



A 37-year follow-up after Skoog lip repair of unilateral cleft lip and palate: what is the long-term verdict on lip appearance by different panels?

Roshan Peroz¹ · Alberto Falk-Delgado² · Stina Arwidson³ · Malin Hakelius¹ · Maria Mani¹

Received: 21 May 2024 / Accepted: 10 September 2024
© The Author(s) 2024

Abstract

Background The study examined long-term lip appearance after Skoog's lip repair in unilateral cleft, lip, and palate (UCLP) patients. To include different views, three different panels respectively, rated level of lip appearance and desire for improvement. The secondary aim was to investigate the correlation between the self-reported satisfaction of UCLP patients and controls with the panel assessments.

Methods Cropped frontal photographs of the lips of 74 UCLP-treated individuals and 61 non-cleft individuals were assessed twice by three panels of five cleft surgeons (Professional panel), nine laymen (Layman panel), and five individuals with CLP (Cleft panel) using a web-based platform, at a mean follow-up of 37 years post-lip surgery.

Results All three panels rated the lip appearance of UCLP patients worse than that of the controls ($p < 0.001$). The Professional panel reported higher satisfaction levels and less desire to change lip appearance than the Layman and Cleft panels. Intra-rater agreements were, in general, moderate in the Professional panel ICC (0.57–0.82) and Laymen panel ICC (0.63–0.73) and poor in the Cleft panel ICC (0.21–0.36). The inter-rater agreement was, in general, poor in all panels. No significant correlation was found between the patient's Self-assessments and any of the panel assessments.

Conclusions All panels' rated lip appearance of adults born with UCLP treated with Skoog's technique as worse than that of non-cleft controls. The study highlights a significant discrepancy between the Professional panel's more favorable assessment and the more critical views of the laymen and cleft panels, contradicting the initial hypothesis of uniform perception across panels. The lack of correlation between patient self-assessment and panel assessments underscores the subjective nature of lip satisfaction, emphasizing the need for personalized patient care strategies in CLP treatment outcomes.

Level of evidence Not gradable.

Keywords Cleft lip and palate · Skoogs lip repair · Panel assessment · Intra-rater reliability · Inter-rater reliability · Correlation analysis

Background

Facial aesthetics is an important outcome in cleft, lip, and palate (CLP) care and has been shown to be an aspect strongly associated with patients' psychological well-being [1–6]. Perception of attractiveness is, however, individual and complex. People with CLP are treated according to specified surgical protocols, which may vary between centers. In most facilities, a surgical lip repair at 3 to 6 months is the first surgical procedure performed on a child born with CLP. Multiple cleft lip repair techniques have been described, including geometric, z-plasty, and rotation advancement flaps [7–11]. The common goal of these procedures is to

✉ Roshan Peroz
roshan.peroz@surgsci.uu.se

¹ Department of Plastic and Reconstructive Surgery, Department of Surgical Sciences, Uppsala University and Uppsala University Hospital, Uppsala 751 85, Sweden

² Department of Plastic and Reconstructive Surgery, Karolinska University Hospital, Stockholm, Sweden

³ Department of Ear, Nose, and Throat Diseases, Hudiksvalls Hospital, Hudiksvall, Sweden

normalize symmetry and function. However, scar formation and contraction can disturb the symmetry, resulting in unfavorable outcomes [12, 13]. When it comes to secondary procedures with the aim to improve lip appearance, several factors may play a role, and the decision for surgery or not may be influenced by the cleft surgeon, the social community, family, and peers, in addition to the patient's own perception and wishes.

Given the limited number of studies evaluating long-term outcomes following lip repair, we aim to report panel assessments of lip appearance, in adults who were treated with Skoog's lip repair technique as a child. This technique has been our institution's standard for all patients for the last 60 years [8]. Previous studies by our team have reported on nasolabial appearance, self-reported satisfaction with lip appearance, complications, and the need for subsequent lip surgeries [14–16]. However, to appreciate different assessors' opinions, and give a broader view of long-term assessment of lip appearance after Skoog's lip repair surgery we have for the first time, asked panels to assess the long-term lip appearance.

There are no universally accepted methods for evaluating facial appearance [17]. Several methods have been described for assessing facial aesthetics in previous studies, such as clinical assessment [18], two-dimensional photographs [19, 20], three-dimensional evaluations [21, 22] and video graphic assessments [23, 24]. Asher-McDade described a widely used method in which different combinations of raters rate cropped photographs of the nasolabial region on a Likert scale [19]. The agreement between specialists and laypeople on assessing facial appearance are inconclusive; some studies found agreement between these groups [25–27], whereas others found disagreement [15, 28].

Attractiveness is highly subjective, and decision-making for further surgical treatment can be influenced not only by patients' own opinions but also by care providers, the social community, and the people around the person. Therefore, evaluating how different groups of people perceive lip appearance and whether there is consensus among them is crucial. Understanding these perceptions helps customize treatment approaches that align with diverse expectations and enhance patient satisfaction. In the clinical setting, the decision for further surgical secondary correction of the lip is based on the subjective judgment of the surgeon in conjunction with the patient and/or the parents. Additionally, peer pressure and instances of teasing can influence individuals. Therefore, it is important to know how professionals with experience in treating people with CLP, lay people, and people with knowledge of having a CLP malformation assess the appearance and need for further surgical intervention as well as to understand the correlation between the assessments.

To the best of our knowledge, no earlier studies have reported on long-term panel-assessed satisfaction with lip appearance among adults treated with Skoog's lip repair technique. The primary aim of the current study is to evaluate the panels' long-term verdict on the lip appearance of UCLP patients treated with Skoog's technique, hypothesizing that there is no difference in the panels' perception of labial aesthetics. The secondary aim is to investigate how peoples' subjective satisfaction with lip appearance (both of cleft patients and controls) correlates to that of the panels. Evaluation strategies employed by the participating panel members is investigated.

Material, patients, and methods

Subjects

The study was conducted at Uppsala University (UU) hospital, Uppsala, Sweden, which serves as the primary provider of tertiary healthcare services for a population of approximately 1,500,000 individuals. Notably, the UU hospital is the sole facility in the region offering specialized care for patients with cleft lip and palate (CLP). The present cohort is part of a long-term follow-up project performed in Uppsala which recruited adult UCLP patients and non-operated controls for assessment during 2007–2009. Previous studies have described the project set-up and reported on parts of the findings and scientific results [4, 15, 29, 30]. In summary, for this long-term follow-up project, all consecutive patients with, non-syndromal, complete unilateral cleft lip and palate (UCLP), treated at UU hospital, born 1960–1987 were considered ($n = 128$). All patients had been treated with primary lip repair at three months of age, using Skoog's lip repair technique (mean follow-up time 37 years (range 20–47 years). Palate closure had been performed using either a one-stage or two-stage method, depending on the birth year. Patients born from 1960 to 1975 had received the one-stage procedure, based on the Veau-Wardill method and later modified by Skoog, at an average age of 1.9 years [31–33]. Those born between 1976 and 1987 had undergone the two-stage technique [34–36]. No primary nose surgery had been performed on the patients until adolescence, during which minimal or no nasal operations occurred. Secondary surgeries on the lip and nose, along with pharyngoplasty, had been tailored to the individual needs of the patients. The surgical treatment protocol has previously been described in detail [30]. Patients with incomplete clefts, associated syndromes, or other malformations were excluded. Nineteen patients were excluded from the study for the following reasons: Six patients had passed away, five patients exhibited physical or mental incapacity, five patients were residing

abroad, and three patients could not be located in the Swedish population registry.

As previously described by Mani et al. [29] the remaining 109 eligible patients were initially invited through an information letter and followed up with a telephone call. Ultimately, eighty-three patients (76%) actively participated in the study, including participating in photography sessions and completing the questionnaires. There were no differences in gender and age between the participating and non-participating groups. Additionally, a control group of 67 individuals with no clefts was recruited and underwent the same study protocol. In the current study, the photographic records of nine patients and six controls were incomplete and thus excluded from the analysis. Consequently, images of 74 UCLP patients and 61 controls without cleft were evaluated.

Self-assessment of lip appearance

For self-assessment of lip appearance, questions from two of the questionnaires used in the follow-up study were included: The Satisfaction with Appearance Scale (SWA) and the modified version of the Body Cathexis- Scale [37, 38]. Within the scope of the current study, a specific component of the SWA, addressing self-assessment of lip appearance, was used; “What do you think of this part of your face? (lips).” Respondents were provided along a Visual Analogue Scale (VAS), with 0 indicating a very high level of satisfaction and 10 indicating a low level of satisfaction. The modified version of the Body Cathexis Score comprises 22 items addressing various aspects of facial appearance, desire for further treatment, and speech function [38, 39]. The question related to the desire for further lip treatment was included in the current study. Specifically, “If it were possible, would you like to change the appearance of your lips?” Respondents used a VAS ranging from 0 to 10 (0 indicated “not at all,” and 10 indicated “very much”) [40].

The questionnaires were distributed to patients and controls to be answered at home. During the subsequent visit, organized for the follow-up project of this cohort, a physician not involved in the study reviewed the questionnaires with the patients to minimize the number of unanswered questions and potential misunderstandings [15]. Self-evaluation of lip appearance has been collected and described in detail in previous studies [16, 30]. The procedures and specific data for the patients and the controls and their reported satisfaction with lip appearance, as well as their self-reported desire to change lip appearance, have previously been reported [16].



Fig. 1 Photograph of subject from frontal view



Fig. 2 A cropped photograph of the lips, as shown to the panel in the assessments

Image processing

A professional photographer at UU hospital took all the photographs at the study visit as described by Mani and co-workers [15]. The photographs had been taken under standardized and reproducible conditions to maintain consistency. As such, a yardstick featuring a ruler and color palette was incorporated into the photographs to facilitate color and size calibration. The yardstick was positioned in a holder at the level of the base of the nose, as illustrated in Fig. 1. For the current study only frontal-view photographs were used in the evaluation process. To eliminate any potential bias stemming from viewing the full faces or surrounding facial features, each photo underwent additional cropping using Photoshop Elements software (Adobe Systems Incorporated, San Jose, CA, USA). This cropping focused on revealing an oval-shaped area, with only the lower and upper lip visible, including the entire length of the philtrum and the base of the nostrils, as seen in Fig. 2. Additionally, based on previous research findings indicating that left-sided clefts are ranked more favorable by human raters

compared to right-sided clefts, all images were presented as left-sided clefts [41]. In total, this yielded 140 images.

Web-based Scoring

The web-based scoring included 140 images featuring the UCLP patients ($n=74$) and the individuals without clefts ($n=61$), as well as ($n=5$) randomly selected duplicate photos. These images were uploaded randomly at each session onto a secure web-based platform, Artologik software (Artisan Global Media, Växjö, Sweden), hosted on the Uppsala University secure portal website.

Invited assessors (members of the different panels described below) were provided with secure, personalized access to the scoring session by email. Each session was open for a 7-day duration, during which assessors were asked to rate the aesthetics of the lip; for this evaluation, they employed the Visual Analogue Scale (VAS) at their convenience at any point they chose within the given time-frame. In addition to assessing lip aesthetics, assessors were asked whether they would like to change the appearance of the lips if given the opportunity, also reported on a VAS scale. The questions presented to the assessors were identical to those included in SWA and Body Cathexis scale questionnaires answered by the UCLP patients.

The assessors were encouraged to approach the evaluation as a “walk into the room” assessment. No further instructions or rules of conduct were given, but assessors were advised to conduct their assessment in a quiet, dark room. Before each formal rating session, assessors were given five practice images of people not part of the study to acquaint themselves with the materials and scoring procedure. For the formal rating session, each photograph was rated one at a time, with no option to go back and revise the assessment. Assessors were given the flexibility to pause and save their session and later resume within the established seven-day period. Following a 2-week interval, the images were re-randomized, and assessors were given access to a second scoring session. A duplicate of five randomly selected cases was included for each scoring session to evaluate intra-rater reliability.

After the second scoring session, two separate questions were posed to the assessors: (1) What methods or strategies did you use when assessing lip appearance? (2) What aspects of the images drew your primary attention? This aimed to collect insight into the different assessment methods used by assessors.

Panels

Three different panels were enrolled to evaluate the lip appearance of the study participants: a Professional panel,

a Layman panel, and a panel of people treated for CLP malformation with no affiliation to the current study (Cleft panel).

Professional panel

The Professional panel consisted of five highly experienced CLP surgeons with > 10 years of practice, each affiliated with different Swedish university hospitals other than Uppsala. None of them had any involvement in the treatment of the patients included in the study or the research. The Professionals were contacted through email and willingly agreed to participate by assessing the images via the online platform. The Professional panel consisted of two male and three female doctors. The mean age was 55 years, with a range spanning from 47 to 60 years.

Layman panel

The Layman panel comprised ten individuals recruited from acquaintances of the authors. They had no medical background or connection to the current study and were of diverse professional backgrounds. One of the ten initial participants did not complete the dual evaluation session and was thus excluded. Consequently, nine individuals completed both sessions. The mean age of the participants was 33 years, ranging from 25 to 49 years old. This group consisted of four men and five women.

Cleft panel

The Cleft panel consisted of people treated for CLP at UU hospital, with no direct affiliation to the current research. Ninety people born between 1960 and 2002 were randomly retrieved from the local patient list. The telephone number was not available, or no one answered when calling for the vast majority of these persons, leaving in the end 17 individuals whom were reached by phone and informed about the online photo assessment and asked to perform the scoring online. Of the 17 individuals communicated with, ten individuals initially agreed to contribute to the study and perform assessments. However, in the end, eight of these contributed of which five individuals completed both the scoring sessions. The Cleft panel’s participants were mean 37 years old, ranging from 26 to 48 years, with one man and four women.

Statistical analysis

The data was tested for normality distribution with the Shapiro-Wilk test and homogeneity of variances with Levene’s test. Continuous variables were presented as medians. To

minimize variability, assessors' scores within each panel were averaged. Comparative analysis of medians employed the appropriate statistical test. Mann-Whitney U-test was used to compare ratings by panels of UCLP patients and control. Friedman's test was used to compare ratings between different panels where, in case of statistical significance, a pairwise test with Bonferroni correction was done.

The intraclass correlation coefficient (ICC) was used to estimate intra-rater correlation for repeated rating of the duplicate images and across both scoring sessions. The inter-rater correlation was calculated between the rater inside the same panel group. The intraclass correlation coefficient type (2,1) two-way random with a single score and with absolute agreement was employed for the responses to the questions: "What do you think about the appearance of the lips?" and "Would you like to change the appearance of the lips?"

Individual ratings were used for all except for calculation between the panel groups, where we calculated the mean value of all raters within the panel groups, resulting in one rating for each individual and panel group.

To assess the strength of association/correlation between the different panel groups and self-assessment ratings, where the value used for the panel was the mean value of each rater within the panel. A non-parametric measure of rank correlation was used. For this, the Spearman rank correlation was used. The correlations were interpreted according to an accepted threshold, with values less than 0.19 indicating weak to no correlation, 0.20–0.39 weak correlation, 0.40–0.59 moderate correlation, and 0.60–0.79 strong

correlation. All analyses were made using the computer software for statistical analysis, IBM SPSS 24.0 (IBM corporation, Somers, NY).

Results

Demographics

The study participants' demographics are presented in Table 1.

Panel assessment of appearance: comparison between UCLP and control

Lip appearance of UCLP patients were rated as worse than that of the controls by all three panels; (Professional, Layman, and Cleft panel) ($p < 0.001$) (Fig. 3a). Additionally, the panels reported a greater desire to change the appearance of the lips of the UCLP patients compared to controls ($p < 0.001$) (Fig. 3b).

Comparison between the different panels' ratings

Professional panel consistently rated lip appearance, of both UCLP patients and the controls, as better (lower scores) compared to the Layman panel and the Cleft panel ($p < 0.001$) (Fig. 4a). The Cleft panel rated the second highest level of satisfaction, while the Layman panel rated the

Table 1 Number (n), age, and gender of patients, controls, and panels

	Patients with UCLP		Controls		Professional panel		Layman panel		Cleft panel	
	n		n		n		n		n	
Total number of patients	128				5		10			
Excluded	-19									
Invited to participate	109						10		8	
Declined participation	26									
Did not complete both sessions							1		3	
Participating individuals	83		67				9		5	
Incomplete photodocumentation	9		6							
Included individuals	74		61		5		9		5	
Right sided cleft, number	27									
Left sided cleft, number	47									
Average age (range), years	34 (20-47)		32 (23-58)		55 (47-60)		33 (25-49)		38 (26-48)	
Male, number	43		25		2		4		1	
Female, number	31		36		3		5		4	

UCLP= unilateral cleft lip and palate

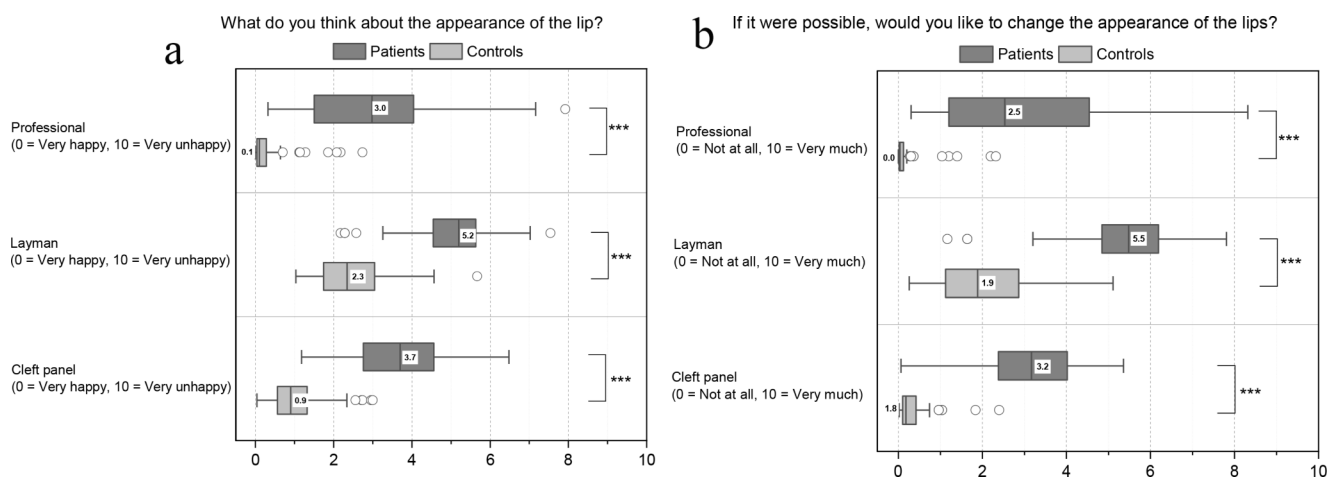


Fig. 3 (a-b) Comparison of all patients and controls. UCLP patients (n=74) and controls (n=61). Questions answered by VAS scale as box plot with median values. * = outlier. Mann Whitney U-test, *** p < 0.001

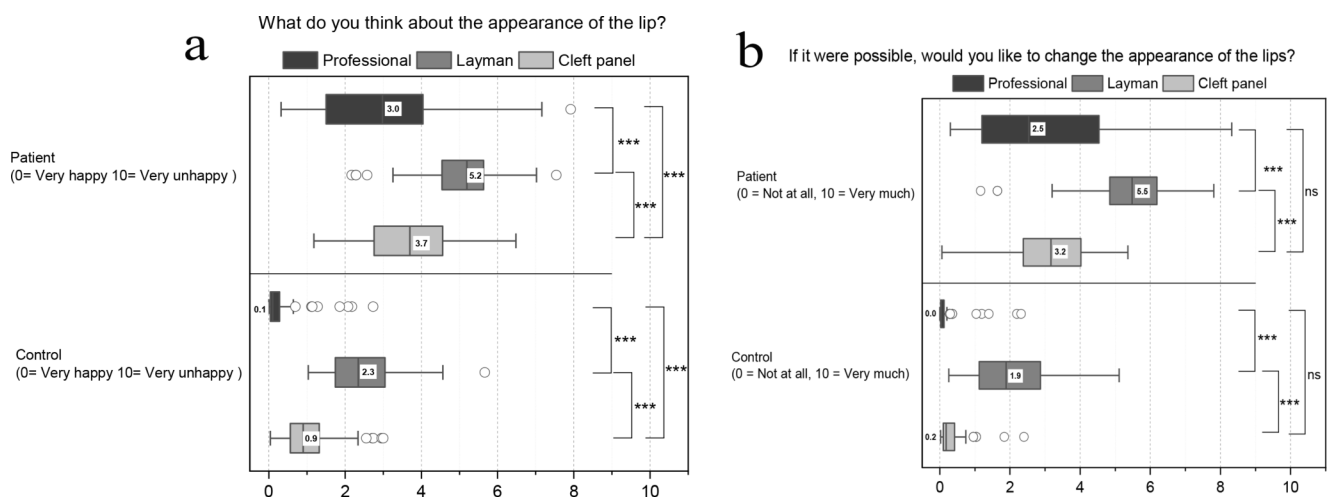


Fig. 4 (a-b) Comparison between panels. UCLP patients (n=74) and controls (n=61). Questions answered by VAS scale as box plot with median values. * = outlier. Friedmans' test, with Bonferroni correction. *** p < 0.001

lowest ($p < 0.001$) (Fig. 4a). Additionally, the Professional panel indicated the least interest to change the lip appearance for both UCLP patients and controls compared to the Layman and Cleft panel ($p < 0.001$) (Fig. 4b). The Layman panel indicated the highest desire to change the lip appearance ($p < 0.001$) (Fig. 4b).

Intra-rater agreement and inter-rater agreement

Intra-rater agreement was generally moderate for both lip appearance and the desire to change it, in both duplicate photographs and between sessions, within the Professional and Layman panels, particularly for UCLP patients, indicating consistent evaluations. Conversely, the Cleft panel showed poor agreement across these variables, reflecting greater variability in assessments, especially for UCLP patients. Agreement for controls varied across panels, with

some improvement noted in the second session. Table 2 shows the agreement of subjective assessment of lip appearance and the desire to change the lip appearance by using different panels for duplicate photos and between sessions (session one and session two).

Inter-rater agreement was generally poor across all panels. The Professional panel showed moderate agreement in the first session for both lip appearance and the desire to change it in UCLP patients, but this declined in the second session. The Layman and Cleft panels consistently exhibited poor agreement in both sessions, highlighting significant variability among raters. Controls also displayed poor inter-rater agreement across panels, with only minimal improvement in the second session. Table 3 shows the agreement in ratings between individuals in each panel group for UCLP patients and controls.

Table 2 Analysis of intra-rater correlation

Duplicate Image Evaluation Agreement				
	UCLP PATIENTS		CONTROLS	
	ICC. (95% CI)	p-value	ICC. (95% CI)	p-value
Session 1				
What do you think of the appearance of the lips?				
Professional panel	0.57 (0.20-0.80)	0.003*	0.99 (0.99-1.00)	<0.001*
Layman panel	0.63 (0.38-0.79)	<0.001**	0.28 (-0.33-0.76)	>0.05
Cleft panel	0.21 (-0.27-0.59)	>0.05	-0.13 (-1.08-0.79)	>0.05
If it were possible, would you like to change the appearance of the lips?				
Professional panel	0.82 (0.61-0.92)	<0.001**	0.00 (-0.81-0.81)	>0.05
Layman panel	0.73 (0.53-0.85)	<0.001**	0.22 (-0.51-0.75)	>0.05
Cleft panel	0.36 (-0.09-0.69)	>0.05	0.39 (-0.55-0.91)	>0.05
Session 2				
What do you think of the appearance of the lips?				
Professional panel	0.52 (0.14-0.78)	0.004*	0.60 (-0.31-0.95)	>0.05
Layman panel	0.64 (0.39-0.80)	<0.001**	0.53 (-0.14-0.87)	>0.05
Cleft panel	0.18 (-0.29-0.58)	>0.05	0.93 (0.44-0.99)	0.002*
If it were possible, would you like to change the appearance of the lips?				
Professional panel	0.53 (0.15-0.78)	0.005*	0.60 (-0.31-0.95)	>0.05
Layman panel	0.62 (0.38-0.79)	<0.001**	0.66 (-0.08-0.91)	<0.019*
Cleft panel	0.41 (0.01-0.71)	0.021*	0.96 (0.62-0.99)	0.002*
Between-Session Evaluation Agreement				
	UCLP PATIENTS		CONTROLS	
	ICC. (95% CI)	p-value	ICC. (95% CI)	p-value
What do you think of the appearance of the lips?				
Professional panel	0.49 (0.41-0.56)	<0.001**	0.22 (0.11-0.32)	<0.001**
Layman panel	0.63 (0.58-0.67)	<0.001**	0.65 (0.60-0.70)	<0.001**
Cleft panel	0.45 (0.26-0.58)	<0.001**	0.45 (0.35-0.53)	<0.001**
If it were possible, would you like to change the appearance of the lips?				
Professional panel	0.53 (0.45-0.60)	<0.001**	0.10 (-0.02-0.21)	0.048*
Layman panel	0.69 (0.63-0.73)	<0.001**	0.62 (0.56-0.67)	<0.001**
Cleft panel	0.48 (0.32-0.60)	<0.001**	0.34 (0.19-0.47)	<0.001**

- 1) ICC: Intra-rater correlation coefficient, based on individual ratings of duplicate images. This is done for the first (1st) and second (2nd) rating sessions for each panel.
- 2) ICC: Intra-rater correlation coefficient based on individual ratings between first and second rating sessions for each panel.
- 3) CI: Confidence interval
- 4) Significance marked in bold *p<0.05 **p<0.001

Table 3 Analysis of inter-rater correlation

	PATIENTS		CONTROLS	
	ICC. (95% CI)	p-value	ICC. (95% CI)	p-value
Session 1				
What do you think of the appearance of the lips?				
Professional panel	0.46 (0.32-0.59)	<0.001**	0.09 (0.01-0.20)	0.012*
Layman panel	0.07 (0.03-0.14)	<0.001**	0.08 (0.03-0.15)	<0.001**
Cleft panel	0.18 (0.07-0.30)	<0.001**	0.01 (-0.03-0.08)	0.312
If it were possible, would you like to change the appearance of the lips?				
Professional panel	0.53 (0.43-0.64)	<0.001**	0.07 (-0.02-0.18)	>0.05
Layman panel	0.08 (0.03-0.15)	<0.001**	0.14 (0.07-0.23)	<0.001**
Cleft panel	0.20 (0.08-0.33)	<0.001**	0.25 (0.14-0.38)	<0.001**
Session 2				
What do you think of the appearance of the lips?				
Professional panel	0.29 (0.12-0.47)	<0.001**	0.07 (-0.00-0.17)	0.029*
Layman panel	0.06 (0.02-0.11)	<0.001**	0.08 (0.03-0.15)	<0.001**
Cleft panel	0.09 (0.02-0.19)	<0.001**	0.14 (0.03-0.27)	<0.001**
If it were possible, would you like to change the appearance of the lips?				
Professional panel	0.29 (0.13-0.45)	<0.001**	0.09 (0.01-0.19)	0.015*
Layman panel	0.07 (0.02-0.13)	<0.001**	0.14 (0.07-0.23)	<0.001**
Cleft panel	0.10 (0.03-0.20)	<0.001**	0.26 (0.12-0.41)	<0.001**

ICC: Inter-rater correlation coefficient between members in each panel. Data reported for the first and second rating sessions separately.

CI: Confidence interval

Significance marked in bold * $p < 0.05$ ** $p < 0.001$

Correlations

All correlation analyses are presented in Table 4.

Panel assessment and self-assessment of lip appearance of UCLP patients

In the first session, the Cleft panel and Layman panel showed a moderate correlation rho (Spearman's rank correlation coefficient) ($r = 0.499$, $p < 0.001$). Similarly, the Cleft panel's correlation with the Professional panel was moderate ($r = 0.481$, $p < 0.001$). A weak yet significant correlation existed between Professional and Layman panels ($r = 0.304$, $p < 0.008$). However, neither panel showed any significant correlation with the UCLP patients' self-assessment. For the second session correlations were similar to the ones for the first session,

If it were possible, would you like to change the appearance of the lips?

In the first session, a strong correlation was seen between the Cleft and Layman panels ($r = 0.667$, $p < 0.001$) and a moderate correlation between Layman and Professional panels ($r = 0.499$, $p < 0.001$). Similarly, a moderate correlation was observed between Professional and Cleft panels ($r = 0.503$, $p < 0.001$). Similar correlations were revealed in the second session.

Panel assessment and self-assessment of lip appearance of controls

The Cleft panel showed a moderate correlation with both the Layman panel ($r = 0.378$, $p < 0.05$) and the Professional panel ($r = 0.329$, $p = 0.05$). The Layman panel and the Professional panel also demonstrated a moderate correlation ($r = 0.547$, $p < 0.001$). The correlations between the Cleft

Table 4 Correlation analysis between different panels and self-assessment

UCLP PATIENTS																	
Assessment session 1																	
What do you think about the appearance of the lips?																	
If it were possible, would you like to change the appearance of the lips?																	
	Self-assessment		Cleft Panel		Layman panel		Professional panel			Self-assessment		Cleft panel		Layman panel		Professional panel	
	rho	p	rho	p	rho	p	rho	p		rho	p	rho	p	rho	p	rho	p
Self-assessment	1.00	.	0.103	0.382	0.109	0.355	0.091	0.439		1.00	.	0.116	0.328	0.121	0.306	0.013	0.913
Cleft panel	0.103	0.382	1.00	.	0.499	<0.001**	0.481	<0.001**		0.116	0.328	1.00	.	0.667	<0.001**	0.503	<0.001**
Layman panel	0.109	0.355	0.499	<0.001**	1.00	.	0.304	0.008*		0.121	0.306	0.667	<0.001**	1.00	.	0.499	<0.001**
Professional panel	0.091	0.439	0.481	<0.001**	0.304	0.008*	1.00	.		0.013	0.913	0.503	<0.001**	0.499	<0.001**	1.00	.
Assessment session 2																	
What do you think about the appearance of the lips?																	
If it were possible, would you like to change the appearance of the lips?																	
	Self-assessment		Cleft panel		Layman panel		Professional panel			Self-assessment		Cleft panel		Layman panel		Professional panel	
	rho	p	rho	p	rho	p	rho	p		rho	p	rho	p	rho	p	rho	p
Self-assessment	1.00	.	0.146	0.214	0.004	0.972	0.054	0.650		1.00	.	0.135	0.254	0.107	0.369	0.059	0.621
Cleft panel	0.146	0.214	1.00	.	0.666	<0.001**	0.577	<0.001**		0.135	0.254	1.00	.	0.708	<0.001**	0.549	<0.001**
Layman panel	0.004	0.972	0.666	<0.001**	1.00	.	0.510	<0.001**		0.107	0.369	0.708	<0.001**	1.00	.	0.565	<0.001**
Professional panel	0.054	0.650	0.577	<0.001**	0.510	<0.001**	1.00	.		0.059	0.621	0.549	<0.001**	0.565	<0.001**	1.00	.
CONTROLS																	
Assessment session 1																	
What do you think about the appearance of the lips?																	
If it were possible, would you like to change the appearance of the lips?																	
	Self-assessment		Cleft panel		Layman panel		Professional panel			Self-assessment		Cleft panel		Layman panel		Professional panel	
	rho	p	rho	p	rho	p	rho	p		rho	p	rho	p	rho	p	rho	p
Self-assessment	1.00	.	0.119	0.368	0.096	0.468	0.086	0.518		1.00	.	0.053	0.687	0.070	0.598	0.206	0.115
Cleft panel	0.119	0.368	1.00	.	0.378	0.003*	0.329	0.010*		0.053	0.687	1.00	.	0.477	<0.001**	0.251	0.051
Layman panel	0.096	0.468	0.378	0.003*	1.00	.	0.547	<0.001**		0.070	0.598	0.477	<0.001**	1.00	.	0.271	0.034*
Professional panel	0.086	0.518	0.329	0.010*	0.547	<0.001**	1.00	.		0.206	0.115	0.251	0.051	0.271	0.034*	1.00	.
Assessment session 2																	
What do you think about the appearance of the lips?																	
If it were possible, would you like to change the appearance of the lips?																	
	Self-assessment		Cleft panel		Layman panel		Professional panel			Self-assessment		Cleft panel		Layman panel		Professional panel	
	rho	p	rho	p	rho	p	rho	p		rho	p	rho	p	rho	p	rho	p
Self-assessment	1.00	.	0.066	0.620	0.077	0.563	0.207	0.116		1.00	.	0.013	0.919	0.031	0.812	0.130	0.321
Cleft panel	0.066	0.620	1.00	.	0.627	<0.001**	0.649	<0.001**		0.13	0.919	1.00	.	0.482	<0.001**	0.263	0.040*
Layman panel	0.077	0.563	0.627	<0.001**	1.00	.	0.578	<0.001**		0.031	0.812	0.482	<0.001**	1.00	.	0.479	<0.001**
Professional panel	0.207	0.116	0.649	<0.001**	0.578	<0.001**	1.00	.		0.130	0.321	0.263	0.040*	0.479	<0.001**	1.00	.

Rho = Spearman's correlation coefficient. Significant correlation is marked in bold (*p < 0.05, **p < 0.001).

panel and both the Layman and the Professional panel were improved to strong correlations in the second session.

No correlation was seen between self-assessment and the panel ratings in any of the sessions.

If it were possible, would you like to change the appearance of the lips?

In the first session, a moderate correlation between the Cleft panel and the Layman panel was observed ($r=0.477, p<0.001$). The Layman and Professional panels showed a weak correlation ($r=0.271, p<0.05$). In the second session, the same trend was seen as in the first session.

Assessment strategies among panel members

Professionals reported to be focusing on symmetry, shape, and volume, with only one specialist not emphasizing these features. Other factors taken into account were the presence of scars (noted by two professionals, the natural look of the lips (mentioned by one), and comparison with what is considered normal lips (by one professional). Cleft panel members had varied approaches: two based their judgments on intuition or first impressions, one emphasized symmetry, another by evenness, and one followed the notion that “normal is beautiful”. Five out of nine Layman panel members focused on symmetry as a primary aspect. Shape and size were also noted (by two individuals), alongside comparisons to normal lips, the health of the lips, and the presence of abnormalities like injuries which were noted once each by different assessors in the Layman panel.

Discussion

The current study on long-term satisfaction of lip appearance in patients operated with Skoog’s lip repair as infants, found that all panels consistently gave higher ratings to the lip appearance of controls (non-cleft individuals) compared to subjects treated for UCLP. This aligns with prior research suggesting that individuals without congenital defects are generally thought to have more aesthetically pleasing faces [22]. Such perceptions are likely shaped by societal beauty norms and standards, which favor symmetry and typical facial features, as discussed by Bongaarts et al. [26]. Furthermore, panels showed a greater desire to improve UCLP patients’ lips than non-operated control. This may reflect the gap between ideal expectations and realistic surgical results and the importance of managing pre-surgery expectations [28].

It is important to capture the diverse perspectives of patients, healthcare professionals, and laymen, as aesthetic evaluation is inherently subjective and influenced by personal preferences and perceptions [17]. Our study highlighted this subjectivity, showing that different assessor groups provided varying ratings for lip appearance and desire to change lip appearance. The Professional Panel, chosen as they had no previous involvement in the care of the patients of the current study, were selected in order to hopefully give a unbiased clinical perspective. The Layman panel, picked from the general public, would represent the broader societal viewpoints. Unlike previous studies where CLP patients’ rated their own appearance, we recruited a separate Cleft panel to assess other UCLP patient’s photos, to reduce factors influenced by patient performing self-assessments [2, 42, 43]. Example of such factors, which

could affect a cleft panel assessment are: overly positive rating of the appearance due to a feeling of thankfulness towards the care providers [39] or underestimation of their appearance and higher desire for further corrections [2].

Contrasting views emerged, with the Professional panel rating lip appearance more positively than the Layman and Cleft panels, likely due to the Professional panel’s understanding of surgical outcomes and potentially more realistic expectations. The non-professional panels (Cleft and Layman panel), however, driven by social and psychological factors were more critical, expressing lower satisfaction with lip appearance and a higher desire to change it, with the Layman panel giving the lowest ratings. In the current study, the ratings from the first session were used to illustrate differences between panels as encounters typically happen between patients and non-experienced assessors (i.e. laymen) in daily life. This observation aligns with the existing literature, suggesting that professionals tend to judge facial aesthetic outcomes better than non-experienced panels [2, 17, 39, 44–46]. Yet, there are studies presenting opposing views, where non-experienced panels have rated appearance as better than professionals [42, 47] and where no difference in ratings was seen [21, 48]. While the cited studies mainly focus on nasolabial appearance, they contribute to understanding the subjective nature of facial aesthetics. In the second scoring session, Layman and Professional panels’ ratings improved slightly, whereas the assessment from the Cleft panel’s ratings was further worsened, reflecting a more significant variability. These differing perceptions, shaped by societal norms and media, underscore the importance of considering multiple perspectives when evaluating cleft lip surgery outcomes.

Limited literature exists on the influence of panels’ gender and age on the subjective assessment of CLP patients. There was an imbalance of gender distribution in the Cleft panel, where 4/5 members were females, which may have influenced the assessments. Females might be more critical in their assessment of beauty due to societal pressure on beauty standards [2, 49]. However, there are studies showing contrasting views where males have been more critical in their rating than females [50]. Additionally, studies also revealed no difference between male and female panellists on photographic assessments of CLP patients [45, 51]. Age differences between the panel groups could also have influenced the assessments, where the older panel, which in this case was the Professional group, was more forgiving in their assessment, and the youngest panel (Layman panel) was most negative in their assessment. Previous studies have shown that gender and age have no influence on subjective assessment among lay people [52, 53]. We conclude that the positive ratings by the Professional panel are more likely

due to their deeper understanding, experience, and realistic expectations rather than their age.

Upon analyzing the responses to our questions about the assessment strategies used in lip appearance assessments, we found that symmetry and shape were the most commonly noted aspects across all panels, with lip volume also being a relevant factor for the Professional panel. Both the Cleft and Layman panels similarly emphasized symmetry and shape while also observing deviations from the norm and the overall expression.

The rating procedure

The evaluation of photographs in cleft care is complex and involves several variables. While there are validated tools for assessing functional outcomes like Golson Yardstick for assessing dental arch relationship [54] and the Cleft Audit Protocol used for evaluation of assessment speech outcome [55], there is no, universally accepted method for evaluating lip aesthetics. Until now, only two other published studies have scored cleft aesthetic outcomes using cropped images showing lips separately [56, 57]. However, a widely adopted technique used for panel judgment of nasolabial area is detailed by Asher McDade et al. [19]. Where a single photograph displaying both the lip and nose is used, with the evaluator assigning separate scores for each feature within the same image. The Asher-McDade method has been employed in numerous studies and has established its credibility as a dependable tool for assessing the nasolabial region aesthetics [58–60]. Prior research suggests that the lips are the primary point of reference for symmetry before refining evaluation based on the nasal condition [57]. It has also been observed that the overall evaluation process is primarily influenced by the aesthetic appearance of the lips when both the lips and nose are scored together [56, 57]. Consequently, with reference to these studies, we chose to evaluate the lips separately in photos and focus on lip aesthetics avoiding biases linked to cleft and non-cleft facial features and to reduce the influence of unrelated facial features. This methodology was supported by earlier research recommendations [17, 19, 25, 26]. Some other studies have used full-view facial photographs, instead [44, 45, 61, 62]. However, in the absence of a universally accepted method for evaluating lip aesthetics, we chose to selectively crop the images, removing any distracting elements while highlighting the lips. We used circular cropping to avoid cropping-induced symmetry artifacts, as suggested in previous studies [25, 63].

In our study, we implemented a visual analog scale (VAS) over the traditionally used Likert-type scale proposed by Asher-McDade [19]. This approach allowed us to ask evaluators questions that mirror those in SWA and Body Cathexis

Scale, answered by UCLP patients and controls, thus simplifying the comparison between external assessments (Panels) and UCLP patients and controls self-evaluations. The visual analog scale is noted for its superior repeatability, objectivity, sensitivity, and reliability compared to fixed-category scales [64–67].

At the time of subject recruitment and questionnaire distribution, no questionnaires such as the CLEFT-Q was available, which today is regarded as a standard tool for reporting patients' reported outcomes in cleft care today [68].

Rater agreement

Intra-rater agreement

The results indicate varying consistency, with the Professional and Layman panels showing the most consistent ratings, while the Cleft panel was the least consistent.

The intra-rater agreement among individual raters across different sessions showed moderate agreement for UCLP patients in the professional group. The poor agreement observed for controls, particularly in the desire to change the lip appearance, highlights a significant variability in professional opinions over time when assessing normal or less affected appearances. These findings resonate with existing research that has scored lip appearance separately [56, 57]. In both of these studies, good intra-rater reliability scores were reported by professional on lip appearance. In the study conducted by Mosmuller et al. six assessors were involved and in the study by Deall et al. 17 cleft surgeons and 28 other cleft professionals. In contrast to the current study, Deal et al. used selected reference photographs which is described to give better reliability [57, 69].

In contrast to the Professional panel, the Cleft panel's poor intra-rater agreement for duplicate images in both sessions and across different sessions on UCLP patient's lips and desire to change lip appearance ratings could indicate a more personal and varied interpretation of cleft aesthetics, potentially shaped by personal experiences or emotional connections to cleft conditions. Interestingly, the almost perfect agreement in the second session for control lip appearance ratings suggests that, over time, individuals with cleft experience may reach a consensus on non-cleft aesthetics, possibly as they refine their judgment based on comparison to the cleft conditions.

High intra-rater consistency has also been reported in previous studies for the aesthetic of the upper lip of adult UCLP by all involved panels (professionals and laypeople) [45]. Similarly, Luyten et al. reported on panel assessment of smile aesthetics and showed good intra-rater agreement scores [70]. Unlike our study, the study by Gkantidis et al. included a smaller sample of 12 UCLP patients, the

photographs were not assessed twice by the panel groups, no duplicate photographs were assessed, and different statistical methods were used for the calculation of intra-panel reliability; all these factors could have contributed to higher intra-rater reliability [45]. The study by Lutyen et al. had a different study design, number of included patients, and the rating focused on smile aesthetics using smile aesthetic index involving four main categories with nasolabial aesthetics being one of them [70].

Inter-rater agreement

The inter-rater agreement observed in the Professional panel suggested a moderate agreement on evaluating lip appearance and the desire to change lip appearance. However, we saw a decline in agreement in the second session, indicating increased judgment variability. Control assessments were inconsistent throughout all sessions, demonstrating a challenge in evaluating lip appearance in nonaffected lips. This may also reflect the difficulties of judging subtle appearance differences and the subjective nature of aesthetic judgments. Similar to the current study, Mosmuller et al. saw the lowest inter-rater reliability for lip scoring [56]. Conversely, Deall et al. reported that lips were scored more reliably than the nose [57].

Both the Layman and Cleft panels had a poor inter-rater agreement in both sessions for UCLP patients and controls. Personal perspectives, cultural pressures, and lack of information on cleft diseases may explain the Layman panel's wide variety of viewpoints. The Cleft panel's diverse perceptions demonstrate the personalized nature of aesthetic evaluation. Cleft lip and palate aesthetic judgment are complex and multifaceted, as seen by the poor inter-rater agreement across panels. These findings support research showing that evaluators with varied experiences and attitudes struggle to agree [17, 56]. In contrast to the findings in the current study, previous studies on facial aesthetics have shown a high degree of agreement between specialists in the professional group [44]. However, it should be noted that the study sample was small, and the professional panel consisted of a professional panel of orthodontists and maxillofacial surgeons.

Correlation

The concordance observed between the Layman and Cleft panels and between both of these panels and the Professional panel may reflect shared expectations of aesthetic outcomes among those with personal cleft experiences, medical professionals, and the general public. Also, this could indicate a common perception of ideal outcomes in cleft surgery or common facial aesthetics. These findings are in line with

previous studies showing similar perspectives on appearance between different panels [71]. Interestingly, neither Professional, Cleft, nor Layman panels significantly correlated to UCLP patients, indicating a discrepancy in patients' perception of their post-surgical appearance compared to external evaluators. This suggests that patients' self-perceptions can differ greatly from external assessments. These findings correlate to earlier studies on nasal aesthetics, where no correlation was found between patient subjective measures and that of the panels [72]. Personal experiences, emotional responses, and challenges related to living with a cleft may influence patients' self-assessments. Two weeks later, correlations among all panels strengthened, showing that a greater consensus regarding lip appearance in UCLP patients was reached with time and consideration. This could be explained by customization to the criteria for assessment.

Conclusions

All panels' rated lip appearance of adults born with UCLP treated with Skoog's technique as worse than that of non-cleft controls. The study highlights a significant discrepancy between the Professional panel's more favorable assessment and the more critical views of the laymen and cleft panels, contradicting the initial hypothesis of uniform perception across panels. The lack of correlation between patient self-assessment and panel assessments underscores the subjective nature of lip satisfaction, emphasizing the need for personalized patient care strategies in CLP treatment outcomes.

Author contributions R. P. contributed to the planning of the study, data collection, data analysis, data interpretation, and manuscript preparation. A.F. contributed to the manuscript preparation. S.A. contributed to the data collection and preparation of photographs. M.H. contributed to the manuscript preparation. M.M. contributed to the planning of the study, data collection, data interpretation, and manuscript preparation. Statistical support has been given by the Region Västernorrland (FOU), Statistician PhD. Erling Englund.

Funding Open access funding provided by Uppsala University. The study has partially been funded by the local hospital research funding (ALF). There are no financial or other relationships that could be considered a conflict of interest.

Data availability The datasets are not publicly available. However, data can be made available by the corresponding author upon reasonable request, subject to approval by the institutional ethics board.

Declarations

Ethical approval The research was conducted in full accordance with the ethical principles outlined in the Declaration of Helsinki. The study was granted approval from the Regional Research Ethics Committee (Reference number 2005:245 (2019–02084) and (2019–04015)).

Patients consent All study subjects gave their written informed consent. Depicted patients provided additional consent to use their images.

Competing interests The authors have no competing interests to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Feragen KB, Borge AI, Rumsey N (2009) Social experience in 10-year-old children born with a cleft: exploring psychosocial resilience. *Cleft Palate Craniofac J* 46(1):65–74
2. Sinko K et al (2005) Evaluation of esthetic, functional, and quality-of-life outcome in adult cleft lip and palate patients. *Cleft Palate Craniofac J* 42(4):355–361
3. Al-Namankany A, Alhubaishi A (2018) Effects of cleft lip and palate on children's psychological health: a systematic review. *J Taibah Univ Med Sci* 13(4):311–318
4. Mani M et al (2013) Factors related to quality of life and satisfaction with nasal appearance in patients treated for unilateral cleft lip and palate. *Cleft Palate Craniofac J* 50(4):432–439
5. Stock NM, Feragen KB (2016) Psychological adjustment to cleft lip and/or palate: a narrative review of the literature. *Psychol Health* 31(7):777–813
6. Sinko K et al (2017) Facial aesthetics in young adults after cleft lip and palate treatment over five decades. *Sci Rep* 7(1):15864
7. Tennison CW (1946) *The repair of the unilateral cleft lip by the stencil method*. *Plast Reconstr Surg* 1952. 9(2): pp. 115–20
8. Skoog T (1958) A design for the repair of unilateral cleft lips. *Am J Surg* 95(2):223–226
9. Randall P (1959) A triangular flap operation for the primary repair of unilateral clefts of the lip. *Plast Reconstr Surg Transpl Bull* 23(4):331–347
10. Millard, Ralph DJ (1976) *Cleft Craft: The Evolution of Its Surgery Volume I: The Unilateral Deformity*
11. Fisher DM (2005) Unilateral cleft lip repair: an anatomical subunit approximation technique. *Plast Reconstr Surg* 116(1):61–71
12. van der Veer WM et al (2009) Potential cellular and molecular causes of hypertrophic scar formation. *Burns* 35(1):15–29
13. Ritter K, Trotman CA, Phillips C (2002) Validity of subjective evaluations for the assessment of lip scarring and impairment. *Cleft Palate Craniofac J* 39(6):587–596
14. Falk-Delgado A et al (2018) The Skoog lip repair for unilateral cleft lip deformity: the Uppsala experience. *Plast Reconstr Surg* 141(5):1226–1233
15. Mani MR, Semb G, Andlin-Sobocki A (2010) Nasolabial appearance in adults with repaired unilateral cleft lip and palate: relation between professional and lay rating and patients' satisfaction. *J Plast Surg Hand Surg* 44(4–5):191–198
16. Peroz R et al (2023) Patient reported Outcome following the Skoog Unilateral Cleft Lip Repair among Adults- a long-term cohort study and comparison to a Non-cleft Population. *Cleft Palate Craniofac J*, : p. 10556656231177139
17. Sharma VP et al (2012) Outcomes in facial aesthetics in cleft lip and palate surgery: a systematic review. *J Plast Reconstr Aesthet Surg* 65(9):1233–1245
18. Friede H, Lilja J, Johanson B (1980) Cleft lip and palate treatment with delayed closure of the hard palate. A preliminary report. *Scand J Plast Reconstr Surg* 14(1):49–53
19. Asher-McDade C et al (1991) Development of a method for rating nasolabial appearance in patients with clefts of the lip and palate. *Cleft Palate Craniofac J*, 28(4): p. 385–90; discussion 390–1.
20. Pigott RW, Pigott BB (2010) Quantitative measurement of symmetry from photographs following surgery for unilateral cleft lip and palate. *Cleft Palate Craniofac J* 47(4):363–367
21. Meyer-Marcotty P, Stellzig-Eisenhauer A (2009) Dentofacial self-perception and social perception of adults with unilateral cleft lip and palate. *J Orofac Orthop* 70(3):224–236
22. Al-Omari I et al (2003) An appraisal of three methods of rating facial deformity in patients with repaired complete unilateral cleft lip and palate. *Cleft Palate Craniofac J* 40(5):530–537
23. Marrant DG, Shaw WC (1996) Use of standardized video recordings to assess cleft surgery outcome. *Cleft Palate Craniofac J* 33(2):134–142
24. Trotman CA, Faraway JJ, Essick GK (2000) Three-dimensional nasolabial displacement during movement in repaired cleft lip and palate patients. *Plast Reconstr Surg* 105(4):1273–1283
25. Prah C et al (2006) Infant orthopedics and facial appearance: a randomized clinical trial (Dutchcleft). *Cleft Palate Craniofac J* 43(6):659–664
26. Bongaarts CA et al (2008) Effect of infant orthopedics on facial appearance of toddlers with complete unilateral cleft lip and palate (Dutchcleft). *Cleft Palate Craniofac J* 45(4):407–413
27. Mulder FJ et al (2018) The cleft aesthetic rating scale for 18-Year-old unilateral cleft lip and palate patients: a Tool for Nasolabial aesthetics Assessment. *Cleft Palate Craniofac J* 55(7):1006–1012
28. Goodacre TE et al (2004) Does repairing a cleft lip neonatally have any effect on the longer-term attractiveness of the repair? *Cleft Palate Craniofac J* 41(6):603–608
29. Mani M, Carlsson M, Marcusson A (2010) Quality of life varies with gender and age among adults treated for unilateral cleft lip and palate. *Cleft Palate Craniofac J* 47(5):491–498
30. Mani M et al (2010) EDITOR'S CHOICE Objective assessment of the nasal airway in unilateral cleft lip and palate—a long-term study. *Cleft Palate Craniofac J* 47 3 217–224
31. Veau V, Borel S (1931) *Division Palatine: anatomie, chirurgie, phonétique*. Masson
32. Wardill WEM (2005) The technique of operation for cleft palate. *Br J Surg* 25(97):117–130
33. Skoog T (1974) *Plastic surgery: New methods and refinements*. Thieme
34. Widmaier W (1961) [The primary elevated velum extension, a method of closing the cleft palate]. *Chirurg* 32:333–341
35. Perko MA (1979) Two-stage closure of cleft palate (progress report). *J Maxillofac Surg* 7(1):46–80
36. Friede H (2007) Maxillary growth controversies after two-stage palatal repair with delayed hard palate closure in unilateral cleft lip and palate patients: perspectives from literature and personal experience. *Cleft Palate Craniofac J* 44(2):129–136
37. Bates A, Emerson M, Spencer S (2004) Relationships between self-esteem, social experiences and satisfaction with appearance: standardisation and construct validation of two cleft audit measures
38. Secord PF, Jourard SM (1953) The appraisal of body-cathexis: body-cathexis and the self. *J Consult Psychol* 17(5):343–347

39. Marcusson A, Paulin G, Ostrup L (2002) Facial appearance in adults who had cleft lip and palate treated in childhood. *Scand J Plast Reconstr Surg Hand Surg* 36(1):16–23
40. Marcusson A, Akerlind I, Paulin G (2001) Quality of life in adults with repaired complete cleft lip and palate. *Cleft Palate Craniofac J* 38(4):379–385
41. Feragen KJ, Semb G, Magnussen S (1999) Asymmetry of left versus right unilateral cleft impairments: an experimental study of face perception. *Cleft Palate Craniofac J* 36(6):527–532
42. Eliason MJ, Hardin MA, Olin WH (1991) Factors that influence ratings of facial appearance for children with cleft lip and palate. *Cleft Palate Craniofac J* 28(2):190–193 193–4 discussion
43. Thomas PT et al (1997) Satisfaction with facial appearance among subjects affected by a cleft. *Cleft Palate Craniofac J* 34(3):226–231
44. Papamanou DA et al (2012) Appreciation of cleft lip and palate treatment outcome by professionals and laypeople. *Eur J Orthod* 34(5):553–560
45. Gkantidis N et al (2013) Aesthetic outcome of cleft lip and palate treatment. Perceptions of patients, families, and health professionals compared to the general public. *J Craniomaxillofac Surg* 41(7):e105–e110
46. Eichenberger M et al (2014) Facial attractiveness of patients with unilateral cleft lip and palate and of controls assessed by laypersons and professionals. *Eur J Orthod* 36(3):284–289
47. Foo P et al (2013) Facial aesthetics and perceived need for further treatment among adults with repaired cleft as assessed by cleft team professionals and laypersons. *Eur J Orthod* 35(3):341–346
48. Tobiasen JM, Hiebert JM, Boraz RA (1991) Development of scales of severity of facial cleft impairment. *Cleft Palate Craniofac J* 28(4):419–424
49. Kapp K (1979) Self concept of the cleft lip and or palate child. *Cleft Palate J* 16(2):171–176
50. Alhayek S et al (2019) Evaluation of facial appearance in patients with repaired cleft lip and palate: comparing the assessment of laypeople and healthcare professionals. *Maxillofac Plast Reconstr Surg* 41(1):5
51. Pausch NC et al (2011) Secondary cleft rhinoplasty rejuvenates the nose: a suggestion from a panel survey. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 111(2):174–180
52. Eliason MJ, Hardin MA, Olin WH (1991) Factors that Influence Ratings of Facial Appearance for Children with cleft lip and palate. *Cleft Palate Craniofac J* 28(2):190–194
53. Tobiasen JM (1987) Social judgments of facial deformity. *Cleft Palate J* 24(4):323–327
54. Mars M et al (1987) The Goslon Yardstick: a new system of assessing dental arch relationships in children with unilateral clefts of the lip and palate. *Cleft Palate J* 24(4):314–322
55. Sell D et al (2009) Cleft audit protocol for speech (CAPS-A): a comprehensive training package for speech analysis. *Int J Lang Commun Disord* 44(4):529–548
56. Mosmuller DG et al (2014) Comparison of two scoring systems in the assessment of nasolabial appearance in cleft lip and palate patients. *J Craniofac Surg* 25(4):1222–1225
57. Deall CE et al (2016) Facial aesthetic outcomes of cleft surgery: Assessment of Discrete Lip and nose images compared with Digital Symmetry Analysis. *Plast Reconstr Surg* 138(4):855–862
58. Fudalej P et al (2009) Nasolabial esthetics in children with complete unilateral cleft lip and palate after 1- versus 3-stage treatment protocols. *J Oral Maxillofac Surg* 67(8):1661–1666
59. Asher-McDade C et al (1992) 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* 29(5):409–412
60. Nollet PJ et al (2007) Nasolabial appearance in unilateral cleft lip, alveolus and palate: a comparison with Eurocleft. *J Craniomaxillofac Surg* 35(6–7):278–286
61. Zhu S, Jayaraman J, Khambay B (2016) Evaluation of facial appearance in patients with cleft lip and palate by Laypeople and professionals: a systematic literature review. *Cleft Palate Craniofac J* 53(2):187–196
62. Kocher K et al (2016) Judgment of Nasolabial Esthetics in Cleft Lip and Palate is not influenced by overall facial attractiveness. *Cleft Palate Craniofac J* 53(3):e45–52
63. Kuijpers-Jagtman AM et al (2009) Reference photographs for nasolabial appearance rating in unilateral cleft lip and palate. *J Craniofac Surg* 20(Suppl 2):1683–1686
64. Grossmann N et al (2005) Nasal airflow and olfactory function after the repair of cleft palate (with and without cleft lip). *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 100(5):539–544
65. Duncan JAL et al (2006) Visual analogue scale scoring and ranking: a suitable and sensitive method for assessing scar quality? *Plast Reconstr Surg* 118(4):909–918
66. Grant S et al (1999) A comparison of the reproducibility and the sensitivity to change of visual analogue scales, Borg scales, and Likert scales in normal subjects during submaximal exercise. *Chest* 116(5):1208–1217
67. Williamson A, Hoggart B (2005) Pain: a review of three commonly used pain rating scales. *J Clin Nurs* 14(7):798–804
68. Klassen AF et al (2018) 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. *CMAJ* 190(15):E455–E462
69. Mercado AM et al (2016) 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* 53(1):30–37
70. Luyten J et al (2023) Reliability of the novel cleft lip and palate smile esthetic index (CLP-SEI). *Eur J Orthod* 45(6):662–670
71. Meng T, Ma L, Wang Z (2015) Comparative Assessment of the Cleft Profile by patients with cleft lip and palate, Cleft surgeons, and Lay people. *Aesthetic Plast Surg* 39(5):757–763
72. Pitak-Arnnop P et al (2011) Panel and patient perceptions of nasal aesthetics after secondary cleft rhinoplasty with versus without columellar grafting. *J Craniomaxillofac Surg* 39(5):319–325

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.