development in children with unilateral clefts of the lip, alveolus, and palate treated according to three different regimes. Assessment of nasolabial appearance. *Scandinavian Journal of Plastic, Reconstructive and Hand Surgery* 26: 313–319
- Clinical Standards Advisory Group Cleft Lip and/or Palate Report 1998 Her Majesty's Stationery Office, London, ISBN: 0113221037

- Eliason M J, Hardin M A, Olin W H 1991 Factors that influence ratings of facial appearance for children with cleft lip and palate. *Cleft Palate-Craniofacial Journal* 28: 190–194
- Farkas L G, Hajnis K, Posnick J C 1993 Anthropometric and anthroposcopic findings of the nasal and facial region in cleft patients before and after primary lip and palate repair. *Cleft Palate-Craniofacial Journal* 30: 1–12
- Fleiss J L 1986 The design and analysis of clinical experiment. John Wiley and Sons, New York, pp. 1–35
- Howells D J, Shaw W C 1985 The validity and reliability of ratings of dental and facial attractiveness for epidemiological use. *American Journal of Orthodontics* 88: 402–408
- Landis J R, Koch G G 1977 The measurement of observer agreement for categorical data. *Biometrics* 33: 159–174
- Morrant D G, Shaw W C 1996 Use of standardised video recordings to assess cleft surgery outcome. *Cleft Palate-Craniofacial Journal* 33: 134–142
- Morris H L, Bardach J 1989 Cleft lip and palate related disorders: issues for future research of high priority. *Cleft Palate Journal* 26: 141–144
- Richman L 1976 Behaviour and achievement of the cleft palate child. *Cleft Palate Journal* 13: 4–10
- Richman L C 1997 Facial and speech relationships to behaviour of children with clefts across three age levels. *Cleft Palate-Craniofacial Journal* 34: 390–395
- Richman L C, Holmes C S, Eliason M J 1985 Adolescents with cleft lip and palate: self perceptions of appearance and behaviour related to personality adjustment. *Cleft Palate Journal* 22: 93–96
- Roberts C T, Richmond S 1997 The design and analysis of reliability studies for the use of epidemiological and audit indices in orthodontics. *British Journal of Orthodontics* 24: 139–147
- 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 R *et al.* 1998 The Clinical Standards Advisory Group (CSAG) cleft lip and palate study. *British Journal of Orthodontics* 25: 21–30
- Schneiderman C R, Harding J B 1984 Social ratings of children with cleft lip by school peers. *Cleft Palate Journal* 21: 219–223
- Shaw W C 1981 The influence of children's dentofacial appearance on their social attractiveness as judged by peers and lay adults. *American Journal of Orthodontics* 79: 399–415
- Shaw W C, Rees G, Dawe M, Charles C R 1985 The influence of dentofacial appearance on the social attractiveness of young adults. *American Journal of Orthodontics* 87: 21–26
- Thomas P T, Turner S R, Rumsey N, Dowell T, Sandy J R 1997 Satisfaction with facial appearance among subjects affected by a cleft. *Cleft Palate-Craniofacial Journal* 34: 226–231
- Tobiasen J M 1987 Social judgements of facial deformity. *Cleft Palate Journal* 24: 323–327
- Tobiasen J M 1991 Commentary. *Cleft Palate-Craniofacial Journal* 28: 193–194
- Tobiasen J M, Hiebert J M, Boraz R A 1991 Development of scales of severity of facial impairment. *Cleft Palate-Craniofacial Journal* 28: 419–424
- Vegter F, Mulder J W, Hage J J 1997 Major residual deformities in cleft patients: a new anthropometric approach. *Cleft Palate-Craniofacial Journal* 34: 106–110
- Williams H B 1968 A method of assessing cleft lip repairs: comparison of LeMesurier and Millard techniques. *Plastic and Reconstructive Surgery* 41: 103–107

Appendix 1

Grade 1: excellent—good in all features

Profile: Class I soft tissue profile
 Frontal: Upper lip: symmetrical, scarring minimal, vermilion continuous
 Nose: nostrils symmetrical in size, shape, and height, columella central, alar bases symmetrical

Grade 2: good—imperfections in two or three features

Profile: Class I or very mild Class III soft tissue profile
 Frontal: Upper lip: slight asymmetry, scarring minor faults, vermilion discontinuity
 Nose: nostrils slightly flawed, columella off central, alar bases asymmetrical

Grade 3: fair—several poor features

Profile: Mild-moderate Class III soft tissue profile
 Frontal: Upper lip: asymmetry, scarring more noticeable, vermilion discontinuity
 Nose: nostril asymmetry or flattening, columella deviated, alar bases asymmetrical

Grade 4: poor—some very poor features

Profile: Moderate-severe Class III soft tissue profile
 Frontal: Upper lip: asymmetry, scarring quite noticeable, vermilion discontinuity
 Nose: nostrils very flattened, columella deviated, alar bases asymmetrical and distorted

Grade 5: very poor—overall very poor features

Profile: Severe Class III soft tissue profile
 Frontal: Upper lip: asymmetry, scarring noticeable, vermilion discontinuity
 Nose: nostrils very flattened, columella deviated, alar bases asymmetrical and distorted

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.