

# Influence of Severity on Aesthetic Outcomes of Unilateral Cleft Lip Repair in 1,823 Patients

Alex Campbell, MD\*†‡  
 Carolina Restrepo, MD\*†‡  
 Eugene Park, MD§  
 Genesis Navas, MD‡  
 Gaurav Deshpande, MDS\*¶  
 Jordan Swanson, MD, MSc†||  
 Bjorn Schonmeyr, MD, PhD\*\*  
 Lisa Wendby, BA†  
 Ruben Ayala, MD, MSc†

**Background:** Although efforts to improve access to care for patients with cleft lip in the developing world have grown tremendously, there is a dearth of data regarding aesthetic outcomes after cleft lip repairs in this setting. Defining severity-outcome relationships has the potential to improve efficiency of care delivery in resource-limited settings, and to improve overall results. In this study, we investigate the relationship between initial cleft lip severity and early aesthetic outcomes following surgical repair of primary unilateral cleft lip.

**Methods:** Using previously validated tools to assess unilateral cleft lip severity and aesthetic outcome after repair, we evaluated 1,823 consecutive patients who underwent primary unilateral cleft lip/nose (UCL/N) repair. Three separate evaluators scored each case for a total of 5,469 total independent evaluations.

**Results:** Our results show that with increasing severity of UCL/N deformity, there is a corresponding decrease in early aesthetic outcome scores. Using our results, we established normative early aesthetic outcomes following repair for each severity grade of UCL/N deformity.

**Conclusions:** In conclusion, this study has achieved a standardized, timely, and cost-effective evaluation of 1,823 surgical cases of primary UCL/N repair. This data set provides a normal distribution of aesthetic results according to initial cleft severity and defines a standard of “expected” aesthetic results after primary UCL/N repair. Our results also show a clear correlation between initial severity and immediate aesthetic result after surgery, though we also show that excellent results are possible regardless of initial cleft severity. (*Plast Reconstr Surg Glob Open* 2019;7:e2083; doi: 10.1097/GOX.0000000000002083; Published online 22 January 2019.)

## INTRODUCTION

The central element of the overall outcome after cleft lip repair is the aesthetic appearance of the lip and nose. Oral clefts are among the most common craniofacial anomalies worldwide, yet there is relatively little evidence on aesthetic results after unilateral cleft lip repair. Severity of the primary unilateral cleft lip/nose deformity

(UCL/N) is postulated to play a key role in aesthetic results after surgery, but this remains controversial in the literature with conflicting reports involving small sample sizes.<sup>1-6</sup> Defining severity-outcome relationships has the potential to improve efficiency of care delivery in resource-limited settings, and to improve overall results. Surgeons and techniques achieving the best results for each severity phenotype can be identified and analyzed to determine best practices across the cleft spectrum.

In this study, we investigate the relationship between initial cleft lip severity and early aesthetic outcomes following surgical repair. Using previously validated indices for cleft lip severity and aesthetic outcome,<sup>7,8</sup> we evaluated the relative prevalence of varying severities of cleft lip deformity, as well as the range of outcomes associated with

*From the \*Division of Plastic Surgery, Penn State Hershey Medical Center, Hershey, Pa.; †Operation Smile, Virginia Beach, Va.; ‡Premium Care Plastic Surgery, Cartagena, Colombia; §Division of Plastic Surgery, Northwestern University, Chicago, Ill.; ¶MGM Dental College and Hospital, Navi Mumbai, India; ||Division of Plastic Surgery, Childrens Hospital of Philadelphia, Philadelphia, Pa.; \*\*Division of Plastic Surgery, Skane University Hospital, Malmo, Sweden.*

*Received for publication August 13, 2018; accepted October 26, 2018.*

*Copyright © 2019 The Authors. Published by Wolters Kluwer Health, Inc. on behalf of The American Society of Plastic Surgeons. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.*

DOI: 10.1097/GOX.0000000000002083

**Disclosure:** *The authors have no financial interest to declare in relation to the content of this article. The Article Processing Charge was paid for by the authors.*

Supplemental digital content is available for this article. Clickable URL citations appear in the text.

each level of severity. To our knowledge, this is the largest study of its kind to review aesthetic results after primary UCL/N repair. One thousand eight hundred twenty-three consecutive patients undergoing primary UCL/N repair received evaluations of preoperative cleft severity and early postoperative aesthetic outcome. Three separate evaluators scored each case for a total of 5,469 total independent evaluations.

The goals of this study are:

- 1.) To examine the implementation of a large-scale evaluation of aesthetic outcomes after primary UCL/N repair, utilizing validated evaluation tools.
- 2.) To define the distribution of aesthetic results according to initial cleft severity.
- 3.) To investigate the influence of initial cleft severity on early aesthetic outcomes.

### METHODS

This retrospective analysis includes 1,823 patients with primary UCL/N deformity presenting to Operation Smile for repair. Secondary cases were excluded from this study. Eight hundred study subjects were drawn from patients treated at the Guwahati Comprehensive Cleft Care Center in Assam, India, between May 2011 and March, 2014. One thousand twenty-three study subjects were drawn from patients receiving surgery at 39 separate Operation Smile mission sites in 18 countries between March 2015 and May 2017 (Table 1).

#### Subjects

This retrospective review gathered key patient demographics, as well as immediate preoperative and postoperative photographs. Photographs were de-identified, cropped to minimize the portion of the face not affected by the cleft, and formatted to be of uniform length and width. All patients' parents or guardians signed an informed consent allowing for the use of their medical records and photographs.

#### Assessment

Preoperative cleft severity was graded utilizing the Unilateral Cleft Lip Severity Index (UCL CSI). This assessment tool is based on defined guidelines that evaluate the overall appearance of the primary deformity, and separates patients into 4 categories according to the progressive degree of lip and nose involvement (Fig. 1; see figure, **Supplemental Digital Content 1**, which display the criteria and examples demonstrating each of the 4 grades of the Cleft Severity Index, <http://links.lww.com/PRSGO/A964>). Postoperative aesthetic outcome was evaluated utilizing the Unilateral Cleft Lip Surgical Outcomes Evaluation (UCL SOE) scale. The UCL SOE scores symmetry of 4 individual anthropomorphic components of the cleft repair (Cupid's bow, lateral lip, nose, and free vermillion), and the scores of each are then summed for a total score of 0 (lowest) to 8 (highest) [Fig. 2; see figure, **Supplemental Digital Content 2**, which displays the unilateral cleft lip surgical outcomes evaluation scale scores of each element

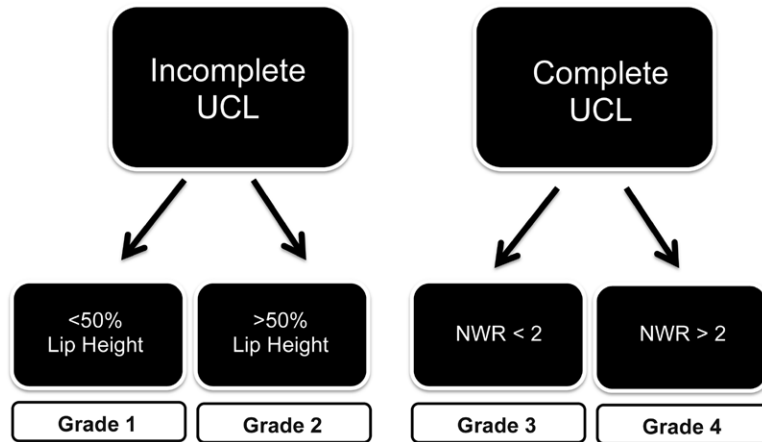
**Table 1. Locations, Dates, and Number of Patients Treated at Each Operation Smile Site Included in the Current Cohort**

Location	Date	No. Patients (Total 1,023)
Bolivia, Santa Cruz	March, 2015	19
Bolivia, Santa Cruz	October, 2015	17
Bolivia, Santa Cruz	March, 2016	27
Brazil, Mossoro	August, 2016	16
Brazil, Mossoro	January, 2017	10
Dominican Republic, Santo Doming	August, 2016	11
Democratic Republic Congo, Kinshasa	July, 2015	131
Ethiopia, Addis Ababa	June, 2015	59
Ethiopia, Mekelle	October, 2015	41
Ghana, Ho	March, 2016	2
Ghana, Ho	November, 2015	66
Guatemala, Guatemala	September, 2016	17
Guwahati, India (GCCC Center)	May 2011 - March 2013	800
Honduras, Comayagu	April, 2016	6
Madagascar, Antananarivo	April, 2015	32
Madagascar, Tamatave	September, 2015	1
Malawi, Zomba	July, 2016	75
Mexico, Guadalajara	February, 2016	15
Mexico, Monterrey	June, 2016	15
Mexico, Monterrey	October, 2015	4
Mexico, Puebla	August, 2015	7
Mexico, San Cristobal de las Casa	April, 2016	35
Morocco, Beni-Mellal	July, 2016	45
Morocco, Casablanca	April, 2016	18
Morocco, El-Jadida	November, 2015	23
Morocco, Ouarzazate	March, 2016	27
Morocco, Tiznit	July, 2015	29
Nicaragua, Esteli	Aprin, 2016	9
Nicaragua, Managua	May, 2017	27
Paraguay, Asuncion	March, 2016	17
Paraguay, Asuncion	September, 2015	10
Philippines, Bacolod	June, 2016	48
Philippines, Cavite	February, 2016	26
Philippines, Isabela	February, 2016	47
Philippines, Isabela	June, 2016	18
Philippines, Roxas City	July, 2015	15
Philippines, South Cotabato	February, 2016	35
South Africa, Nelspruit	September, 2015	15
Thailand, Sisaket	September, 2015	7
Venezuela, Maracaibo	July, 2016	2

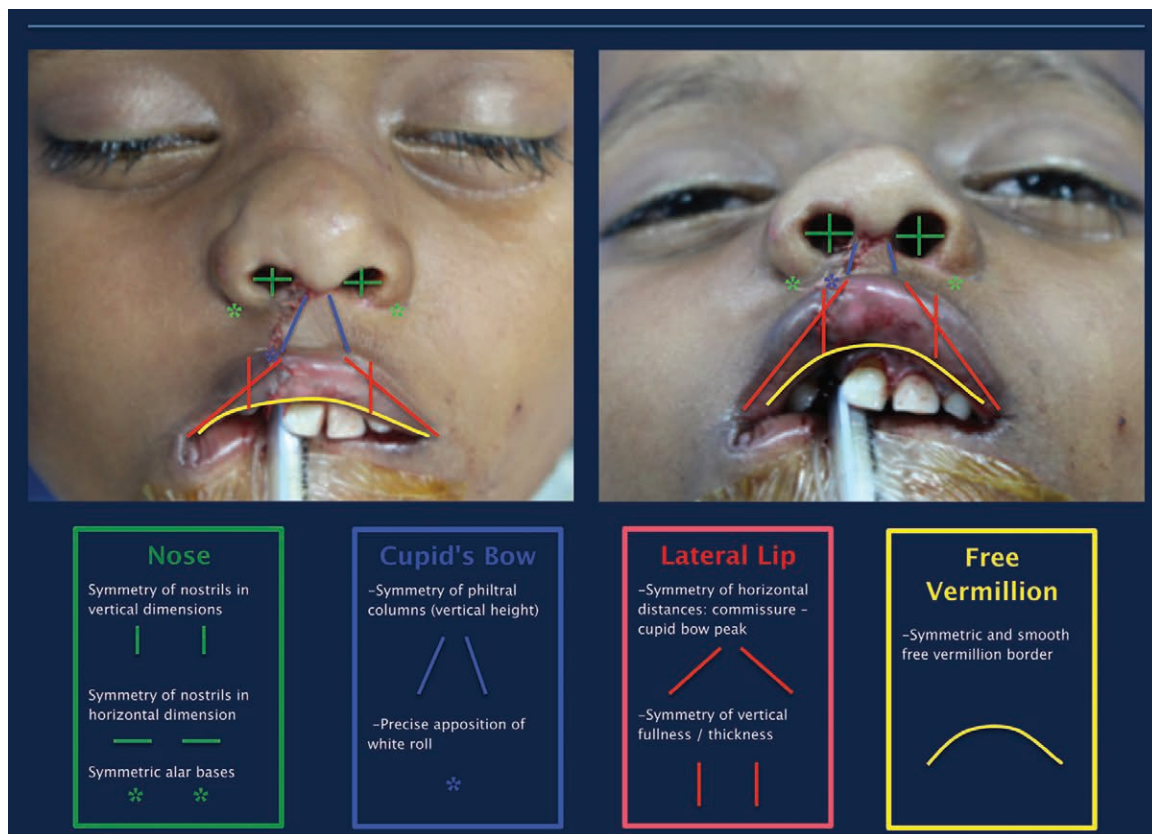
(Cupid's bow, lateral lip, nose, and free vermillion) on a 3-point scale: 2 (excellent), 1 (mild asymmetry), 0 (unsatisfactory), <http://links.lww.com/PRSGO/A965>]. Both tools have previously been validated and found to have good (>0.7) to very good (>0.8) inter-rater reliability.<sup>7,8</sup>

#### Evaluators

All cases were reviewed and graded by 3 separate lay-person evaluators, as recommended by prior validation studies. Eighteen evaluators were recruited from the institutions of the authors. Eleven evaluators were dental students at the Mahatma Gandhi Mission Dental College and Hospital in Navi Mumbai, India, and 7 evaluators were general physicians working in Cartagena, Colombia. Written and video instructions on how to use the UCL CSI and UCL SOE were provided to all evaluators, and a 2-hour teaching session was provided to describe the scale and perform practice cases. Three separate evaluators were provided with frontal and basal photo-



**Fig. 1.** The unilateral cleft lip severity index for surgeons and laypersons uses an algorithm that separates patients into 4 categories according to the progressive degree of lip and nose involvement. NWR, nostril width ratio; UCL, unilateral cleft lip.



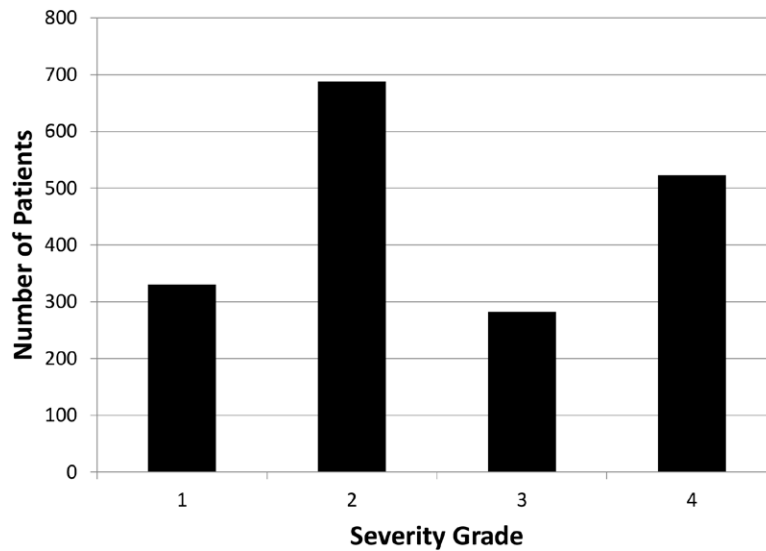
**Fig. 2.** The unilateral cleft lip surgical outcomes evaluation scale for surgeon and laypersons scores the symmetry of 4 individual anthropomorphic components of the cleft repair (Cupid's bow, lateral lip, nose, and free vermillion). The scores are then summed for a total score of 0 (lowest) to 8 (highest).

graphs before surgery and after surgery for each of the 1,823 cases. On-table, immediate preoperative photographs were utilized to grade initial cleft severity according to the UCL CSI. On-table, immediate postoperative photographs were used to grade the aesthetic result after primary UCL/N repair according to the UCL SOE. Evaluators recorded their scores of cleft severity and aesthet-

ic outcome in tables provided adjacent to the pictures, saved the evaluations, and sent results electronically back to the researchers for analysis.

**Data Analysis**

All data were tabulated using Microsoft Excel (Microsoft, Redmond, Wash.), and statistical analyses



**Fig. 3.** Distribution of unilateral cleft lip patients according to cleft severity.

were performed using SPSS 20 (IBM, Armonk, N.Y.). For cleft severity, the mode (most frequent) score of the 3 evaluations of an individual case was reported as the final cleft severity for that case. For aesthetic outcome, the average (mean) score of the 3 evaluations of an individual case was calculated and reported as the final aesthetic outcome score for that case. Cases were then separated into 4 groups according to cleft severity, and the mean and SD were calculated for each group. Homogeneity of variances was violated, as assessed by Levene’s Test of Homogeneity of Variances ( $P < 0.0005$ ). Thus, differences in means were evaluated using Welch’s analysis of variance, with Games-Howell post hoc analyses performed to analyze differences between each group.

### RESULTS

The distribution of unilateral cleft lip patients according to cleft severity is shown in Figure 3. Fifty-six percent of patients presented with incomplete clefts (grade 1 and grade 2) and 44% presented with complete clefts

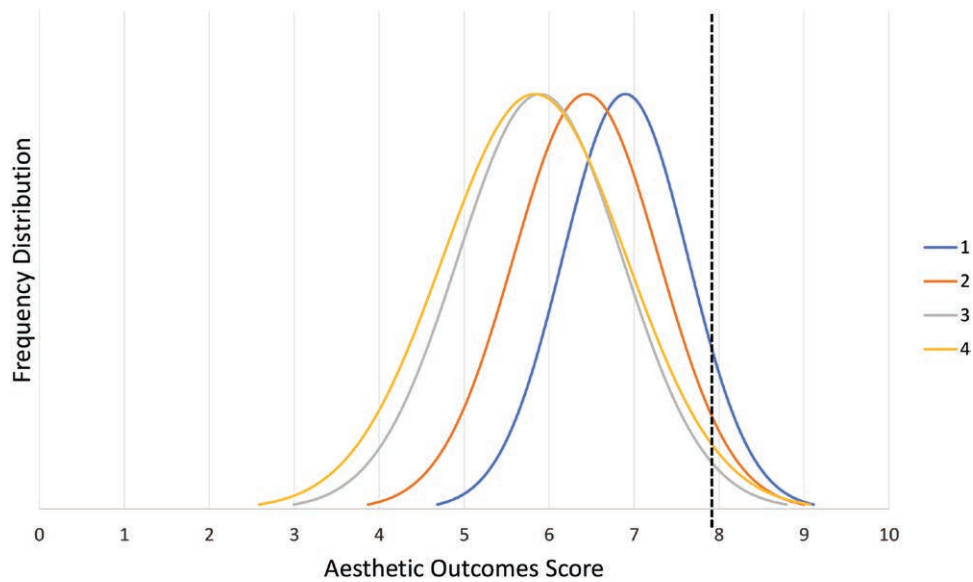
(grade 3 and grade 4). With incomplete cleft lips, the more severe phenotype (grade 2) is more than twice as common as mild incomplete cleft lips (grade 1). For complete clefts of the lip, the more severe phenotype (grade 4) likewise dominates over the less severe (grade 3), in a similar ratio.

The distribution of aesthetic outcome scores according to initial cleft severity are presented in Table 2 and Figure 4. The results demonstrate that as severity of the initial UCL/N deformity worsens, the aesthetic outcome score likewise worsens, and with a wider distribution of results ( $P < 0.0005$ ). For grades 1, 2, and 3, there was a statistically significant decrease in average outcome score with each increase in the grade of severity ( $P < 0.0005$ ). Grade 4 patients had significantly lower scores compared with grades 1 and 2 ( $P < 0.0005$ ). Scores for grade 4 patients tended to be lower than for grade 3 patients, but this did not reach statistical significance (Table 2).

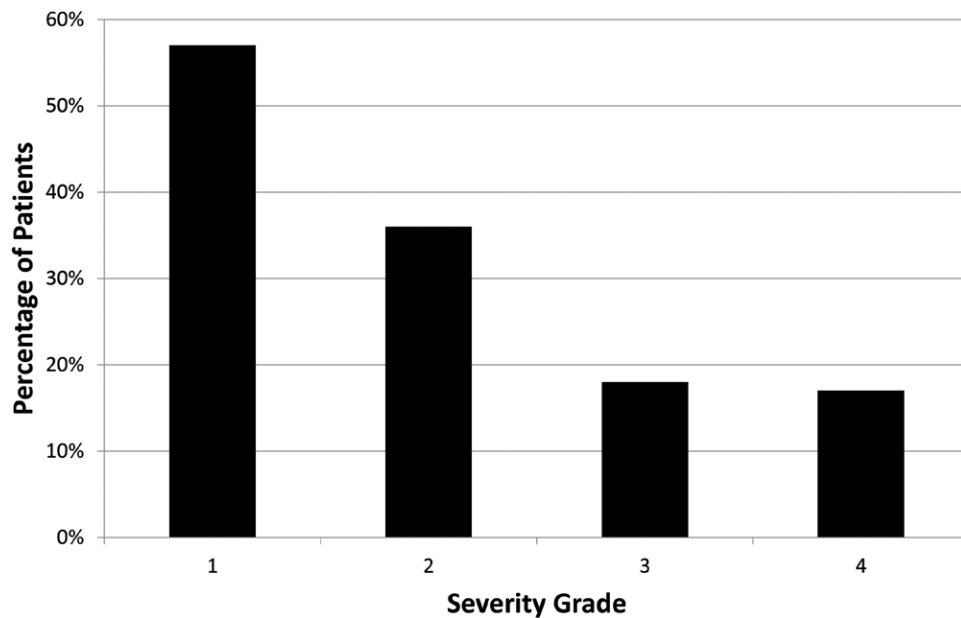
Results also demonstrate that attaining a “perfect” or “near perfect” score (total averaged aesthetic outcome score of 7.00–8.00) becomes much less frequent as cleft

**Table 2. List of Average Aesthetic Outcome Scores for Each Grade of Cleft Lip Severity and Comparison of Means across Severity Grades**

Grade	Average (Mean ± SD)	Comparison Grade	Mean Difference	P	95% CI Lower Bound	95% CI Upper Bound
1	6.90±0.74	2	0.46	<0.0005	0.33	0.60
		3	1.00	<0.0005	0.82	1.18
		4	1.07	<0.0005	0.91	1.23
2	6.44±0.86	1	-0.46	<0.0005	-0.60	-0.33
		3	0.54	<0.0005	0.37	0.71
		4	0.60	<0.0005	0.46	0.75
3	5.89±0.97	1	-1.00	<0.0005	-1.18	-0.82
		2	-0.54	<0.0005	-0.71	-0.37
		4	0.06	0.83	-0.13	0.25
4	5.83±1.08	1	-1.07	<0.0005	-1.23	-1.28
		2	-0.60	<0.0005	-0.75	-0.46
		3	-0.06	0.83	-0.25	0.13



**Fig. 4.** Visual representation of the distribution of aesthetic outcome scores for each severity grade of unilateral cleft lip (dotted line represents highest possible score).



**Fig. 5.** Frequency of perfect (8.0) and near perfect (>7.0) aesthetic outcome scores for each severity grade of unilateral cleft lip.

severity increases, dropping from 57% of cases for severity grade 1 to 17% of cases for severity grade 4 (Fig. 5).

### DISCUSSION

Many techniques and protocols exist in cleft surgery, and there is a central need for institutions, hospitals, organizations, and individuals involved in cleft care to systematically measure and compare outcomes to guide best practices. Assessment of the appearance of the cleft-related deformity and the impact of surgical treatment is critical to the quality of life for this patient pool. However,

limitations of reliable, valid, and meaningful ways of measuring preoperative severity and postoperative aesthetic results have historically inhibited our ability to determine the best strategies for treatment.

The UCL CSI and UCL SOE were validated and published in 2017, and shown to be intuitive, easy to apply, and reliable for use by surgeons and laypersons. This study successfully implements these tools to systematically analyze aesthetic results after primary UCL/N repair in a large cohort of 1,823 patients. Three separate trained layperson evaluators reviewed all cases for a total of 5,469

total independent evaluations of initial cleft severity and final aesthetic outcome.

Perhaps most importantly, this study demonstrates that large-scale outcomes analysis after cleft lip repair is now possible in a manner that is simple, reproducible, relevant, and surgically applicable. Our experience also shows that it is time and cost effective to utilize layperson evaluators. Prior studies, including validation studies for the UCL CSI and UCL SOE, have generally found that surgeon and layperson rankings of pre- and postoperative cleft patients are consistent with good inter- and intrarater reliability.<sup>7-11</sup> In this study, layperson evaluators were provided minor but meaningful monetary compensation (USD \$10 per 50 cases). This incentivized them to complete evaluations quickly, generally within 1 week of receiving the cases. The total cost of for completing aesthetic outcomes evaluations of 1,823 patients in this study (5,469 evaluations) was USD \$1,093.80 (\$0.60 per patient).

This is the largest study of its kind, and the first published report to present normative distributions of aesthetic outcome scores after primary UCL/N repair controlled for initial cleft severity. These distribution curves now define a standard of “expected” aesthetic results after primary UCL/N repair. The distribution of aesthetic outcomes according to cleft severity clearly shows that aesthetic results worsen as cleft severity increases, with results more widely distributed and less predictable. It becomes increasingly less frequent for surgeons to achieve “perfect” or “near perfect” results (averaged aesthetic outcome score 7.00–8.00) for more severe clefts, and there are larger numbers of scores at the lower end of the distribution curves for more severe cases.

Although it is less common for surgeons to achieve “perfect” or “near perfect” results for more severe clefts, it is possible. Even among the most severe clefts (grade 4) in this series, 17% of postoperative results were scored as “perfect” or “near perfect.” This indicates that these surgeons are utilizing techniques capable of achieving excellence even with the most severe of cases.

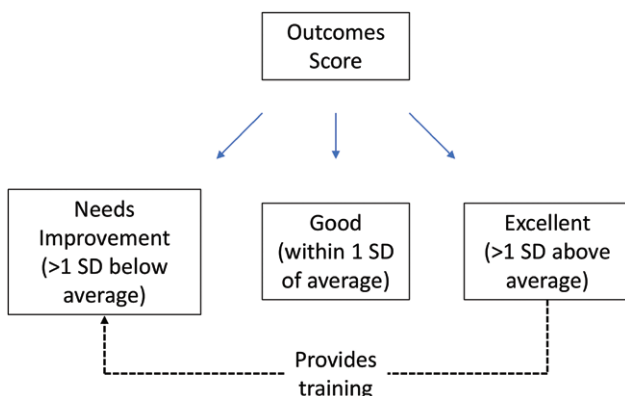
We have translated the data presented here into a guide for quality control for aesthetic outcomes after cleft

lip surgery (Fig. 6). Utilizing the methodology presented, all cases are evaluated by 3 layperson evaluators. All final scores falling within 1 SD of the average for a given cleft severity are classified as “good.” All final scores falling more than 1 SD above the average are termed “excellent,” and scores falling more than 1 SD below the average are termed “needs improvement.” Surgeons performing at the top of the curves are being identified to learn from their techniques and to utilize them as surgical educators. Surgeons receiving scores at the lower end of the curves (“need improvement”) are directed to educational resources and opportunities for additional training.

Our study is limited by several factors, including the nature of providing care to patients in a resource-limited setting. Patient follow-up is often more difficult in these settings, and thus our analysis of aesthetic outcomes is limited to early results. It is not clear how these results will translate into medium and long-term aesthetic outcomes, and further efforts will be required to improve the longitudinal scope of this research. The UCL CSI was used to classify preoperative severity and the UCL SOE was used to score postoperative outcome. Although these scales have been validated in previous studies, human evaluations are inherently prone to inconsistency and inaccuracy. Our methodology utilizes 3 reviewers to grade each case, and the large volume of 1,823 patients is intended to largely overcome these limitations, though no system using human evaluators is perfect. Also, despite the large number of patients evaluated, our outcomes tool was not able to detect statistically significant differences in scores for patients with grade 3 and grade 4 severity. We suspect that with even larger cohorts, we will be able to detect differences between these higher severity groups.

Continued evaluation of aesthetic outcomes on a large scale will pave the way for research that evaluates and compares various surgeons, centers, techniques, and protocols. These tools have additional value to track patient results through time and also to monitor surgical development during training and practice. The ability to objectively measure UCL/N surgical outcomes will provide insight into the factors that contribute to differences in outcomes among patients of similar initial severity, allowing identification of best practices that lead to superior results. This should also allow for development of an algorithmic approach to repair, employing specific maneuvers proven to best correct anatomic irregularities across the spectrum of cleft severity. Our goal is that over time, outcomes curves based around each cleft severity will shift toward higher median scores with more narrow distribution of results, translating to better outcomes for patients.

In conclusion, this study has achieved a standardized, timely, and cost-effective evaluation of 1,823 surgical cases of primary UCL/N repair. This data set provides a normal distribution of aesthetic results according to initial cleft severity and defines a standard of “expected” aesthetic results after primary UCL/N repair. Our results also show a clear correlation between initial severity and immediate aesthetic result after surgery, though we also show that excellent results are possible regardless of initial cleft severity.



**Fig. 6.** Scoring and quality improvement strategy using the cleft lip severity index and surgical outcomes evaluation scale. Surgeons who average greater than 1 SD above the mean after controlling for the severity of cases are recruited to provide additional training or assistance to surgeons who average less than 1 SD below the mean.

**Alex Campbell, MD**  
Division of Plastic Surgery  
Penn State Hershey Medical Center  
500 University Dr.  
PO Box 850  
Hershey, PA 17033  
E-mail: alexcampbellmd@gmail.com

#### REFERENCES

1. Fisher DM, Tse R, Marcus JR. Objective measurements for grading the primary unilateral cleft lip nasal deformity. *Plast Reconstr Surg*. 2008;122:874–880.
2. Witt PD, Marsh JL. Advances in assessing outcome of surgical repair of cleft lip and cleft palate. *Plast Reconstr Surg*. 1997;100:1907–1917.
3. Johnson N, Williams A, Singer S, et al. Initial cleft size does not correlate with outcome in unilateral cleft lip and palate. *Eur J Orthod*. 2000;22:93–100.
4. Holtmann B, Wray RC. A randomized comparison of triangular and rotation-advancement unilateral cleft lip repairs. *Plast Reconstr Surg*. 1983;71:172–179.
5. Tanikawa DY, Alonso N, Rocha DL. Evaluation of primary cleft nose repair: severity of the cleft versus final position of the nose. *J Craniofac Surg*. 2010;21:1519–1524.
6. Haley CB, S., Jones C. Unilateral cleft lip severity affects aesthetic outcome. 32nd Annual Meeting, Northeast Society of Plastic Surgeons; September 18, 2015, 2015; Philadelphia, PA.
7. Campbell A, Restrepo C, Deshpande G, et al. Validation of the unilateral cleft lip severity index for surgeons and laypersons. *Plast Reconstr Surg Glob Open*. 2017;5:e1479.
8. Campbell A, Restrepo C, Deshpande G, et al. Validation of a unilateral cleft lip surgical outcomes evaluation scale for surgeons and laypersons. *Plast Reconstr Surg Glob Open*. 2017;5:e1472.
9. Meng T, Ma L, Wang Z. Comparative assessment of the cleft profile by patients with cleft lip and palate, cleft surgeons, and lay people. *Aesthetic Plast Surg*. 2015;39:757–763.
10. Tobiasen JM, Hiebert JM, Boraz RA. Development of scales of severity of facial cleft impairment. *Cleft Palate Craniofac J*. 1991;28:419–424.
11. Yao CA, Imahiyero T, Swanson J, et al. The smile index: Part 1. A large-scale study of phenotypic norms for preoperative and postoperative unilateral cleft lip. *Plast Reconstr Surg*. 2018;141:137–146.