Reference Photographs for Nasolabial Appearance Rating in Unilateral Cleft Lip and Palate

Anne Marie Kuipers-Jagtman, DDS, PhD,*† Pieter J.P.M. Nollet, DDS, PhD,*
Gunvor Semb, DDS, PhD,‡§ Ewald M. Bronkhorst, PhD,¶ William C. Shaw, DDS, PhD,¶ and Christos Katsaros, DDS, PhD*†¶¶

Abstract: A popular method for nasolabial rating in unilateral cleft lip and palate (UCLP) is the Asher-McDade system consisting of a 5-point ordinal scale assessing nasal form, nasal symmetry, nasal profile, and vermilion border. The aim of the current study was to identify reference photographs illustrating this scale to facilitate its use.

Four observers assessed nasolabial appearance on frontal and profile photographs of the nasolabial area of 42 children of Caucasian origin with a repaired UCLP at age 9 years. Cronbach's α, based on the individual scores of the 4 observers, ranged from 0.73 to 0.82 for the 4 nasolabial ratings, indicating a good reliability. The reliability of the overall score (mean of the 4 component scores) was also high (Cronbach's α, 0.83). Both for the nasolabial component ratings and for the overall score, duplicate measurement errors were small. The reliability for the mean of the 4 observers' scores was good, Spearman rank correlation coefficients ranging from 0.56 to 0.96.

Subsequently, photographs were selected that showed the highest agreement among observers. For each of the 4 components (eg, nasal form, nasal deviation, nasal profile, and shape of the vermilion border), 5 photographs were selected to illustrate the whole range of the scale (score, 1–5), resulting in the selection of 20 pictures.

It was concluded that nasolabial appearance rating can be performed reliably using a panel of judges and averaging the scores of all observers. Reference photographs, as developed from this study, may facilitate the rating task.

Key Words: Cleft palate, nasolabial appearance, outcome, maxillofacial surgery, orthodontics

(Original Article)

Individuals with cleft lip and palate (CLP) may be confronted with a less attractive facial appearance than their peers, speech difficulties, and hearing impairment. Stigma experiences are frequently reported by youth with facial deformities. The general assumption that follows is that children with orofacial clefts are probably at risk to develop problems in their social-emotional functioning as a result of their condition. In a systematic review on psychosocial effects of CLP, however, overall, most children and adults with orofacial cleft do not experience major psychosocial problems, but some specific problems may arise such as dissatisfaction with facial appearance, depression, and anxiety. Of 11- to 13-year-olds, 75% cited teasing or bullying about their appearance as causing considerable distress. Many individuals with CLP are not satisfied with their facial appearance and desire further treatment. In 17-year-olds with repaired complete unilateral CLP (UCLP) of the Eurocleft study, lip and nose corrections were the most frequently desired options for further treatment by patients and their parents.

These studies show that facial appearance is an important factor in treatment outcome in CLP. This implies that an objective and reliable instrument for assessment of facial appearance and, more specific, nasolabial appearance is required. This is also needed when comparisons among surgical treatment procedures and treatment centers are to be made. However, there is still no widely accepted rating method to evaluate facial aesthetics in CLP. Direct clinical assessment and two- and three-dimensional approaches have been used. Rating of two-dimensional facial photographs is often used, as it is a simple, cheap, and noninvasive, and almost always available, as it is a routine record for diagnosis and treatment planning.

A method that gained some popularity over the past decades is a nasolabial rating system for UCLP, which was used in the Eurocleft, the Clinical Standards Advisory Group, and the Eurocleft follow-up studies. It was further used, sometimes in a slightly modified form, in several other studies. The scoring system is a 5-point ordinal scale assessing vermilion border, nasal form, nasal symmetry, and nasal profile. No reference photographs are available that illustrate the 5 categories of the scale. This might explain why the reliability coefficients for individual judges are rather low and the reason why the developer of the scale recommended to pool the scores of a panel of judges rather than using the individual scores.

Therefore, the aim of the current study was to identify reference photographs illustrating the Asher-McDade scale for nasolabial appearance rating in UCLP to facilitate the use of this scoring system.
PATIENTS AND METHODS

Patients

The sample consisted of 42 children of Caucasian origin, with a repaired complete unilateral cleft lip, alveolus, and palate. Patients with Simonart bands and patients with syndromes were excluded. All patients were treated at the Nijmegen Cleft Palate Craniofacial Unit from birth on. The treatment protocol has been described elsewhere.16,17

Methods

Of all patients, frontal and profile two-dimensional images were available. The mean age at which the photographs were taken was 9 years (range, 7.9–10.3 y). The pictures were loaded into PowerPoint (Microsoft, Inc, Redmond, WA), and each slide contained a profile and frontal view of one patient together with an identification number. Only the nasolabial area was shown, whereas the surrounding facial features were masked to reduce the influence of background facial appearance. In addition, duplicate photographs of 14 randomly selected patients were placed between the photographs of the 42 patients to assess the intraobserver reliability.

The Asher-McDade aesthetic index9 was used to score nasolabial appearance. In this index, 4 components of the nasolabial area are scored separately:

1. Nasal form (frontal view),
2. Deviation of the nose (frontal view),
3. Shape of the vermilion border, and
4. Nasal profile including upper lip.

To rate these 4 components, a 5-point ordinal scale is used. Score 1 represents a very good appearance; score 2, a good appearance; score 3, a fair appearance; score 4, a poor appearance; and score 5, a very poor appearance.

Four observers (A.K., P.N., G.S., and W.S.) scored all slides. Two of them (G.S. and W.S.) had additional experience with this method, as they served also as raters in the Eurocleft study. Before the experimental assessments, a training and calibration scoring session took place.

Statistical Methods

Cronbach’s $\alpha$ was calculated for each individual nasolabial component and for the sum of the 4 component scores to test for coherence among the observers. Because there was sufficient coherence (Table 1) in the further analysis, the mean scores for the 4 observers were computed for each nasolabial component separately and for the sum of the 4 component scores to reduce variability.

<table>
<thead>
<tr>
<th>TABLE 1. Reliability and Duplicate Measurement Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cronbach $\alpha$ (n = 42)</td>
</tr>
<tr>
<td>Nasal form</td>
</tr>
<tr>
<td>Nasal deviation</td>
</tr>
<tr>
<td>Nasal profile</td>
</tr>
<tr>
<td>Shape vermilion border</td>
</tr>
<tr>
<td>Overall score</td>
</tr>
</tbody>
</table>

Cronbach’s $\alpha$ is based on the individual scores of the 4 observers. The duplicate measurement error and Spearman correlation coefficient are based on the mean scores of the 4 observers for each separate component to be scored.

RESULTS

Reliability of the Method

An overview of the reliability and the duplicate measurement errors is given in Table 1. Both for the separate nasolabial component ratings and for the overall score duplicate measurement errors (n = 14) were small. The reliability for the mean of the 4 observers’ scores was good, Spearman rank correlation coefficients ranging from 0.56 to 0.96 (Table 1).

Selection of the Photographs

Based on the scores of the 4 observers for each of the 4 components (eg, nasal form, nasal deviation, nasal profile, and shape...
of the vermilion border), 5 photographs were selected to illustrate the whole range of the scale (score, 1–5), resulting in the selection of 20 pictures (Fig. 1A–D). Photographs that showed the highest inter-observer agreement were selected for the scale.

**DISCUSSION**

Subjective evaluation of nasolabial appearance does not allow comparisons of the standard of care for patients with CLP among centers, treatment protocols, and surgical techniques. In addition on patient level, the decision to surgically correct a lip or nose in a secondary procedure is still almost always based on the subjective judgment of the surgeon in conjunction with the patient and/or the parents. Therefore, the use of a standardized objective rating method for nasolabial appearance is preferred, which was already recognized by other investigators. One of the more widely accepted methods for rating nasolabial appearance in UCLP is the system developed by Asher-McDade et al. The developers of the scale suggested to compute the mean scores over the observers to reduce this variability, which we also performed in the current study.

To further facilitate the use of this scoring system, we aimed to identify reference photographs for each of the 4 components of nasolabial appearance to illustrate the 5-point scale. We were able to build 4 full-color scales (Fig. 1A–D), illustrating scores 1 to 5 (from very good to very poor appearance) for each component of nasolabial appearance. These reference photographs can help observers to calibrate themselves against the scales. Furthermore, the scales can be used as anchor sets to compare with photographs of patients who have to be rated during the scoring sessions. The method is comparable to the well-known Goslon Yardstick, which is used to categorize dental casts of patients with UCLP according to their dental arch relationship using anchor models as reference during the scoring session.

When applying the scale to patients with UCLP, an important prerequisite is to use the whole range of the scale and to keep in mind that “very good appearance” means very good for a child with a cleft. Furthermore, generally, the lip scar should be disregarded when assessing the vermilion border, as the quality of a scar is difficult to assess on a picture owing to variability in photography, lighting, and color. Standardization is the key to good medical photography. This applies to background, lighting, magnifications, and patient positioning. The Institute of Medical Illustration, which can be downloaded from www.imi.org.uk, give an excellent overview of photography of cleft audit patients. Standard views that are recommended by Eurocleft are frontal, both laterals, 3/4 oblique, and inferior columellar. The latter is not part of the Asher-McDade rating system, and this view was also not used in the current study. Although the worm’s-eye view probably is the best way to assess symmetry of the nostrils from a professional point of view, such a picture does not express a person’s whole nasolabial attractiveness, as this view is rarely shown in social circumstances.

The scales were developed for 9-year-olds with UCLP for reasons of availability of the material as in the Nijmegen CLP Unit where at age 9 years, standard documentation is made of all patients with CLP. The validity and reliability of these scales for other age groups have to be investigated further. To facilitate intercenter comparisons, it is urgently needed to reach worldwide consensus on standardization of time points for record taking. As a result of the Eurocleft study, it has been advocated to take standardized facial photographs before any surgery, at ages 5 or 6, 10, and older than 18 years for bilateral and UCLP and for patients with a cleft lip at the same time points, except at 10 years of age.

The disadvantage of still photography is of course that it gives while it assesses form and not function. In a comparative study into the reliability and validity of using two- and three-dimensional stimuli in the assessment of facial appearance in CLP, the facial deformity of the full face, lip, nose, and midface were scored using a 5-point ordinal scale. The study showed fair- to-moderate agreement between the assessment modalities, although the equivalence of using color slides and three-dimensional media relative to the clinical assessment depended on the region of the face being considered. Two other studies that compared clinical examination of the cleft deformity in the life patient with two-dimensional photographic assessment showed a moderate-to-good agreement between judgments of facial appearance. Results of studies using three-dimensional techniques for the assessment of facial morphology and growth are becoming rapidly available now. Although this technology has great potential for the documentation of the cleft face, it remains to be seen whether these techniques are applicable in large scale intercenter clinical studies. Digital two-dimensional photographs still are the most widely available clinical records, cheap to make and easy to handle. When purchasing costs of more sophisticated three-dimensional techniques decrease, the acquisition of three-dimensional images will be facilitated, resulting in an increase in its use in the future.

In conclusion, nasolabial appearance rating can be performed reliably using a panel of judges and obtaining the mean scores of all observers. Reference photographs, as developed from this study, to illustrate the rating scale for 9-year-old children with UCLP may facilitate the rating task.

**ACKNOWLEDGMENTS**

The authors wish to express their gratitude to Athina Chatziigianni, DDS, Department of Orthodontics, University of Athens, for the scanning and preparation of the patient slides.

**REFERENCES**


