Defining the Aesthetic Range of Normal Symmetry for Lip and Nose Features in 5-Year-Old Children Using the Computer-Based Program SymNose

Nirvana S. S. Kornmann, MD1,2, Robin A. Tan, BSc2, Frans J. Mulder, MD2, Joseph T. Hardwicke, PhD, FRCS(Plast.)1, Bruce M. Richard, FRCS(Plast.), FRCS1, Brian B. Pigott, BAI, DIC, MICE3, and Ronald W. Pigott, FRCSI, FRCS(Plast.)4

Abstract

Objective: To provide a normal comparison group against which to judge symmetry results after cleft surgery and to introduce the thin lip correction (TLC) feature in SymNose. A lip–aspect ratio algorithm has been added to the latest version of SymNose to compensate for the higher degree of overlap in thicker lips when compared to thin lips.

Design: Retrospective analysis of symmetry in healthy participants, using the computer-based program SymNose on both anteroposterior (AP) and base view images. Photographs of 91 noncleft children were traced twice by 3 independent investigators experienced with SymNose.

Participants: Five-year-old healthy participants from a local state school in Tavistock (West Devon, United Kingdom).

Main Outcome Measure: Asymmetry expressed as the perimeter mismatch percentage for nose and lip features on AP view images and for nose features on base view images.

Results: The perimeter mismatch reference range for the nose (AP view) was 2.65% to 30.91%, for the lip 2.13% to 15.44%, for the nose (base view) 1.69% to 14.84%, for the nostrils 4.68% to 26.6%, and for the width–height ratio 1.15% to 1.80%. The perimeter mismatch percentage for the lip without TLC was significantly higher compared to the perimeter mismatch percentage with TLC (P < .001).

Conclusion: This article provides a noncleft reference range for all perimeters drawn from SymNose against which to compare results after cleft surgery at 5 years of age. Furthermore, it shows the importance of correcting for variance in lip volume per child.

Keywords
cleft lip, SymNose, thin lip compensation, normal controls

Introduction

The recent Cleft Care UK study has demonstrated improved outcomes in dentofacial growth and speech in unilateral cleft lip and palate children treated in designated cleft units after the government’s reconfiguration of cleft services in the United Kingdom after the Clinical Standards Advisory Group (CSAG) report of 1998 (Di Biase and Markus, 1998; Al-Ghatam et al., 2015; Ness et al., 2015; Persson et al., 2015; Sell et al., 2015). These improvements were measurable because the 5-Year Index for dentoalveolar relationship and the Cleft Audit Protocol for Speech (CAPS-A) are both reliable and validated outcome measures. The changes in facial aesthetic outcomes were
less impressive as the outcome tool used is less robust (Sharma et al., 2012; Mosmuller et al., 2013). To assess changes in aesthetic outcomes after cleft surgery, there is a need for a reliable outcome measure.

Three-dimensional (3D) imaging is frequently suggested as the hoped-for long-term objective measure of facial aesthetics, but to date, tools are not yet widely available for clinical use. The aesthetic assessment is commonly performed on 2-dimensional (2D) photographs using some form of the Asher-McDade system with a 5-point Likert scale ranging from “excellent” to “very poor” (Asher-McDade et al., 1991). However, the inter- and intrarater reliability of this scoring system remains only moderately reliable and is only slightly improved by using reference photographs (Kuipers-Jagtman et al., 2009; Mercado et al., 2015) or assessment of discrete lip and nose (Mosmuller et al., 2014; Deall et al., 2016).

Symmetry in the areas close to the midline seems to play an important role in facial aesthetics (Springer et al., 2007). In order to measure the asymmetry of the lip and nose, thus providing a more objective aesthetic outcome measure after cleft surgery, Pigott and Pigott (2010) developed the computer-based program SymNose. This program allows measurement of asymmetry on 2D images by tracing the outline of the upper lip and the lower border of the nose on frontal view images and by tracing around the alar bases over the upper nasal perimeter on base view images. By reflecting the left side of the midline over the right, the percentage mismatch of the nonoverlapping area is calculated. As this program enables rapid and reliable assessment of the lip and nose, SymNose has proven to be a useful tool in the measurement of asymmetry after unilateral and bilateral cleft lip and palate repair (Freeman et al., 2013; McKearney et al., 2013; Russell et al., 2014). To compare their results, these studies used control groups of noneleft children.

Over the years, the program has been under constant development in order to improve the accuracy of “measuring” the aesthetic outcome after cleft surgery. As visible scarring plays a role in facial aesthetics, SymNose was further developed to enable the calculation of a subjective scar area between the central half of the upper lip and the lowest outline of the nose (Pigott and Pigott, 2016). Another topic of interest has been variations in lip volume per child. To compensate for the higher degree of overlap in thicker lips when compared to thin lips, a lip–aspect ratio algorithm has been added to the latest version of the program. The lip–aspect ratio is incorporated in SymNose by automatically dividing the lip horizontally by a great number of vertical columns giving a different height of vermillion for each vertical line across the lip from one commissure to the other. Subsequently, each height is averaged and then divided by the intercommissure distance, providing a linear relationship for the aspect ratio. The objectives of this study were to provide a normal comparison group against which to judge results after cleft surgery and to introduce the thin lip correction (TLC) feature.

### Materials and Methods

For this study, the following equipment were used:

- SymNose (version 6.30; © Brian Pigott 2007-2015)
- Apple iMac (Intel chip) running Mac OSX 10.8.5 or later
- Digitizing pad
- Apple iWorks or Microsoft Office for Mac
- Adobe Photoshop Elements software (Adobe Systems Inc, San Jose, California)

### Participants

Both anteroposterior (AP) and base view photographs of 117 healthy participants were obtained, subdivided into 62 males and 55 female participants. After excluding poor quality images, AP view images of 48 males and 43 females and base view images of 48 males and 43 females were left for assessment, resulting in 91 AP photographs and 91 base view photographs in total. The Index of Multiple Deprivation for West Devon ranged from 9.057 to 32.064 (3rd-10th decile), with the majority of postal codes (90%) in the 4th to 7th decile (Ministry of Housing, Communities and Local Government).

### Protocol

Both AP and basal view images were taken according to a standard protocol, using the same camera and similar lighting. Children were instructed to keep a neutral facial expression. This was important as children in this age category are known to either press their lips together, resulting in an even smaller upper lip, or keep their mouth open, showing a larger proportion of the upper lip and thus a larger upper lip volume. Especially the pressed thin lip may influence the normal reference range, despite the TLC feature. After cataloguing photographs, poor quality images were rejected. Quality was considered poor when photographs were out of focus or in case of saliva or mucous obstructing the view of the nose or lip. All original photographs were cropped rectangular-shaped using Photoshop Elements software (Adobe Systems Inc, San Jose, California) showing only the medial canthi, nose, and lips.

Images were assessed by 3 independent investigators (N.S.S.K., R.A.T., and F.J.M.) experienced with SymNose. Rater 1 (N.S.S.K.) traced all images twice, rater 2 (R.A.T.) traced all male images twice, and rater 3 (F.J.M.) traced all female images twice. Repeat tracings were performed with a minimum interval of 2 weeks. Roundles were placed prior to the assessment according to the user manual. On the AP images, the complete upper lip and lower border of the nose were traced; on the base view, both nasal and nostril outline were traced. A vertical axis was created bisecting a line joining the medial canthi. By reflecting the left side over the right side, the total area where the left and right sides did not overlap (percentage mismatch), measured in pixels as a percentage of
the traced area of the upper lip, was calculated by the program. Perfect symmetry would result in 0% mismatch.

**Thin Lip Compensation and Reference Scale**

The traced images (4 images per case) were imported in turn, correlated by superimposing the canthal roundels, and an average percentage mismatch was calculated. The lip perimeter mismatch percentage was calculated with and without the TLC to show the difference between both measurements. Reference ranges for nose perimeter mismatch percentage and lip perimeter mismatch percentage (with TLC) were subsequently constructed from these data.

**Statistical Analysis**

Data were analyzed using SPSS incl. PASW statistics 24.0. Intra- and interobserver agreement was tested with the intraclass correlation coefficient, using the absolute agreement for a 2-way random model. Statistical differences in lip perimeter mismatch percentages between the calculations with and without TLC were tested with the Wilcoxon signed-rank test. Statistical differences in gender for perimeter mismatch of all perimeters were studied using the Student $t$ test for normally distributed data and Mann-Whitney $U$ test for the remaining continuous data. Normal distribution was tested using the Shapiro-Wilk test. Significance was set at $P < .05$.

**Results**

**Inter- and Intra-Assessor Reliability**

In Table 1A, the intrarater reliability is given for all assessors on every perimeter constructed from SymNose. Rater 1 scored the most consistent, whereas rater 2 scored the least consistent. The nostrils were found the most difficult to score by rater 1 and rater 3, while rater 2 showed more difficulties with the lip.

Table 1B shows an interrater score of 0.80 and 0.78 between rater 1 and 2 and between rater 1 and 3, respectively, on the first assessment, and an intrarrater reliability of 0.81 and 0.83 between rater 1 and 2 and between rater 1 and 3, respectively, on the second assessment.

**Thin Lip Compensation**

The lip perimeter mismatch percentage without TLC ranged from 6.53 to 60.66 (median: 19.97), and the lip perimeter mismatch percentage with TLC ranged from 2.13 to 15.44 (median: 4.9). The perimeter mismatch percentage without TLC was significantly higher compared to the perimeter mismatch percentage with TLC ($P < .001$). In Figure 1, examples are shown of 2 noncleft individuals and 1 patient with cleft, whose perimeter mismatch percentages were calculated with and without TLC. This figure shows that without TLC, participant A with a thin lip has a disproportional higher mismatch percentage compared to participant B with a full lip (63.44% vs 15.07%). It also shows that without TLC, participant A has an even higher percentage mismatch compared to participant C (63.44% vs 49.35%), whereas participant C shows a poor result after cleft surgery and has a subjective higher degree of asymmetry compared to the noncleft individual.

**Reference Range**

Table 2 demonstrates the reference range for the noncleft comparison group in total and when subdivided into males and females. Female children had a significantly higher perimeter mismatch percentage for the lip and the nose base compared to male participants ($P = .005$ and $P = .048$, respectively).

**Discussion**

SymNose is a useful tool in the assessment of facial aesthetics after cleft surgery. This article introduces the TLC. The asymmetry of the lip as calculated with SymNose is the result of the nonoverlap between the left and right side as a percentage of the total upper lip volume measured in pixels. However, the same amount of nonoverlap will understandably result in a
higher percentage mismatch in a very thin lip, compared to the percentage mismatch of a thicker upper lip. As an attempt to limit the inevitable mismatch of thinner lips, a linear denominator was programmed using an aspect ratio. In SymNose version 6.3 and later, the TLC is automatically on but can be switched off in the analytical mode if preferred or for research purposes.

In previous studies with SymNose, normal controls were used to compare the symmetry results after cleft lip and palate repair. One of the aims of cleft surgery is to make the child look "normal," minimizing the stigmatic features patients with cleft have, to improve self-esteem and decrease psychosocial issues. However, as the phenotype of children varies, it is important to define "normal." With their crowd sourcing paper, Tse et al. (2016) recently showed that both lay people and cleft professionals rank-ordered the noncleft controls as the best aesthetic outcome in a mixed group of 46 patients with a unilateral cleft lip and palate (UCLP) and 4 normal controls. As a start to objectify "normal," this study provides the reference scale of 91 noncleft children against which to compare results after cleft surgery at 5 years of age.

In this study, a substantial to almost perfect inter- and intraobserver agreement was found. The nose on AP views was scored most reliable, whereas the nostrils and the lip for rater 2 were assessed less reliable. Mosmuller et al. (2016) found the lowest reliability on the nose base perimeter due to different perceptions on the shape of the nose. Although SymNose is proven to be reliable, there is some subjectivity to the recording of the nose and lip perimeter. Rater 2 scored a very poor intraobserver reliability on the lip (0.18). The explanation for this reliability lies in the perception of the shape of the lip. Choosing the commissure point is frequently an issue of contention between different tracers because the outer corner of the lip is often darkened by shadow and depth. This of course is even worse in patients with cleft as the upper lip vermilion is often inverted. In Figure 2A and B, an example of tracings of the same participant by rater 2 on assessment moment 1 and 2 is shown. In SymNose, photographs are rotated based on a horizontal line between the medial canthi. The roundels, that is, landmarks, must be placed by the assessor, which makes establishing the plane of rotation sensitive for errors. However, as the medial canthi are close to the midline, we believe that possible errors are negligible. Unlike most programs, the midline of the nose and lip is not established based on the medial canthi or orbits, but by an automatic measurement between the outer corners of the lips, at the commissure points, divided in half. The reason for this is preventing bias when the midline of the nose and/or lip does not coincide with the midline based on the medial canthi. However, when a rater chooses a corner point closer to the midline, the midline will automatically shift toward the other side. When subsequently the left side of the midline is reflected over the right side to calculate the nonoverlapping area, a different percentage mismatch is calculated. During a panel meeting, the authors agreed to perform tracings to the outer corners of the lip.

![Figure 1](image.png)

**Figure 1.** Examples of perimeter mismatch percentage with and without thin lip correction (TLC). A, A noncleft child with thin lips. The mismatch percentage with TLC 9.25% and without TLC 63.44%. B, A noncleft child with full lips. The mismatch percentage with TLC 5.53% and without TLC 15.07%. C, A patient with cleft having an average lip volume. The mismatch percentage with TLC 13.16% and without TLC 49.35%.

| Parameter       | Total Group | Reference Range | Median (IQR) | Reference Range | Median (IQR) | Reference Range | Median (IQR) | Reference Range |
|-----------------|-------------|-----------------|--------------|-----------------|--------------|-----------------|--------------|-----------------|
| Nose (AP)       | 14.06 (9.26) | 2.65-30.91      | 14.98 (9.18) | 5.81-30.26      | 12.85 (6.35) | 2.65-30.91      |
| Lip             | 4.9 (2.39)   | 2.13-15.44      | 4.57 (2.23)  | 2.13-11.03      | 5.53 (2.55)* | 2.59-15.44      |
| Nose (base)     | 4.75 (3.57)  | 1.69-14.84      | 4.16 (3.05)  | 1.69-10.85      | 5.17 (3.66)* | 2.36-14.84      |
| Nostril         | 11.63 (5.98) | 4.68-26.6       | 11.5 (6.5)   | 4.74-26.6       | 11.92 (4.84) | 4.68-21.34      |
| W/H ratio       | 1.46 (0.14)  | 1.15-1.80       | 1.46 (0.15)  | 1.17-1.80       | 1.47 (0.18)  | 1.15-1.72       |

Abbreviations: AP, anteroposterior; IQR, interquartile range.

*P < .05.
part of the overall aesthetic assessment, using a 3-point or 5-point Likert scale as proposed by Asher-McDade et al (Paiva et al., 2014; Pausch et al., 2016). Deall et al. (2016) found no significant relation between the aesthetic assessment of the nose by human raters and the asymmetry assessed with SymNose. Freeman et al. (2013) also found that the asymmetry results with SymNose for nose front perimeter (AP view) were in contrast with the perimeters as measured on the base view images, illuminating that aesthetic assessment of the nose is more complex, possibly because more features such as the shape of the alar base, the shape of the nostrils, septal deviation, and the width height ratio should be taken into account. It will be interesting in future research to compare SymNose asymmetry results on base view images with the subjective assessment by humans on this particular view.

The 5-year-old children in this study came from a single geographical region. The Index of Multiple Deprivation Decile divides all areas, ranked from 1 (most deprived) to 32.884 (least deprived), in 10 equal groups, showing which areas are among the most deprived 10% (first decile) or least deprived 10% (10th decile). The Index of Multiple Deprivation Decile in this region varied from 3 to 10, representing almost all economical classes, except for the first and second decile. These data are important as it is generally accepted that a relation exists between attractiveness and wealth (Gilmore et al., 1986; Hamermesh and Biddle, 1994; Jackson et al., 1995; Duarte et al., 2012; Pareek and Zuckerman, 2013; Ravina, 2012). Having a large variability in this study means that these data are applicable on a larger scale rather than solely on the “most attractive” or “least attractive” people.

The limitation of this study is that the reference scale was drawn from a Caucasian population aged 5. Therefore, caution is recommended with the interpretation of results when comparing patients with cleft in other age categories or from different ethnicities. In the future, it will be interesting to expand the normal control group to make this reference range applicable on a broader scale. Age categories 16 to 20 might be suitable as this period marks the end of treatment if no additional surgeries such as rhinoplasty or orthognathic surgery are chosen. Because of increasing migration and in order to use the normal asymmetry reference scale worldwide, asymmetry reference ranges for different noneleft ethnicities must be studied and, if divergent from the Caucasian population, be added to the reference scale.

Conclusion

This study provides a nonleft reference range for all perimeters drawn from SymNose against which to compare results after cleft surgery at 5 years of age. Although SymNose is proven to be reliable, there is some subjectivity to the recording of the nose and lip perimeter. The authors recommend performing tracings to the outer corners of the lip, indicated by the outer corners of the lower lip, and to cutoff 10% of the outer corners of the lip when calculations are performed. In newer versions of SymNose, this will be incorporated in standard
settings. Furthermore, this study shows the importance of correcting for a variance in lip volume per child.

**Declaration of Conflicting Interests**
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**
The author(s) received no financial support for the research, authorship, and/or publication of this article.

**References**
Al-Ghatam R, Jones TE, Ireland AJ, Atack NE, Chawla O, Deacon S, Albery L, Cobb AR, Cadogan J, Leary S, et al. Structural outcomes in the cleft care UK study. Part 2: dento-facial outcomes. *Orthod Craniofac Res*. 2015;18(suppl 2):14-24.
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(4):385-390; discussion 390-391.
Bella H, Kornmann NSS, Hardwicke JT, Wallis KL, Wearn C, Su TL, Richard BM. Facial aesthetic outcome analysis in unilateral cleft lip and palate surgery using web-based extended panel assessment. *J Plast Reconstr Aesthet Surg*. 2016;69(11):1537-1543.
Deall CE, Kornmann NSS, Bella H, Wallis KL, Hardwicke JT, Su TL, Richard BM. Facial aesthetic outcomes of cleft surgery: assessment of discrete lip and nose images compared with digital symmetry analysis. *Plast Reconstr Surg*. 2016;138(4):855-862.
Di Biase D, Markus A. Cleft lip and palate care in the UK: the CSAG report. *Br Dent J*. 1998;185(7):320-321.
Duarte J, Siegel S, Young L. Trust and credit: the role of appearance in peer-to-peer lending. *Rev Financ Stud*. 2012;25(8):2455-2484.
Freeman AK, Mercer NS, Roberts LM. Nasal asymmetry in unilateral cleft lip and palate. *J Plast Reconstr Aesthet Surg*. 2013;66(4):506-512.
Gilmore DC, Beehr TA, Love KG. Effects of applicant sex, applicant physical attractiveness, type of rater and type of job on interview decisions. *J Occup Psychol*. 1986;59(2):103-109.
Hamermesh DS, Biddle JE. Beauty and the labor market. *Am Econ Rev*. 1994;84(5):1174-1194.
Jackson LA, Hunter JE, Hodge CN. Physical attractiveness and intellectual competence: a meta-analytic review. *Soc Psychol Quartely*. 1995;58(2):108-122.
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 Craniofac Surg*. 2009;20(suppl 2):1683-1686.
McKeaney RM, Williams JV, Mercer NS. Quantitative computer-based assessment of lip symmetry following cleft lip repair. *Cleft Palate Craniofac J*. 2013;50(2):138-143.
Mercado AM, Russell KA, Daskalogiannakis J, Hathaway RR, Semb G, Ozawa T, Smith A, Lin AY, Long RE Jr. The Americleft project: a proposed expanded nasolabial appearance yardstick for 5- to 7-year-old patients with complete unilateral cleft lip and palate (CUCLP). *Cleft Palate Craniofac J*. 2015;53(1):30-37.
Ministry of Housing, Communities and Local Government. Index of multiple deprivation. Available at: www.opendatacommunities.org. Published, September 2015.
Mosmuller DG, Bijnen CL, Don Griot JP, Kramer GJ, Disse MA, Prahl C, Kuik DJ, Niessen FB. Comparison of two scoring systems in the assessment of nasolabial appearance in cleft lip and palate patients. *J Craniofac Surg*. 2016;44(10):1515-1521.
Mosmuller D, Maa T, Prahl C, Tan R, Mulder F, Schwirz R, de Vet H, Bergé S, Don Griot J. Comparison of two- and three-dimensional assessment methods of nasolabial appearance in cleft lip and palate patients: do the assessment methods measure the same outcome? *J Craniofac Surg*. 2017;45(8):1220-1226.
Mosmuller D, Tan R, Mulder F, Bachour Y, de Vet H, Don Griot J. The use and reliability of SymNose for quantitative measurement of the nose and lip in unilateral cleft lip and palate patients. *J Craniofac Surg*. 2016;44(10):1515-1521.
Pigott RW, Pigott BB. Quantitative measurement of symmetry from photographs following surgery for unilateral cleft lip and palate. *Cleft Palate Craniofac J*. 2010;47(4):363-367.
Pigott RW, Pigott BB. Quantifying asymmetry and scar quality of children with repaired cleft lip and palate using SymNose 2. *Cleft Palate Craniofac J*. 2016;53(3):298-301.
Persson M, Sandy JR, Waylen A, Wills AK, Al-Ghatam R, Ireland AJ, Hall AJ, Hollingworth W, Jones T, Peters TJ, et al. A cross-sectional survey of 5-year-old children with non-syndromic unilateral cleft lip and palate: the cleft care UK study. Part 1: background and methodology. *Orthod Craniofac Res*. 2015;18(suppl 2):1-13.
Ravina E. *Am Econ Rev*. 1986;74(5):1174-1194.
Russell JH, Kiddy HC, Mercer NS. The use of SymNose for quantitative assessment of lip symmetry following repair of complete bilateral cleft lip and palate. *J Craniofac Surg*. 2014;45(2):454-459.
Sell D, Mildinhall S, Albery L, Wills AK, Sandy JR, Ness AR. The Cleft Care UK Study. Part 4: perceptual speech outcomes. *Orthod Craniofac Res*. 2015;18(suppl 2):36-46.
Sharma VP, Bella H, Cadier MM, Pigott RW, Goodacre TE, Richard BM. Outcomes in facial aesthetics in cleft lip and palate surgery: a systematic review. *J Plast Reconstr Aesthet Surg*. 2012;65(9):1233-1245.

Springer IN, Wannicke B, Warnke PH, Zernial O, Wiltfang J, Russo PA, Terheyden H, Reinhardt A, Wolfart S. Facial attractiveness: visual impact of symmetry increases significantly towards the mid-line. *Ann Plast Surg*. 2007;59(2):156-162.

Tse RW, Oh E, Gruss JS, Hopper RA, Birgfeld CB. Crowdsourcing as a novel method to evaluate aesthetic outcomes of treatment for unilateral cleft lip. *Plast Reconstr Surg*. 2016;138(4):864-874.