The Americleft Study: An Inter-Center Study of Treatment Outcomes for Patients With Unilateral Cleft Lip and Palate Part 4. Nasolabial Aesthetics

Ana Mercado, D.M.D., Ph.D., Kathleen Russell, D.D.S., M.Sc., Ronald Hathaway, D.D.S., M.S., John Daskalogiannakis, D.D.S., M.Sc., F.R.C.D.(C.), Hani Sadek, D.D.S., Ross E. Long, Jr., D.M.D., M.S., Ph.D., Marilyn Cohen, B.A., Gunvor Semb, D.D.S., Ph.D., William Shaw, B.D.S., Ph.D.

Objective: **To compare the nasolabial aesthetics for individuals with nonsyndromic complete unilateral cleft lip and palate between the ages of 5 and 12 years.**

Design: **Retrospective cross-sectional study.**

Setting: **Four cleft centers in North America.**

Subjects: **A total of 124 subjects with repaired complete unilateral cleft lip and palate who were treated at the four centers.**

Methods: **After ethics approval was obtained, 124 preorthodontic frontal and profile patient images were scanned, cropped to show the nose and upper lip, and coded. Using the coded images, four nasolabial features that reflect aesthetics (i.e., nasal symmetry, nasal form, vermilion border, and nasolabial profile) were rated by five examiners using the rating system reported by Asher-McDade et al. (1991). Intrarater and interrater reliabilities were determined using weighted kappa statistics. Mean ratings, by center, were compared using analysis of variance.**

Results: **Intrarater reliability scores were good to very good and interrater reliability scores were moderate to good. Total nasolabial scores were Center B = 2.98, Center C = 3.02, Center D = 2.80, and Center E = 2.87. No statistically significant differences among centers were detected for both total aesthetic scores and for any of the individual aesthetic components.**

Conclusion: **There were no significant differences in nasolabial aesthetics among the centers evaluated. Overall good to fair nasolabial aesthetic results were achieved using the different treatment protocols in the four North American centers.**

KEY WORDS: *Americleft, intercenter study, nasal form, nasal symmetry, nasolabial aesthetics, nasolabial profile, treatment outcome measures, vermilion border*

Dr. Mercado is Clinical Assistant Professor, Division of Orthodontics, College of Dentistry, The Ohio State University, Columbus, Ohio. Dr. Russell is Associate Professor and Head, Division of Orthodontics, Dalhousie University, and Staff Orthodontist and Chair, Cleft Palate Team, IWK Health Centre, Halifax, Nova Scotia, Canada. Dr. Hathaway is Medical Director, Craniofacial Center, Peyton Manning Children's Hospital at St. Vincent, Indianapolis, Indiana. Dr. Daskalogiannakis is Staff Orthodontist, SickKids Hospital, and Assistant Professor, Department of Orthodontics, University of Toronto, Toronto, Ontario, Canada. Dr. Sadek is Graduate Orthodontics Resident, College of Dentistry, The Ohio State University, Columbus, Ohio. Dr. Long is Director, Lancaster Cleft Palate Clinic, Lancaster General Health, and Clinical Professor, Department of Surgery, Penn State College of Medicine, Lancaster, Pennsylvania. Ms. Cohen is Administrative Director and Patient Care Coordinator, Regional Cleft–Craniofacial Program, Cooper University Hospital, Moorestown, New Jersey. Dr. Semb is Adjunct Professor, University of Oslo, and Affiliate, Oslo Cleft Team, Department of Plastic Surgery, Oslo, Norway; and Senior Lecturer in Craniofacial Anomalies, University of Manchester, Manchester, United Kingdom. Dr. Shaw is Professor of Orthodontics and Dentofacial Development, University of Manchester, Manchester, United Kingdom.

This article was presented, in part, at the 64th American Cleft Palate–Craniofacial Association meeting in Boulder, Colorado, April 2007; the 65th American Cleft Palate–Craniofacial Association meeting in Philadelphia, Pennsylvania, April 2008; and the 66th American Cleft Palate–Craniofacial Association meeting in Scottsdale, Arizona, 2009. In addition, portions were presented at the 8th European Craniofacial Congress, Bilbao, Spain, June 2007; Craniofacial Society of Great Britain and Ireland, Chester, United Kingdom, April 2008; the Japanese Cleft Palate Association, Hirsoshima, Japan, May 2008; Cleft 2008–5th International Congress of the World Cleft Palate Foundation, Dallas, Texas, September 2008; and the International Society of Craniofacial Surgery, XIII Biennial Congress, Oxford, United Kingdom, September 2009.

The authors and Americleft group would like to acknowledge and thank the American Cleft Palate–Craniofacial Association and the Cleft Palate Foundation for their annual generous financial support from 2006 to the present. Additional funding was provided by the Trout Family Trust, Lancaster, Pennsylvania; the Mellinger Medical Research Fund, Lancaster, Pennsylvania; and the H.G. Barsumian Memorial Fund, Winston-Salem, North Carolina.

Submitted September 2009; Accepted July 2010.

Address correspondence to: Dr. K. A. Russell, Division of Orthodontics, Faculty of Dentistry, Room 5164, 5981 University Avenue, Dalhousie University, Halifax, Nova Scotia, Canada, B3H 3J5. E-mail Kathy.Russell@Dal.ca.

DOI: 10.1597/09-186.1

Patients with cleft lip and palate undergo several surgeries, beginning in infancy, to repair and restore basic morphology and function to the nose, lip, and palate. These surgeries potentially lead to tissue distortion and may further result in less-than-optimal facial aesthetics. This distorted form and lack of symmetry influence physical appearance and attractiveness and have been shown to have significant impact on psychosocial health (Jacobson, 1984). Hence, facial aesthetics is a crucial aspect in the evaluation of treatment outcomes for patients with cleft lip and palate (Shaw et al., 1985).

An enormous number of treatment protocols for the management of children with cleft lip and palate are in existence at the more than 250 cleft centers across North America (Mølsted, 1999). However, a comparison of the impact of different treatment protocols on nasolabial aesthetics has never been conducted previously in North America. The purpose of this study, as part of the Americleft Intercenter Outcome Study, was to evaluate and compare nasolabial aesthetics among four large North American cleft centers that each uses a distinct treatment protocol for patients with repaired complete unilateral cleft lip and palate (CUCLP).

Methods

The records of those patients from part 2 (Goslon scores, model study) and part 3 (craniofacial hard and soft tissue morphology) of this Americleft study were examined for good quality frontal and profile color facial images for use in the nasolabial aesthetics study. From the original sample of 154 consecutive patients, 30 were excluded due to lack or poor quality of color facial images, and thus the final sample for the nasolabial aesthetics study was 124. The sample characteristics from each center are shown in Table 1 of this paper. Center A was excluded from this part of the study due to an insufficient number of patients with the required records (fewer than 18 subjects). (The complete sample characteristics and treatment protocols for all centers are described in Table 1 of part 1 in this series of papers.) In Centers B, C, and D, photos corresponded to the same age when the records for parts 2 and 3 were taken. In Center E, however, some of the photos were taken at a younger age in relation to the records for parts 2 and 3 (an average of 1.5 years younger). Excluding the young patients from the sample of Center E would have disqualified this center from the study based on inadequate sample size. Instead, it was decided to include Center E in the nasolabial aesthetics study because, in doing so, there was an opportunity to examine dentoalveolar scores (Goslon scores), craniofacial morphology (cephalometric analysis), and nasolabial aesthetics in the four centers with the larger sample sizes.

TABLE 1 Sample Characteristics for the Nasolabial Aesthetic Outcome Study*

Center	*Sample Size*	*Mean Age (y:mo)*	*Age Range (y:mo)*	*M:F*	*Cleft R:L†*	*Images of Affected Side*
B	37	8:4	5:2–10:11	25:12	6:31	6
C	39	8:7	6:0–10:0	23:16	15:24	15
D	27	8:10	6:7–12:6	19:8	3:24	6
E	21	7:7	5:1–11:2	14:7	6:15	21

* $p < .05$.
† R = right; L = left.

All facial photographic prints or slides provided by the four centers were scanned at a high resolution (1400 dpi), and the resulting images, as well as any digital images provided, were saved as JPEGs. The images were leveled based on the interpupillary line using Adobe Photoshop 6.0 software (San Jose, CA). The eyes were then masked on each facial image in order to protect the patients' identities, and the images were coded to ensure blinding prior to the analyses. Using the same software, images were cropped to show only the nasolabial area including the inner canthus, nasal bridge, nostrils, philtrum, and upper lip (Fig. 1). Thus, surrounding features were excluded in order to eliminate any influence from facial features such as the eyes, hair, and general facial attractiveness. All backgrounds shown in the profile images were standardized to a blue shade (Adobe Photoshop 6.0 color code 88b7f6) to further prevent differentiation among centers.

The images were saved into Microsoft Office PowerPoint 2007 (Mountain View, CA) as individual slides. Because the size of the resulting images varied according to the patients' sizes, target distance from lens, and type of media (slide, print, or digital), all the images were scaled to equal dimensions with proportions preserved using the PowerPoint object scaling tool. Each slide contained a single patient's frontal and profile images (Fig. 1). The final PowerPoint file, containing all 124 images to be assessed, was saved on CDs. The slides were coded and the order of the slides was randomized among the CDs, eliminating patient or center identifiers. Thus, the raters had access to view and rate the images on their own computers.

Five raters were recruited for the study: four orthodontists and one plastic surgeon, all experienced in cleft care. For each image, four features were rated: nasal form,

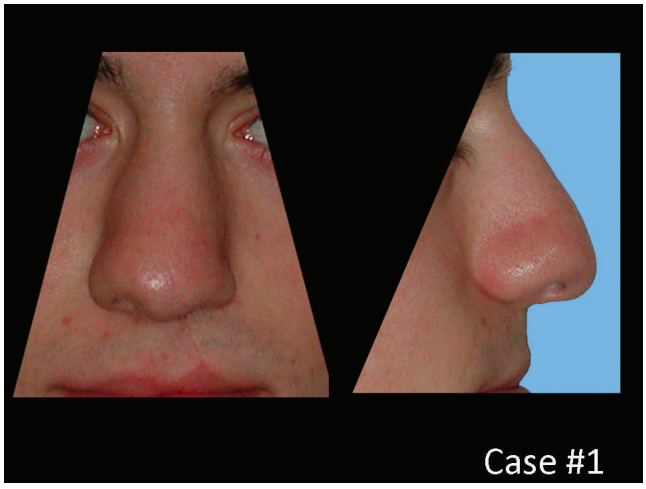

FIGURE 1 Example of a slide to be rated, with matched frontal and profile cropped images.

symmetry, profile including the upper lip, and the shape of the vermilion border (Asher-McDade et al., 1991). The raters were asked to rate each feature on a 5-point scale as follows: 1 = *very good*, 2 = *good*, 3 = *fair*, 4 = *poor*, and 5 = *very poor*. Color reference pictures for each feature were provided on paper for the raters to distinguish between the categories on the severity scale from 1 to 5 (Kuijpers-Jagtman et al., 2009).

Raters were allowed to browse through all slides prior to rating them so as to become familiar with the types of images and features to be evaluated. A practice session with 10 slides was conducted with the raters to allow for calibration in the Asher-McDade system.

Statistics

To assess for intrarater reliability, the five raters were asked to rate an additional 40 randomly selected images. Weighted kappa statistics were performed to evaluate intrarater and interrater reliability.

Ratings of each patient's features by the five raters were averaged. For each image, a total nasolabial score also was calculated by averaging the means of all features. Comparison of least square means among the different centers was performed using analysis of variance followed by the Tukey-Kramer test. These analyses also were conducted on the nasolabial profile scores of images that had the profile pictures taken strictly from the affected (cleft) side to evaluate whether there were any differences in the nasolabial profile outcome between the cleft and noncleft sides. The statistically significant level was set at $p < .05$.

Results

The intrarater reliability test revealed an overall agreement of good to very good (mean, .739; range, .638 to .823) for each of the five raters. The interrater reliability test showed moderate to good overall agreement of five raters (mean, .645; range, .587 to .708) (Landis and Koch, 1977).

The frequency of each raw score from 1 to 5 for each center is shown in Figure 2a. The mean scores for each feature for each center are shown in Table 2 and Figure 2b. Scores for nasal form ranged from 3.05 to 3.45 and are considered fair in the severity scale. There were no statistically significant differences among centers for nasal form. Nasal form was scored as 3 or less in 54% to 67% of patients from four centers. Nose symmetry scores ranged from 2.50 to 2.86 and are considered good. Again, there were no significant differences among centers for nose symmetry. Nose symmetry was scored as 3 or less in 77% to 86% of patients from four centers. Scores for vermilion border ranged from 2.72 to 3.14 and are considered good to fair. Once more, as for nasal form and nose symmetry, there were no significant differences between centers for vermilion border. Vermillion border was scored as 3 or less in 60% to 79% of patients from four centers.

Scores for nasolabial profile ranged from 2.52 to 3.00 and are considered good to fair. This feature was scored as 3 or less in 69% to 78% of patients from four centers.

When comparing all profile images including those taken from both the noncleft and cleft sides, there were no statistically significant differences among centers. After eliminating profile images taken from the noncleft side, the sample sizes from Centers B, C, and D were reduced dramatically because photographs at these centers were routinely taken from the right profile for most patients (regardless of the side of the cleft; Table 1). At Center E, all profile images were taken according to the side of the cleft, thus allowing for the entire sample of 21 to be evaluated. Intercenter comparison of images taken strictly from the affected cleft side revealed no statistically significant differences among centers.

The scores for total nasolabial aesthetics were derived from the sum of the component scores (nasal form, nasal symmetry, nasolabial profile with all images, and vermilion border). Total scores ranged from 2.80 to 3.02 and are considered good to fair on the severity scale. Total scores were 3 or less in 69% to 73% of patients from four centers. There were no significant differences for the total nasolabial scores among the four centers.

Discussion

The original sample size from the four centers was 154, but this was reduced to 124 for the nasolabial study due to lack of facial images or to inadequate facial images. Although the frontal and profile facial photographs that were used in the study are noninvasive and commonly used diagnostic records, it is surprising that all centers did not have these images for their patients. Finally, the image quality of some slides, Polaroid prints, and digital images was less than adequate and led to further exclusions from the sample. A worm's-eye view of the nose also would have aided in evaluation of nostril symmetry; however, this view was not consistently available for the centers and was not therefore included in the study. This is a consideration for future research endeavors. The practice of leveling the frontal images based on the interpupillary line is a rather artificial way to compensate for differences in head tilting and posture. None of the centers used any kind of head stabilizer to standardize the position of the patients' heads. This is one of the shortcomings of nonstandardized, retrospective intercenter studies: Records obtained by one center may be similar to those obtained by other centers but not identical in quality and modality. A good suggestion for future prospective studies is to create a method for standardized record taking that could be used at each participating center.

Nasolabial aesthetic outcome evaluation carries a high degree of subjectivity. Moreover, using still photographic images to assess aesthetics has many limitations (Asher-McDade et al., 1991). Recognizing these limitations, the Asher-McDade method that has been used in previous studies for the evaluation of nasolabial aesthetics and has

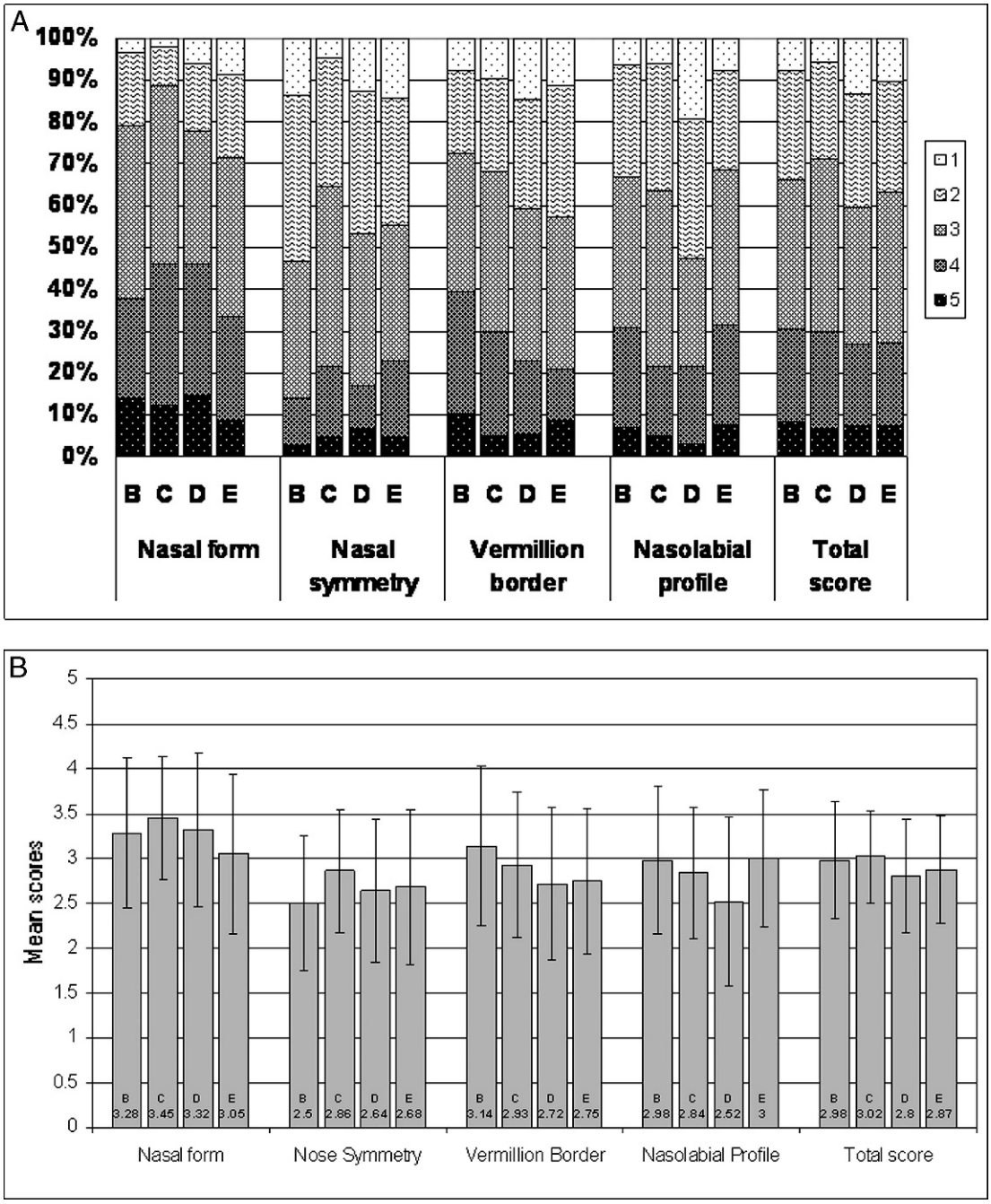

FIGURE 2 Ratings scores. A: Frequency of each raw score from 1 to 5 for each center. B: Mean scores for each feature for each center.

been shown to provide a reasonably reliable and reproducible rating system was used in this study (Asher-McDade et al., 1991). Besides the method's assessment of vermilion border and nasolabial profile, there are more characteristics of the repaired upper lip that could be evaluated (scar, philtrum, white line). Evaluation of such detailed characteristics requires the use of photographic images taken with high-resolution cameras under standardized lighting and patient positioning. In retrospective intercenter studies, it is impossible to control the consistency in the resolution, brightness, contrast, and sharpness of the images that the centers provide. Some images were scanned from Polaroid photos, others were scanned from slides, and some were taken with digital cameras of different models. It was determined that the images included in this intercenter study were of adequate quality for the evaluation of the continuity of the vermilion border, the symmetry of the cupid's bow, and the lateral profile of the upper lip, based on the Asher-McDade method. Future prospective intercenter studies should standardize the type of image-capturing method used by the centers so that the upper lip can be evaluated in greater objective detail.

TABLE 2 Least Square Mean Scores (Standard Deviations) for Each Center; There Were No Significant Differences Between Centers at the $p < .05$ Level

	Center			
	B	*C*	*D*	*E*
Nasal form	3.28 (0.84)	3.45 (0.69)	3.32 (0.86)	3.05 (0.89)
Nose symmetry	2.50 (0.75)	2.86 (0.69)	2.64 (0.80)	2.68 (0.87)
Vermilion border	3.14 (0.89)	2.93 (0.81)	2.72 (0.85)	2.75 (0.81)
Nasolabial profile (all images)	2.98 (0.82)	2.84 (0.73)	2.52 (0.94)	3.00 (0.77)
Nasolabial profile (images from affected cleft side only)	3.33 (0.69)	3.21 (0.66)	3.10 (1.08)	3.00 (0.77)
Total score (all images)	2.98 (0.65)	3.02 (0.51)	2.80 (0.63)	2.87 (0.60)

An overall intrarater reliability of good to very good was achieved (.638 to .823). The lowest overall agreement was achieved by the plastic surgeon (.638); whereas, the highest overall agreement was achieved by one of the orthodontists (.823). The moderate to good range of interrater agreement from .587 to .708 (mean, .645; five raters) was consistent when comparing the four orthodontists and also the various raters (orthodontists and plastic surgeon). These results are similar to the level of reliability achieved by previous studies using this rating method (Asher-McDade et al., 1991; Asher-McDade et al., 1992; Williams et al., 2001; Brattström et al., 2005; Nollet et al., 2007). Ideally, more standardized and objective assessment methods will be developed in order to improve the accuracy and reliability of evaluations of nasolabial aesthetic outcomes.

The component features of the nasolabial aesthetic outcome did not differ significantly among the centers. It was recognized that three of the four centers (B, C, D) submitted profile images that were not consistently taken from the cleft (affected) side; only Center E submitted images taken strictly from the cleft side. This introduced possible error into the study from incompatible images. Nasolabial profile rating, therefore, was reevaluated using only the profile images taken from the cleft (affected) side, and still there were no significant differences among centers. As expected, the ratings of the nasal images from the cleft side from Centers B, C, and D were higher (worse) than the ratings of all images combined from both the cleft and noncleft sides. Although evaluating the profile images from the cleft side should be more discriminating of the associated deformity, the results need to be interpreted with caution because the adjusted sample sizes for the centers varied considerably, and some were lower than optimal. In Center B only 19% of the profile images were taken from the affected side, compared with 38% in Center C, 22% in Center D, and 100% in Center E. Comparing nasal profile aesthetics using the original samples introduced analyses bias into the study from variation in image type; whereas, comparing the samples using only cleft-side profile images introduces sample bias from variation in and small sample sizes. Thus the results should be viewed with caution. It is recommended that cleft palate centers routinely take profile images from both the cleft and noncleft side for possible use to evaluate nasolabial aesthetic outcomes.

All centers produced overall good to fair nasolabial results with no statistically significant differences detected among the centers for total nasolabial score (range, 2.80 to 3.02). This suggests that the different protocols described in Table 1 of part 1 of this series, involving presurgical orthopedics, primary bone grafts, various types and timing of lip and palate surgery, and various numbers of surgeons, did not necessarily influence the overall nasolabial aesthetic outcome. On the other hand, it is possible that the Asher-McDade rating system lacked sufficient sensitivity to evaluate subtle differences in the outcomes among centers. Figure 2b shows that there is great variability among patients with clefts in all centers. This could have contributed to the lack of statistically significant differences among centers. However, the present study intended to mirror the design of the Eurocleft studies by using similar outcome measures proven to be reliable and reproducible, such as the Asher-McDade rating system. In future prospective studies designed to control the uniformity in the quality of the photos from all centers, it will be possible to look at more detailed objective measures of the lip and nose for each patient.

It is important to note that most of the patients in Centers B, D, and E had secondary lip and nose revisions prior to the timing of the photographs used for rating; whereas, patients from Center C had secondary surgical revisions after the time of the photographic records used in the study. Images provided by Center C, therefore, showed the aesthetic outcomes from the primary surgeries without any enhancements from lip and nose revisions. Centers B, D, and E revised the results of the primary surgery prior to when the images were taken; yet, no significant improvement for nasolabial outcome was seen compared with the results from Center C. The addition of surgical revisions for three centers did not show significantly improved nasolabial aesthetics compared with the other center and thus was not shown to be comparatively beneficial.

It may be speculated that the protocol in Center C produces comparatively better aesthetic outcomes prior to age 4, the time of revision surgeries at the other centers. By the age the photographic images were taken at Centers B, D, and E, surgical revisions had occurred that could mask potential intercenter differences that existed prior to this age when revisions were completed. Comparison of facial images before any surgical lip and nose revisions could potentially reveal intercenter differences in nasolabial aesthetic outcome related to primary management protocols that were not identified in this study. Future studies should control for this aspect of treatment regarding the timing of photographic records related to surgical revisions and the age when images are taken.

As is characteristic of many retrospective studies, sample sizes and the comparability of the centers can be viewed as limitations of this study. Optimal sample size for inter-center comparisons is estimated between 30 and 40 subjects per

center (discussed in Part 1 of this series). When evaluating the original cohort of centers (A, B, C, D, and E), it was found that center A had a small number of subjects (18), and this number would be further reduced by the lack of facial images available from this center. Therefore, center A was excluded from the present study of nasolabial esthetic outcomes based on lack of records. Centers D and E had suboptimal sample sizes (27 and 21, respectively). They were not excluded from the present study based on sample size because it was considered valuable to do an 'initial' comparison of the nasolabial outcomes among the four larger centers used in the dental arch relationship study (GOSLON, Part 2) and the craniofacial morphology study (cephalometrics, Part 3). The facial images of patients in Center E were taken at a younger age than the records of the same patients taken for Parts 2 and 3 of this series. Indeed, in the present study, patients from Center E were, in average, younger than those from the other centers. Apart from the assumption that some of the young patients in Center E may not have reached the age for lip revisions according to the center's treatment protocol, it is unclear what the effect of the age difference could be on the results. Center E was not excluded from the present study on the basis of age differences in order to compare outcomes consistently among the same four centers with the larger sample sizes. Results from the present study should thus be interpreted with caution in view of the above limitations.

Although the overall nasolabial outcomes observed in the four North American centers studied ranged from good to fair, centers should strive to improve upon their own results in order to achieve outcomes that are consistently superior (ideally, very good). Several centers advocate methods of primary cleft lip and palate management that offer significantly improved nasolabial aesthetics, such as primary nasal repair and nasoalveolar molding. None of the participating centers in the present study performed these two procedures. Therefore, we are unable to speculate on the effect that those methods could have had on the outcomes of our sample. Furthermore, there currently are no objective, controlled, randomized intercenter comparative studies supporting the use of any different primary management methods that could have improved the outcomes obtained in this study. In the future, Americleft studies should be expanded to include centers that use nasoalveolar molding and other surgical/orthopedic methods, evaluating and comparing their outcomes with the same scrutiny that was used in this study. Although the investigators in the present study were affiliated with four different centers with different management protocols, a concerted effort was made to exclude preconceived experimenter bias that could have tainted the data obtained and obscured the true results of the study. Experimenters involved in this study strived to be objective, uniform, and nonbiased in measuring and comparing outcomes. It is hoped that the results obtained in such nonbiased, retrospective, cross-sectional comparisons can be used to design a randomized, controlled clinical trial of primary management protocols that can fairly assess nasolabial aesthetic outcomes of cleft lip and palate.

Conclusions

There were no significant differences for nasolabial aesthetics and its components (i.e., nasal form, nasal symmetry, nasal profile including the upper lip, and the shape of the vermilion border) among the four centers. It was identified that comparisons among the centers should control for the sidedness of the profile images related to the location of the cleft as well as the age when images are taken relative to aspects of treatment such as lip and nose revision surgery. Accounting for these variations in both the timing and the characteristics of records analyzed could identify variations in nasolabial aesthetic outcomes related to treatment protocols not identified in this study.

Acknowledgments. The authors of the Americleft study would like to acknowledge Dr. Robert Rashid for his assistance with the statistical analysis of the nasolabial aesthetics data.

References

Asher-McDade C, Brattström V, Dahl E, McWilliam J, Molsted K, Plint DA, Prahl-Andersen B, Semb G, Shaw WC, The RP. A six-center international study of treatment outcome in patients with clefts of the lip and palate: part 4. Assessment of nasolabial appearance. *Cleft Palate Craniofac J*. 1992;29:409–412.

Asher-McDade C, Roberts C, Shaw WC, Gallager C. Development of a method for rating nasolabial appearance in patients with clefts of the lip and palate. *Cleft Palate Craniofac J*. 1991;28:385–390, discussion 390–391.

Brattström V, Mølsted K, Prahl-Andersen B, Semb G, Shaw WC. The Eurocleft study: intercenter study of treatment outcome in patients with complete cleft lip and palate. Part 2: craniofacial form and nasolabial appearance. *Cleft Palate Craniofac J*. 2005;42:69–77.

Jacobson A. Psychological aspects of dentofacial esthetics and orthognathic surgery. *Angle Orthod*. 1984;54:18–35.

Kuijpers-Jagtman AM, Nollet PJ, Semb G, Bronkhorst EM, Shaw WC, Katsaros C. Reference photographs for nasolabial appearance rating in unilateral cleft lip and palate. *J Craniofacial Surg*. 2009;20:1683–1686.

Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.

Mølsted K. Treatment outcome in cleft lip and palate: issues and perspectives. *Crit Rev Oral Biol Med*. 1999;10:225–239.

Nollet PJ, Kuijpers-Jagtman AM, Chatzigianni A, Semb G, Shaw WC, Bronkhorst EM, Katsaros C. Nasolabial appearance in unilateral cleft lip, alveolus and palate: a comparison with Eurocleft. *J Craniomaxillofac Surg*. 2007;35:278–286.

Shaw WC, Brattström V, Mølsted K, Prahl-Andersen B, Roberts CT, Semb G. The Eurocleft study: intercenter study of treatment outcome in patients with complete cleft lip and palate. Part 5: discussion and conclusions. *Cleft Palate Craniofac J*. 2005;42:93–98.

Shaw WC, Rees G, Dawe M, Charles CR. The influence of dentofacial appearance on the social attractiveness of young adults. *Am J Orthod*. 1985;87:21–26.

Williams AC, Bearn D, Mildinhall S, Murphy T, Sell D, Shaw WC, Murray JJ, Sandy JR. Cleft lip and palate care in the United Kingdom—the Clinical Standards Advisory Group (CSAG) study. Part 2: dentofacial outcomes and patient satisfaction. *Cleft Palate Craniofac J*. 2001;38:24–29.