Measuring Symmetry in Children With Cleft Lip. Part 2: Quantification of Nasolabial Symmetry Before and After Cleft Lip Repair

Jia Wu, Ph.D., Shu Liang, B.Sc., Linda Shapiro, Ph.D., Raymond Tse, M.D.

Objective: The first part of this study validated an automated computer-based method of identifying the three-dimensional midfacial plane in children with unrepaired cleft lip. The purpose of this second part is to develop computer-based methods to quantify symmetry and to determine the correlation of these measures to clinical expectations.

Participants: A total of 35 infants with unrepaired unilateral cleft lip and 14 infant controls.

Interventions: Six computer-based methods of quantifying symmetry were developed and applied to the three-dimensional images of infants with unilateral cleft lip before and after cleft lip repair and to those of controls.

Main Outcome Measure: Symmetry scores for cleft type, changes with surgery, and individual subjects ranked according to cleft severity were assessed.

Results: Significant differences in symmetry scores were found between cleft types and found before and after surgery. Symmetry scores for infants with unilateral cleft lip approached those of controls after surgery, and there was a strong correlation with ranked cleft severity.

Conclusions: Our computer-based three-dimensional analysis of nasolabial symmetry correlated with clinical expectations. Automated processing made measurement convenient. Use of these measures may help to objectively measure cleft severity and treatment outcome.

KEY WORDS: anthropometry, cleft lip, infant, symmetry, 3D stereophotogrammetry

Cleft lip (CL) occurs in approximately 1 in 1000 newborns and results in a complex three-dimensional (3D) deformity of lip and nose. Treatment occurs over multiple stages, and although changes can be dramatic, methods of objective assessment of preoperative severity and postoperative outcome are limited. In turn, the lack of objective measures hampers progress in determining the best strategies for longitudinal treatment. Providers continue to debate the best care, and decisions are often based upon experience rather than robust evidence.

Our eventual goal is to develop valid and practical ways to measure cleft severity and treatment success in infants undergoing primary CL repair. Conventional photography is prone to error from parallax (Oh et al., 2011), and direct anthropometry and dentofacial casts are impractical for routine use. Recent advances in 3D stereophotogrammetry allow rapid, accurate, and practical capture of 3D facial form in children with CL (Hood et al., 2004; Krimmel et al., 2006; Tse et al., 2014). However, current methods of analysis are labor-intensive and cumbersome because they rely upon manual placement of landmarks on a 3D mesh (Yamada, 2003; Hood et al., 2004; Singh et al., 2005; Bilwatsch et al., 2006; Krimmel et al., 2006; Nkenke et al., 2006; Singh et al., 2007; Schwenzer-Zimmerer et al., 2008; Stauber et al., 2008; Hoefer et al., 2010; Ayoub et al., 2011a, 2011b; Oh et al., 2011; Simanca et al., 2011; Bugaighis, Mattick, et al., 2013, 2014; Li et al., 2013; Bell et al., 2014; Bugaighis, Tiddeman, et al., 2014; Othman et al., 2014; Tse et al., 2014; Tse and Lien, 2015). Although a few methods of analysis that do not rely on indirect anthropometry have been described (Ferrario et al., 2003; Nkenke et al., 2003; van Loon et al., 2010; Dixon et al., 2013; Djordjevic et al., 2014), none of these have been applied to infants and children with unrepaired CL.

We have chosen to involve automated computer-based methods so that analysis can be convenient to permit widespread use. Our initial goal is to quantify nasolabial symmetry, given that symmetry is a major component of facial attractiveness [insert reference] and goal of cleft treatment.

The first part of this project involved the assessment of several methods to establish a vertical midfacial reference
plane for infants with unrepaired clefts (Wu et al., 2015) so that symmetry could be quantified. The purpose of this subsequent project was to develop automated quantitative measures of 3D symmetry and to apply these analyses to infants with unilateral cleft (UCL) before and after primary CL repair.

We specifically wanted to determine whether application of these measurements detected statistically significant differences in (1) subjects with UCL versus infants with no craniofacial differences; (2) subjects with UCL grouped according to cleft type; and (3) subjects with UCL before and after primary CL repair.

We also wanted to determine how well these measurements correlated with the preoperative rank order of cleft severity for individual subjects with UCL.

Given our objective of automation, 3D image processing (cropping and pose-normalization) and analysis (reference plane placement and symmetry measurement) were performed using computer-based processes.

**METHODS**

This prospective study was approved by the Seattle Children’s Hospital Institutional Review Board and informed consent to participate was obtained for each subject.

**Subjects**

Families of consecutive patients presenting to the surgeon’s clinic for primary UCL repair were approached to participate. There were no exclusion criteria and subjects were included regardless of any associated syndrome of a craniofacial condition. Four families declined to participate and four families left the clinic before consent could be obtained. Consent could not be acquired from six families due to language barriers. A total of 35 subjects gave consent and were included in this study. The boy to girl ratio was 19:15, laterality of cleft (left to right) was 18:17, and three subjects had associated anomalies (craniofacial microsomia, popliteal pterygium, and a chromosomal anomaly of uncertain significance).

Cleft extent and type were documented as part of standard clinical care. Eight subjects had CL, 10 had cleft lip and alveolus (CLA), and 17 had cleft lip and palate (CLP). A complete cleft lip with band (i.e., “Simonart”) was differentiated from an incomplete CL by the presence of a complete alveolus (Semb and Shaw, 1991). The type of CL included 17 incomplete UCL, eight complete UCL with band, and 10 complete UCL.

Patients were treated according to standard clinical protocol. Presurgical molding involved lip taping for infants with complete CL and involved nasoalveolar molding for infants with complete CLP. Of the 10 infants with complete UCL, two had CLA and eight had CLP. Neither of the two subjects with CLA underwent taping because one did not tolerate it and the other was an older child adopted from abroad. Of the eight subjects with CLP, five underwent nasoalveolar molding; whereas, three had no molding (two did not tolerate taping and one had other comorbidities that precluded molding). Cleft lip repair was planned at 6 months of age unless an infant was undergoing nasoalveolar molding (in which case surgery was planned once molding was completed). When patients presented at an older age, as in the case of international adoption, surgery was planned as soon as possible. Mean age at surgery was 6.6 months (range, 3 to 24 months).

All subjects with UCL underwent cleft lip repair using a previously described approach (Tse, 2013; Tse and Lien, 2015) that includes skin markings according to the anatomical subunit approximation (Fisher, 2005; Tse and Lien, 2015), an emphasis on wide surgical release, nasal floor and sidewall reconstruction, primary septoplasty (Anderl et al., 2008; Gosla-Reddy et al., 2011; Ridgway et al., 2011), and no tip dissection. We collected preoperative and postoperative images in the clinic according to our center’s protocol using the 3dMD Cranial System (3dMD, Atlanta, GA).

In addition to subjects with UCL, we collected 3D images of 14 age-matched normal control infants who had no craniofacial differences for comparison (mean age, 7.1 months; range, 4 to 10 months). Boy to girl ratio in this group was 10:4.

**Image Processing**

In order to mimic routine clinic conditions, all patients were included regardless of captured image quality. We applied a previously described computer-based algorithm that automatically removes structures other than the face and corrects alignment into a standard facial frontal plane (Wu et al., 2014). No manual image processing was performed.

**Rank Order of Cleft Severity**

Subjects with UCL were ranked in order of decreasing cleft lip nasal deformity by the cleft surgeon (R.T.). We created a digital “sorting board” in which mesh images could be sorted. Each mesh image was enlarged and rotated synchronously with the adjacent ordered images so that 3D form could be examined and compared.

**Quantification of Symmetry**

We defined the midfacial reference plane as the geometric midline of the face that ignores the nasolabial region. This plane was identified using a previously described automated computer algorithm (the deform-
tion method; Liang et al., 2013) that calculates the midline of the eye and chin points (Wu et al., 2015). We used this plane as the $x = 0$ plane.

We quantified symmetry by measuring nasolabial surface differences across the midline (point, radius, or angle) and by measuring deviation of the nose from midline (columellar deviation, nasal tip position, and nasal axis deviation). We defined the nasolabial region as that bounded by the interchelion width and extending superiorly to the interendocanthion line (Fig. 1A). For measurements that are based upon specific landmarks, the landmarks were identified by computer-based template mesh deformation as previously described (Liang et al., 2013).

For all measures, higher values reflected greater asymmetry and lower values reflected greater symmetry.

**Nasolabial Symmetry—Point Difference**

For every point $p$ with coordinate $(x, y, z)$, the symmetric point with coordinate $(-x, y, z)$ was labeled $p_s$. With $q$ being the point on the facial mesh closest to $p_s$, the overall point difference (PD) across the midline was calculated as

$$PD = \frac{1}{n} \sum_{p \text{ in the area}} \text{distance}(p_s, q),$$

where $n$ is the number of points in the nasolabial region.
and distance is the Euclidean distance between points (Fig. 1A).

**Nasolabial Symmetry—Radius Difference**

The nasolabial region was divided into grid patches (Fig. 1B), with half of the area divided into \( m \times m \) squares that were equally divided in the \( z \) and \( \theta \) directions.

The radius difference (RD) defined for a grid patch at position \((\theta, z)\) is

\[
RD(\theta, z) = r(\theta, z) - r(-\theta, z),
\]

where \( r \) takes the average radius value in a given grid patch, and \((-\theta, z)\) is the reflected grid patch of \((\theta, z)\) across the midline.

The overall RD is the average of all the grids on one side of the face:

\[
AD = \frac{1}{m \times m} \sum_{(\theta, z)\text{in area}} AD(\theta, z).
\]

**Nasolabial Symmetry—Angle Difference**

Similar to RD, the nasolabial region was divided into grids with the angle difference (AD) for a grid patch \((\theta, z)\) defined as

\[
AD(\theta, z) = \cos(\beta_{r(\theta, z), r(-\theta, z)}),
\]

where \( \beta_{r(\theta, z), r(-\theta, z)} \) is the angle between the surface normal vector of the mesh at grid patch \((\theta, z)\) and its reflected grid patch. This measures how different the two patches are oriented. The overall AD is the average of all the grids on one side of the face:

\[
AD = \frac{1}{m \times m} \sum_{(\theta, z)\text{in area}} AD(\theta, z).
\]

**Statistical Analysis**

Differences in measurements between groups were assessed by \( t \) test with \( P < .05 \) considered significant. The correlation of measurements and rank order of severity was assessed using a Pearson correlation coefficient.

**RESULTS**

**Nasolabial Symmetry Before and After Primary Cleft Lip Repair**

Mean nasolabial symmetry measures for subjects with UCL before and after surgery are summarized in Table 1. The measures changed by 48%, 33%, and 61% for PD, RD, and AD, respectively. All of these changes
were statistically significant. Similarly, the measures changed by 54%, 58%, and 44% for columellar deviation, nasal tip deviation, and nasal axis deviation, respectively. All of these changes were also statistically significant.

Relative to controls, subjects with UCL had significantly worse symmetry scores before surgery (Table 1). There was no difference in symmetry measures between controls and subjects with UCL after cleft lip repair (Table 1).

**Comparisons of Cleft Types**

Mean symmetry measures for subjects grouped according to cleft type, prior to cleft lip repair, are contained in Table 2. Nasolabial symmetry measures for point, radius, and angle differences followed clinical expectations. Symmetry measures for complete UCL were similar to complete with band UCL. Symmetry measures for incomplete UCL were significantly better than those for both complete UCL and complete with band UCL. In turn, symmetry measures for controls were better than those for incomplete UCL. These differences were statistically significant for all measures except for PDs.

Measurement of columellar deviation, nasal tip deviation, and nasal axis deviation followed similar trends. Measurements for complete versus complete with band were similar. Measurements for incomplete were significantly better than for complete UCL for columellar deviation and nasal tip deviation and were significantly better than both complete UCL and complete with band UCL for nasal axis deviation. Symmetry scores for controls were significantly better than those for all type of UCL.

Mean measurements of nasolabial symmetry and nasal deviation for subjects grouped according to cleft type, after surgery, are contained in Table 2. Differences in symmetry scores between subject groups were less common and inconsistent. Differences in nasal tip deviation and nasal axis deviation were found between subject groups.

**Correlation to Ranking**

Measures of nasolabial symmetry and nasal deviation followed expected trends in which more severe clefts had greater preoperative scores, and less severe clefts had lower preoperative scores that approached those of normal controls (Fig. 1C). Preoperative and postoperative measurements of each subject in rank order of preoperative cleft severity are displayed in Figs. 3 and 4. The measured preoperative nasolabial symmetry and nasal deviation was closely associated with the preoperative rank order of severity. Correlation coefficients were .72, .71, and .7 for PD, RD, and AD, respectively. The correlation coefficients were .45, .76, and .72 for

---

### TABLE 1 Measures of Symmetry (Standard Deviation) for Subjects With Unilateral Cleft Lip (UCL) Before and After Cleft Lip Repair and for Controls

<table>
<thead>
<tr>
<th>Subject</th>
<th>Nasolabial Surface Differences</th>
<th>Nasal Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Point</td>
<td>Radius</td>
</tr>
<tr>
<td>UCL preoperative</td>
<td>4.33</td>
<td>2.04</td>
</tr>
<tr>
<td>UCL postoperative</td>
<td>1.67</td>
<td>1.07</td>
</tr>
<tr>
<td>Control</td>
<td>1.89</td>
<td>0.91</td>
</tr>
</tbody>
</table>

* Significant difference compared with postoperative and control (t test, P < .05).

---

### TABLE 2 Nasolabial Symmetry Measures by Type Before and After Cleft Lip Repair

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Nasolabial Surface Differences</th>
<th>Nasal Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Point</td>
<td>Radius</td>
</tr>
<tr>
<td>Before surgery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complete</td>
<td>6.19</td>
<td>2.58</td>
</tr>
<tr>
<td>Complete + band</td>
<td>6.26</td>
<td>2.65</td>
</tr>
<tr>
<td>Incomplete</td>
<td>2.32</td>
<td>1.44‡</td>
</tr>
<tr>
<td>Control</td>
<td>1.88‡</td>
<td>0.91§</td>
</tr>
<tr>
<td>After surgery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complete</td>
<td>1.37</td>
<td>1.01</td>
</tr>
<tr>
<td>Complete + band</td>
<td>1.51</td>
<td>0.99</td>
</tr>
<tr>
<td>Incomplete</td>
<td>1.92</td>
<td>1.14</td>
</tr>
<tr>
<td>Control</td>
<td>1.88</td>
<td>0.91</td>
</tr>
</tbody>
</table>

* Significant difference compared with incomplete (t test, P < .05).
† Significant difference compared with complete (t test, P < .05).
‡ Significant difference compared with complete and with complete + band (t test, P < .05)
§ Significant difference compared with complete, complete + band, and incomplete (t test, P < .05)
FIGURE 3 Symmetry scores quantifying A: point, B: radius, and C: angle differences across the midfacial reference plane. Subjects with UCL are arranged according to their ranked severity of cleft lip nasal deformity (x-axis). Symmetry scores, before and after surgery, are displayed. Symmetry scores for control subjects are also displayed on the far right. Symmetry scores for subjects with UCL approached those for control subjects following cleft lip repair regardless of initial severity.

FIGURE 4 Symmetry scores quantifying A: columellar deviation, B: nasal tip deviation, and C: nasal axis deviation. Subjects with UCL are arranged according to their ranked severity of cleft lip nasal deformity (x-axis). Symmetry scores, before and after surgery, are displayed. Symmetry scores for control subjects are also displayed on the far right. Symmetry scores for subjects with UCL approached those for control subjects following cleft lip repair regardless of initial severity.
columellar deviation, nasal tip deviation, and nasal axis deviation, respectively.

As expected, postoperative measures of nasolabial symmetry and nasal deviation approached those of controls for most subjects (Figs. 3 and 4). The correlation of postoperative nasolabial symmetry and nasal deviation measurements with the preoperative rank order of severity was greatly reduced. Correlation coefficients were .19, .27, and .02 for PD, RD, and AD, respectively. Correlation coefficients were .13, .35, and .04 for columellar deviation, nasal tip deviation, and nasal axis deviation, respectively.

As expected, there was a close correlation of preoperative severity with the magnitude of change in measurements before and after surgery. Correlation coefficients were .70, .71, and .61 for PD, RD, and AD, respectively. Correlation coefficients were .3, .76, and .64 for columnellar deviation, nasal tip deviation, and nasal axis deviation, respectively.

Discussion

Our understanding of the relative effect of various techniques, treatments, and protocols on aesthetic cleft outcomes is limited by a lack of accurate, objective, and practical means to assess changes through treatment and time. For example, the effect and longevity of primary septoplasty, primary tip rhinoplasty, “school-age” tip rhinoplasty, and septrhinoplasty is poorly documented. The paucity of data and inability to compare studies results in protocols that are mostly based upon expert opinion rather than scientific evidence. In turn, optimization of a given treatment protocol to achieve ideal outcomes, with few interventions through a given patient’s childhood, is difficult to determine.

Three-dimensional stereophotogrammetry allows convenient capture of form and offers an opportunity for objective analysis. However, methods to measure cleft deformity and outcome have yet to be established. The validity of indirect anthropometric measurement on 3D images has been studied (Tse et al., 2014), but these measurements require manual placement of landmarks that is tedious and labor intensive. In addition, the point-to-point measurements are a surrogate of actual form and do not take advantage of the abundant data available in a given 3D facial surface mesh. Automated analysis would allow adoption and more widespread use, and consideration of the full surface data would provide a better indication of form.

Djordjevic et al. (2014) reported a pilot study using landmark-independent analysis of 5-year-olds with repaired UCL with or without cleft palate. They found no measureable difference in symmetry when compared with normal controls; however, their analysis involved the full face rather than focusing on the nasolabial region. In another study, van Loon et al. (2010) performed volumetric analysis of the nose using 3D surface data before and after secondary cleft rhinoplasty and reported measurable changes in volume. Although both of these studies examined surface data, both required considerable manual processing to crop and pose-normalize subjects. Both of these studies also involved much older subjects than ours and neither considered the initial CL nasal deformity prior to primary repair. Given the need for longitudinal analysis of form over a patient’s childhood, we focused on infants prior to their initial CL repair.

We have developed a system of computer-based algorithms that automates image processing, orientation, and analysis. First, using computer vision techniques, we developed an algorithm that recognizes the face, removes nonfacial components of the 3D image, and pose-normalizes images into frontal view (Wu et al., 2014). In our next study, we compared available computer-based methods of defining a reference midfacial plane (Wu et al., 2015). In the current study, we have examined ways of quantifying nasolabial symmetry. Used together, these methods allow for automation that reduces the burden of analysis and permits assessment of large sample sizes. What would take weeks to process and analyze (Tse et al., 2014) can take just a few minutes. The convenience of computer analysis allows for routine application and may encourage wider adoption by other institutions. In turn, reports of objective measures may improve the available evidence used for clinical protocols.

All of the measures that we developed and assessed followed clinical expectations. Differences were found preoperatively versus postoperatively, between cleft types, and when compared with normal control subjects. We also found a close correlation of measurements with preoperative CL nasal severity.

Among the measures assessed, those that considered the entire nasolabial surface (PD, RD, and AD) performed better than those that relied upon specific landmarks (columellar deviation, nasal tip deviation, nasal axis deviation), as reflected by better and more consistent correlation coefficients. This is not surprising given that computer-based landmarking is still less accurate than human perception (Liang et al., 2013). Surface-based measurements that incorporate all of the available data in a given region reduce error, thereby yielding measures that approximate clinical expectations. Among these symmetry measures, the RD seemed to perform best. Among the landmark-based measures, the nasal tip deviation seemed to perform best. We propose that these measures can potentially be used as objective measures of preoperative cleft severity.

Whereas preoperative measures correlated nicely with initial severity (correlation coefficient >.7 for all but columellar deviation), postoperative measures did not.
This lack of postoperative correlation with preoperative severity may have two explanations. The first may be that all patients, regardless of preoperative severity, underwent successful repair, thereby making them difficult to distinguish. Postoperative measures approximated those of controls and greater changes in symmetry scores were recorded for subjects with greater initial cleft severity. All patients underwent surgical repair by a single surgeon and in a previous study, normalization of anthropometric measures was achieved following cleft lip repair regardless of cleft type (Tse and Lien, 2015). The second explanation is that our measures are not sensitive enough to detect more subtle differences in form and symmetry among subjects postoperatively. Although nasolabial form has been improved as expected, our measures of symmetry may not detect residual aberrations of form. Future work will focus on more specific shape-based analysis and detailed analysis of form so that postoperative results can be better differentiated.

Although we have devised a system of analysis that is objective, convenient, and seemingly valid, there are several factors that may limit accuracy. First, our subjects were very young (mean, 6 months old), and the size of their facial surfaces is small relative to the range for which 3D stereophotogrammetry systems are designed (adult full head). In addition to limited resolution, multiple shadow regions and the complex form of the un repaired cleft deformity result in artifacts and regions devoid of surface data. We previously found these to limit the accuracy of anthropometric measurement (Tse et al., 2014). With advances in technology, fine resolution will likely improve. Second, image registration was performed using facial surface data rather than bony landmarks. Our goal was to develop a system of analysis that could be incorporated into routine clinical care, and placement of landmarks would be outside of this realm. We relied, therefore, upon recognition of 3D facial features. In some cases, our method may be more appropriate than use of bony landmarks if a given patient has plagiocephaly, hemifacial microsomia, craniosynostosis, or other asymmetries that are rarely, but can be, associated with CL. Third, our analysis involved capture of a single moment in time, usually of an infant who was awake. Although our image technologist captured multiple images, most images had some degree of facial animation, which likely results in variations in measurement of symmetry. Analysis of multiple images may reduce this error; however, the additional analysis adds further burden. Newer “four-dimensional” (4D) imaging systems capture 3D video; however, these are expensive, and dynamic analysis adds an additional layer of complexity. We currently do not have access to a 4D system, and use of such imaging at this time would be outside of our routine clinical care.

**Conclusions**

Automated computer-based processes can reduce the burden of manual image processing, normalization, and analysis to allow for convenient measurement of form, even with large sample sizes.

Landmark-independent measures of nasolabial symmetry and nasal deformity that we tested could detect 3D changes from before and after surgery. Using these measurements we found significant differences between infants with clefts and normal controls and between groups of subjects divided by cleft type. Measurements of symmetry correlated well with ranked severity of preoperative CL nasal deformity and normalized to approach those of control subjects postoperatively.

Among the measures of symmetry that we studied, the RD and nasal tip deviation performed best in approximating clinical expectations.

Development of accurate and convenient means to quantify changes from treatment may help to determine better treatment techniques, strategies, and protocols.

**References**


