

Perth scoring system for the surgical audit of the repaired unilateral cleft lip

Linda Monshizadeh MBBS(Hons) FRACS (Plast Surg) MS (Craniofacial),¹ Vijith Vijayasekaran MBBS(Hons) FRACS (Plast Surg)¹

¹ Department of Plastic Surgery
Perth Children's Hospital
Perth, Western Australia
AUSTRALIA

OPEN ACCESS

Correspondence

Name: Linda Monshizadeh

Address: Department of Plastic Surgery
Perth Children's Hospital
15 Hospital Avenue
Nedlands, Western Australia, 6009
AUSTRALIA

Email: lindamonshizadeh@gmail.com

Phone: +61 (0)8 6456 2222

Citation: Monshizadeh L, Vijayasekaran V. Perth scoring system for the surgical audit of the repaired unilateral cleft lip. *Australas J Plast Surg.* 2020; 3(1):22–29. <https://doi.org/10.34239/ajops.v3n1.133>

Received: 3 February 2019

Accepted for review: 30 April 2019

Accepted for publication: 13 November 2019

Copyright © 2020. Authors retain their copyright in the article. This is an open access article distributed under the Creative Commons Attribution Licence which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

Section: Cleft lip and palate

Abstract

Background: Currently, there is no standardised assessment tool to assess facial aesthetics in cleft lip and palate surgery. Inter-centre comparison is hampered by the use of different aesthetic indices with low intra- and inter-rater reliability.

Aim: The Perth scoring system is a new assessment tool for unilateral cleft lip which scores four key components of the cleft lip/nose repair: lip length, white roll, alar insertion point and vermillion. The aim of this study was to validate the Perth scoring system as a reliable and useful new assessment tool and to demonstrate the use of the scoring system to measure improvements after cleft lip revision.

Method: Nineteen patients who underwent cleft lip revision by the senior author were selected. Pre- and postoperative photos were presented to a panel of raters to score. Scores were analysed to determine the intra- and inter-rater reliability and to compare outcomes.

Results: Almost all patients (15/16) had improvement in scores (range 1.09–5.59) after cleft lip revision. Intra-raters' agreement scores from lowest to highest were: lip length (0.65), white roll (0.7), alar insertion point (0.78) and vermillion (0.78). The total intra-class correlation coefficient was 0.96 (0.94–0.98, 95% CI, $P < 0.000$).

Conclusion: This new scoring system is a valid and useful tool for assessment of the unilateral cleft lip. The high rate of intra- and inter-rater reliability allow it to serve as a useful tool to compare surgical outcomes both within and between centres. Further field testing with a larger cohort of patients is required.

Keywords: cleft lip, surveys and questionnaires, observation variation, research design, outcome assessment

Introduction

There is no doubt that society has become obsessed with physical image. In the advent of the 'selfie' we have been bombarded by images of glamour and beauty. Advances in technology and the editing and manipulation of photos driven by social media has reinforced the drive towards perfection. This has led to users becoming more anxious, less confident and feeling less physically attractive afterwards.¹⁻³ Needless to say, patients with any physical deformity are left feeling marginalised compared to their peers in a world where even 'normal' appears not to be good enough. Patients with a cleft lip and/or palate continue to be faced with the stigma of their appearance.

Early psychometric findings from the Cleft-Q study show that participants who were unhappy with how they looked reported significantly lower scores on all appearance and health-related quality of life scales. Mean scores for health-related quality of life and speech scales were highest for the group without a speech problem and lowest for the group with a moderate or severe speech problem.⁴ Appearance and speech therefore appear to be the two important factors influencing quality of life.

Assessment of aesthetic outcomes is a difficult problem. Nevertheless, it should be routinely included as part of the 'total' assessment for all cleft patients. Assessment may be performed to audit one surgeon, to compare different surgical techniques by one or multiple surgeons, or to make comparison among different centres. Timing of this assessment may be taken, as needed, at different stages along the cleft treatment pathway to answer a proposed clinical question.

One of the inherent uncontrollable factors affecting outcomes for cleft patients is the severity of the deformity. Although sharing many similarities, unilateral and bilateral clefts have unique outcomes and should be assessed independently. Similarly, the isolated cleft lip may be considered a distinct entity⁵ to the cleft lip and palate deformity. Despite this, however, external differences cannot be appreciated by a layperson in a social setting allowing both types to be assessed together with the same assessment tool.

The final aesthetic outcome for any cleft patient is the totality of treatments spanning a lifetime from birth to maturity along the cleft protocol pathway and may include some or all of the following: use of presurgical orthopaedics/nasoalveolar molding primary lip/nose repair, with or without postoperative nasal splinting, orthodontics, alveolar bone grafting, orthognathic surgery, cleft lip and/or nose revision. How each of these factors contributes to the final outcome remains to be clear. Nevertheless, it is important to remember that primary lip and nose repair is only one part of a patient's overall aesthetic outcome. In future, acknowledging all the techniques used by one centre while assessing outcomes will allow for better comparison through multivariate analysis. Lastly, and most relevant to this study, is that comparing aesthetic outcomes is further complicated by centres using different aesthetic scoring methods, the majority of which have low intra- and inter-rater reliability.⁶⁻⁹

The most commonly used aesthetic measurement is the Asher-McDade index¹⁰ which uses a five-point scale to assess four facial features: nasal form, nasal symmetry, alignment of the vermilion border (frontal view) and nasolabial profile (lateral view). More recently, multiple new reference photographs were added to a modified Asher-McDade index; the result was named the Nasolabial yardstick after the Goslon yardstick.¹¹ The Asher McDade index has been used in many large, multicentre trials.^{12,13} The main problems reported with this scoring system are its' cumbersome nature, lack of objectivity, need for expert opinion for better reliability and focus on the nose.^{8,13}

Since the Asher-McDade index there have been numerous other scales of various reliability, however, a full literature review is beyond the scope of this article. Systematic reviews of published articles⁶⁻⁹ have found the majority of these scales use qualitative scoring systems of a subjective nature and reference photographs that are similar to the Asher-McDade index. In summary, there is a large number of rating scales requiring validation. Newer 3D and computer assisted technologies are emerging but are currently costly and not feasible.

No doubt these techniques will be more important in the future.

Aim

The authors sought to create a new tool for the assessment of the unilateral cleft named the Perth scoring system. The main aim was to ensure the assessment tool was easy to use for both surgeons and lay people, objective, reliable, effective in capturing a true assessment of a unilateral cleft lip and could be used at any stage along the cleft treatment protocol.

This scoring method is used to assess symmetry of four vital components of the cleft lip/nose repair (Figure 1), namely alar insertion point, lip length, white roll and vermillion. Each of the four components are scored using clear and objective descriptive criteria either from (0–3) or (0–2). The total score ranges from zero (best score) to nine (worst score). The purpose of this study was to validate the Perth scoring system as a reliable tool to assess the unilateral cleft lip/nose deformity.

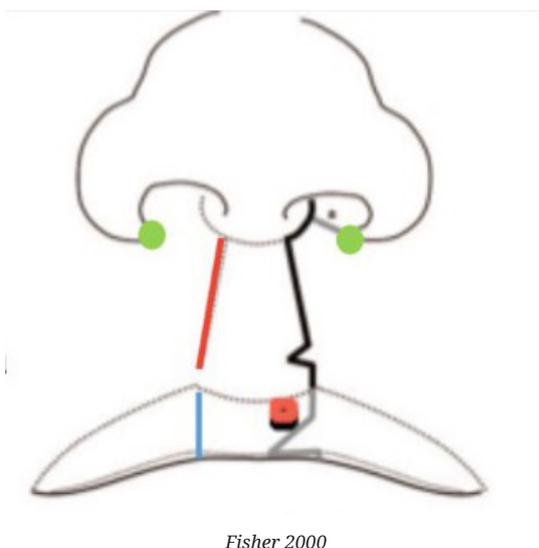
Method

Patients who underwent cleft lip revision by the senior author from 2008–2014 were identified using the hospital’s craniofacial clinical diary and hospital theatre database. Exclusion criteria included patients with bilateral cleft, prior revision surgeries, syndromes or inadequate photographs.

Pre- and postoperative revision photographs were collected and duplicates of each photo made. Photos were cropped to show the nose and lips only (figures 2–4). The shuffled photos were presented as projected slides at a conference. Raters were unaware that the photos included the same patients who had undergone surgical revision.

Prior to commencement, the scoring system was explained to the raters and three examples were shown (figures 2–4) with reasons as to why certain scores had been given for each case. Time was given to answer any questions regarding the scoring system. Slides were then presented in two sessions to reduce rater fatigue.

Scores were statistically analysed using Cohen’s kappa coefficient (IBM SPSS Statistics®, 1 New Orchard Road, Armonk, New York, 10504-1722, United States) to assess intra-rater and inter-rater reliability. Overall reliability of the scale was calculated using the intraclass correlation coefficient (ICC). Results were classified as poor (<0.2), fair (0.21–0.40), moderate (0.41–0.60), good (0.61–0.8) or very good (0.81–1.00). Final analysis was also made to compare preoperative to postoperative scores to determine if there had been significant improvements in appearance following revision.



- Alar insertion point:** ●
 - 0 Similar position as symmetry in nostrils
 - 1 Similar position but unequal nostril size
 - 2 Different position and unequal nostril size

- Lip length:** —
 - 0 Equal lip height and good scar
 - 1 Equal lip height but poor scar
 - 2 Long lip height irrespective of scar
 - 3 Short lip height irrespective of scar

- White roll:** —
 - 0 Continuous and well defined white roll (Nordoofs point identified correctly)
 - 1 Continuous but lacking well defined white roll
 - 2 Discontinuous and lacking well defined white roll

- Vermillion:** —
 - 0 Symmetrical lip fullness
 - 1 Vermillion excess
 - 2 Vermillion deficiency / whistle deformity

Fig 1. Scoring index. Scores can range from best (0) to worst (9)



Fig 2. Example one

Alar insertion point:

- 0 Similar position as symmetry in nostrils
- 1 Similar position but unequal nostril size
- 2 Different position and unequal nostril size

Lip length:

- 0 Equal lip height and good scar
- 1 Equal lip height but poor scar
- 2 Long lip height irrespective of scar
- 3 Short lip height irrespective of scar

White roll:

- 0 Continuous and well defined white roll (Nordoofs point identified correctly)
- 1 Continuous but lacking well defined white roll
- 2 Discontinuous and lacking well defined white roll

Vermillion:

- 0 Symmetrical lip fullness
- 1 Vermillion excess
- 2 Vermillion deficiency / whistle deformity



Fig 3. Example two

Alar insertion point:

- 0 Similar position as symmetry in nostrils
- 1 Similar position but unequal nostril size
- 2 Different position and unequal nostril size

Lip length:

- 0 Equal lip height and good scar
- 1 Equal lip height but poor scar
- 2 Long lip height irrespective of scar
- 3 Short lip height irrespective of scar

White roll:

- 0 Continuous and well defined white roll (Nordoofs point identified correctly)
- 1 Continuous but lacking well defined white roll
- 2 Discontinuous and lacking well defined white roll

Vermillion:

- 0 Symmetrical lip fullness
- 1 Vermillion excess
- 2 Vermillion deficiency / whistle deformity



Fig 4: Example three

Alar insertion point:

- 0 Similar position as symmetry in nostrils
- 1 Similar position but unequal nostril size
- 2 Different position and unequal nostril size

Lip length:

- 0 Equal lip height and good scar
- 1 Equal lip height but poor scar
- 2 Long lip height irrespective of scar
- 3 Short lip height irrespective of scar

White roll:

- 0 Continuous and well defined white roll (Nordoofs point identified correctly)
- 1 Continuous but lacking well defined white roll
- 2 Discontinuous and lacking well defined white roll

Vermillion:

- 0 Symmetrical lip fullness
- 1 Vermillion excess
- 2 Vermillion deficiency / whistle deformity

Results

Nineteen patients were identified through the hospital database who had undergone a cleft lip revision during the allocated time period and met the inclusion criteria. Each case had four photographs (duplicates of pre- and postoperative photographs) resulting in 76 slides.

Thirteen audience members completed the assessment and eleven of these had suitable results (there were two accidentally incomplete papers). Raters included four plastic surgery consultants, five plastic surgery trainees, one radiologist and one business representative.

Tables 1–4 show statistical results for intra-rater reliability for the four components of the scoring index (Figure 1); lip length, white roll, alar insertion point and vermillion. Cohen’s kappa results are shown for each rater. Intra-rater agreement scores from lowest to highest were: lip length (0.65), white roll (0.7), alar insertion point (0.78) and vermillion (0.78). Inter-rater reliability was extremely high for all four components of the scoring index (range 0.84–0.97, refer **Table 5**). Total ICC was 0.96 (P<0.000)

In comparing pre- and postoperative scores, three cases were excluded as postoperative photos were taken too early and scars appeared immature. Almost all the resultant cases (15/16) showed significant improvement in scores post revision (Table 6). The mean sum of preoperative scores for all photographs was 46.97 (mean 3.13) compared to 31.07 (mean 2.07) post revision surgery (p <0.007 paired t-test). The range of improvement was 1.09–5.59.

Discussion

Any new tool used in the analysis of an outcome must agree with the requirements of scientific reproducibility.

The Perth scoring system appears to be a valid and reliable tool to assess components of the unilateral cleft lip with results showing good/very good intra-rater and very good inter-rater reliability. The use of a single anterior 2D photograph was chosen as a more simple and practical option which better reflected the social scenario of daily life where

Table 1: Intra-rater alar insertion point

Rater	Intra-rater reliability	95% CI
1	0.717	(0.455–0.853)
2	0.788	(0.581–0.893)
3	0.640	(0.308–0.813)
4	0.621	(0.271–0.803)
5	0.748	(0.515–0.869)
6	0.894	(0.791–0.946)
7	0.829	(0.672–0.911)
8	0.831	(0.675–0.912)
9	0.837	(0.686–0.915)
10	0.703	(0.429–0.846)
11	0.739	(0.497–0.864)
Total	0.780	(0.550–0.850)

Table 2: Intra-rater lip length

Rater	Intra-rater reliability	95% CI
1	0.618	(0.227–0.811)
2	0.736	(0.465–0.864)
3	0.777	(0.537–0.892)
4	0.371	(0.233–0.679)
5	0.664	(0.340–0.828)
6	0.801	(0.601–0.900)
7	0.579	(0.157–0.790)
8	0.031*	(0.920–0.511)
9	0.607	(0.237–0.798)
10	0.882	(0.778–0.978)
11	0.229**	(0.599–0.628)
Total	0.650	(0.550–0.850.)

*=radiologist; **=trainee (outliers)

Table 3: Intra-rater white roll

Rater	Intra-rater reliability	95% CI
1	0.816	(0.626–0.909)
2	0.385	(0.219–0.689)
3	0.729	(0.474–0.860)
4	0.781	(0.578–0.886)
5	0.910	(0.827–0.953)
6	0.633	(0.280–0.813)
7	0.590	(0.136–0.785)
8	0.668	(0.334–0.834)
9	0.646	(0.306–0.820)
10	0.896	(0.800–0.946)
11	0.423**	(-0.111–0.700)
Total	0.700	(0.550–0.850)

**=trainee (outliers)

Table 4. Intra-rater vermillion

Rater	Intra-rater reliability	95% CI
1	0.896	(0.796–0.947)
2	0.845	(0.682–0.924)
3	0.941	(0.883–0.970)
4	0.504	(0.045–0.742)
5	0.840	(0.689–0.918)
6	0.747	(0.499–0.872)
7	0.940	(0.882–0.969)
8	0.686	(0.390–0.838)
9	0.584	(0.192–0.786)
10	0.964	(0.930–0.981)
11	0.477	(0.006–0.728)
Total	0.780	(0.670–0.850)

Table 5. Inter-rater reliability

AIP	Lip length	White roll	Vermillion
0.97	0.84	0.9	0.96

Total ICC for total scores 0.963 (0.941–0.980) 95% CI (p<0.000).
AIP= alar insertion point CI=confidence interval

Table 7. Interpretation of Kappa statistics

Strength of agreement	Kappa value
Poor	<0.20
Fair	0.21–0.40
Moderate	0.41–0.60
Good	0.61–0.80
Very good	0.81–1.00

Table 6. Overall preoperative versus postoperative scores

Paired samples test								
Paired diff	Mean	Stand dev	Stand err Mean	95% CI of the difference		t	df	Sig(2-tailed)
				Lower	Upper			
Total preop score-total postop score	15.9	19.66523	5.07754	5.00976	26.79024	3.131	14	p=0.007

CI=confidence interval; df=degrees of freedom; paired diff=paired difference; preop=preoperative; postop=postoperative; sig(2-tailed)=statistical significance; stand err=standard error; stand dev=standard deviation; t=t-test.

patients have to interact ‘face-to-face’ with their peers. Furthermore, photographs are cropped to show only the nasolabial area by to reduce the influence of background facial attractiveness upon the assessment of cleft impairment.^{14,15}

In contrast to previous methods, we used more objective criteria to establish the scores that are described using text alongside each rating. The categories chosen for assessment correspond to key areas in cleft lip and nose repair. In effect, the accuracy of repair and hence aesthetic outcome is translated into a scoring system. As a result, our scoring system is more useful in assessing the need for revision as it assesses— and places emphasis (with a higher score)— on key problems such as an asymmetrical alar base, short lip (highest score), discontinuous white roll and significant vermillion deficiency. Although some studies report 2D images as insufficient to assess scar quality, we believe it is still an important component and should be included in the scoring system.

Intra-rater agreement ranged from good to very good in our pilot study for each component (Table 7). Further field studies are required to test the reproducibility of this scoring method. It is well known that reproducibility is further improved by familiarity. Practice rater tasks are therefore highly recommended prior to commencing the scoring of subjects.^{6,16,17}

Comparison of this tool to other published studies of cleft lip and palate (CLP) aesthetic indices are favourable. The most frequently used index developed by Asher-McDade reported fair to good inter-examiner rater agreement from their pilot study. The greater number of categories and reference photographs make the scoring system cumbersome and associated with poorer reproducibility. This index was subsequently utilised in the euro-cleft study with only the frontal and lateral photograph where lower inter-rater agreement was reported.¹⁸

Our review of published articles found one aesthetic index that was similar to our scoring system.¹⁹ The unilateral cleft lip surgical outcomes evaluation (UCLSOE) index scores symmetry of four individual components of the cleft repair (Cupid's bow, lateral lip, nose and free vermillion). Each element is scored on a three point scale: excellent (2), (mild asymmetry (1), unsatisfactory (0). The four individual scores are then summed for a total score of lowest (0) to highest (8).

This index would be a useful tool to compare different surgical techniques as it assesses detailed parts of the lip repair, for example, comparison of both horizontal and vertical height of the lateral lip and assessment of the nose and lip also from a 'worm'-eye view. Weaknesses of this index pertain to the need for two photographs in assessment as well as the use of subjective terminology such as 'mild' or 'marked' in the decision making. In addition, the index fails to distinguish more important problems from one another during the decision process. For example, in their assessment of the lateral lip both vertical and horizontal asymmetry are grouped together. This fails to recognise that a shorter vertical height is far more significant compared to a horizontal discrepancy which is usually not revised.

In contrast, our scoring system only allows a choice for a problem to be either absent or present by excluding use of subjective terminology such as 'mild', 'moderate', 'severe' and so on. This forces the rater to give the worst score for only significant discrepancies which translates to identification of obvious pathology. This makes it easy for different raters to choose a similar score which is shown by the very high intra- and inter-rater reliability. In addition, greater weighting has been given to stigmatising features such as short lip compared to a longer lip, significant deficiency of the vermillion (whistle deformity) or a discontinuous (step) white roll. A high score of nine would indicate failure in all key areas of the lip and nose repair necessitating a full cleft lip revision. The average scores for our own case series of patients who required revision was 3.13 with a mean improvement by one point post-revision (2.07).

Use of figure 1 rather than reference photographs also makes it easier to visualise the areas for each component being assessed with colours to help with faster identification. This is less time consuming than looking at reference photographs, particularly for large case numbers needing assessment.

One question raised has been whether raters with different levels of expertise differ significantly with their scoring and outcomes. Our sample size of lay people was too small to show any significant discrepancies and this will need to be studied further with a larger test group. A systematic review by Zhu examined this question.²⁰ Eleven articles were studied and the results were inconclusive with three studies reporting that laypeople were more critical than professionals, three studies reporting no significant difference between laypeople and professionals, and five studies reporting that professionals were more critical than laypeople when assessing facial appearance of patients with CLP.

Conclusion

The Perth scoring system appears to be a valid and useful tool for assessing unilateral cleft lip at any stage along the cleft treatment protocol. It was created to address problems found in previous rating systems which were too subjective and/or cumbersome. The high intra- and inter-rater reliability allow it to serve as a useful tool to compare surgical outcomes both within and between centres. Although designed primarily for cleft patients, it may be applied to all types of lip analysis (for example, after trauma). It will be necessary to conduct further field testing with a larger cohort to ensure the system's surgical reproducibility among different centres and also for use by lay people.

References

- 1 Lonergan AR, Bussey K, Mond J, Brown O, Griffiths S, Murray SB, Mitchison D. Me, my selfie, and I: the relationship between editing and posting selfies and body dissatisfaction in men and women. *Body Image*. 2019 Mar;28:39–43. <https://doi.org/10.1016/j.bodyim.2018.12.001> PMID:30572289
- 2 Fardouly J, Rapee RM. The impact of no-makeup selfies on young women's body image. *Body Image*. 2019 Mar;28:128–134. <https://doi.org/10.1016/j.bodyim.2019.01.006> PMID:30665030

- 3 Mills JS, Musto S, Williams L, Tiggemann M. 'Selfie' harm: effects on mood and body image in young women. *Body Image*. 2018 Dec;27:86–92.. <https://doi.org/10.1016/j.bodyim.2018.08.007> PMID:30149282
- 4 Klassen AF, Riff KWW, Longmire NM, Albert A, Allen GC, Aydin MA, Baker SB, Cano SJ, Chan AJ, Courtemanche DJ, Dreise MM, Goldstein JA, Goodacre TEE, Harman KE, Muni-ll M, Mahony AO, Aguilera MP, Peterson P, Pusic AL, Slator R, Stiernman M, Tsangaris E, Tholpady SS, Vargas F, Forrest CR. Psychometric findings and normative values for the CLEFT-Q based on 2434 children and young adult patients with cleft lip and/or palate from 12 countries. *Can Med Assoc J*. 2018;190(15):e455–e462. <https://doi.org/10.1503/cmaj.170289> PMID:29661814 PMCID:PMC5903887
- 5 Carroll K, Mossey PA. Anatomical variations in clefts of the lip with or without cleft palate. *Plast Surg Int*. 2012;54:2078. <https://doi.org/10.1155/2012/542078> PMID:23251795 PMCID:PMC3517834
- 6 Mosmuller DG, Griot JP, Bijnen CL, Niessen FB. Scoring systems of cleft-related facial deformities: a review of literature. *Cleft Palate Craniofac J*. 2013;50:286–296. <https://doi.org/10.1597/11-207> PMID:23030761
- 7 Al-Omari I, Millett DT, Ayoub AF. Methods of assessment of cleft related facial deformity: a review. *Cleft Palate Craniofac J*. 2005;42:145–156. <https://doi.org/10.1597/02-149.1> PMID:15748105
- 8 Al-Omari I, Millett DT, Ayoub A, Bock M, Ray A, Dunaway D, Crampin L. An appraisal of three methods of rating facial deformity in patients with repaired complete unilateral cleft lip and palate. *Cleft Palate Craniofac J*. 2003;40:530–537. https://doi.org/10.1597/1545-1569_2003_040_0530_aaotmo_2.0.co_2
- 9 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;65:1233–1245. <https://doi.org/10.1016/j.bjps.2012.04.001> PMID:22591614
- 10 Asher-McDade C, Roberts C, Shaw WC, Gallagher 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. https://doi.org/10.1597/1545-1569_1991_028_0385_doamfr_2.3.co_2
- 11 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;20:1683–1686 <https://doi.org/10.1097/SCS.0b013e3181b3ed9c> PMID:19816333
- 12 Shaw WC, Asher-McDade C, Brattstrom V, Dahl E, McWilliam J, Molsted K, Plint DA, Prah-Andersen B, Semb G, 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. https://doi.org/10.1597/1545-1569_1992_029_0409_asciso_2.3.co_2 PMID:1472518
- 13 Mercado A, Russell K, Hathaway R, Daskalogiannakis J, Sadek H, Long RE Jr, Cohen M, Semb G, Shaw W. The Americleft study: an inter-center study of treatment outcomes for patients with unilateral cleft lip and palate part 4. Nasolabial aesthetics. *Cleft Palate Craniofac J*. 2011; 48:259–264. <https://doi.org/10.1597/09-186.1> PMID:21219227
- 14 Mosmuller DG, Bijnen CL, Kramer GJ, Disse MA, Prah C, Kuik DJ, Niessen FB, Don Griot JP. The Asher-McDade aesthetic index in comparison with two scoring systems in nonsyndromic complete unilateral cleft lip and palate patients. *J Craniofac Surg*. 2015 Jun;26(4):1242–5.. <https://doi.org/10.1097/SCS.0000000000001784> PMID:26080166
- 15 Asher-McDade C, Roberts C, Shaw WC, Gallagher C. Development of a method for rating nasolabial appearance in patients with clefts of the lip and palate. *Cleft Palate Craniofac J*. 1991;28:385–391. https://doi.org/10.1597/1545-1569_1991_028_0385_doamfr_2.3.co_2
- 16 Tobiasen JM, Hiebert JM, Boraz RA. Development of scales of severity of facial impairment. *Cleft Palate Craniofac J*. 1991;28: 419–424. [https://doi.org/10.1597/1545-1569\(1991\)028<0419:DOSOSO>2.3.CO;2](https://doi.org/10.1597/1545-1569(1991)028<0419:DOSOSO>2.3.CO;2) PMID:1742313
- 17 Howells DJ, Shaw WC. The validity and reliability of ratings of dental and facial attractiveness for epidemiological use. *Amer J Orthodontics*. 1985;88:402–408. [https://doi.org/10.1016/0002-9416\(85\)90067-3](https://doi.org/10.1016/0002-9416(85)90067-3)
- 18 Shaw, WC. Semb G, Nelson P, Brattström V, Mølsted K, Prah-Andersen B, Gundlach KK. The Eurocleft project 1996–2000: overview. *J Cranio Maxillo Surg*. 2001(29);131–140, discussion 141–132. <https://doi.org/10.1054/jcms.2001.0220> PMID:11403549
- 19 Campbell A, Restrepo C, Deshpande G, Tredway C, Bernstein SM, Patzer R, Wendby L, Schonmeyr B. Validation of a unilateral cleft lip surgical outcomes evaluation scale for surgeons and laypersons. *Plast Reconstr Surg Glob Open*. 2017;5(9):e1472. <https://doi.org/10.1097/GOX.0000000000001472> PMID:29062644 PMCID:PMC5640349
- 20 Zhu S, Jayaraman J, Khambay B. Evaluation of facial appearance in patients with cleft lip and palate by laypeople and professionals: a systematic literature review. *Cleft Palate Craniofac J*. 2016;53(2):187–196. <https://doi.org/10.1597/14-177> PMID:25650654