

# Developing a visual rating chart for the esthetic outcome of unilateral cleft lip and palate repair

Access this article online

Website:

www.amsjournal.com

DOI:

10.4103/2231-0746.161060

Quick Response Code:



A. Olusanya Adeola, A. Akadiri Oladimeji<sup>1</sup>

Department of Oral and Maxillofacial Surgery, University College Hospital, Ibadan, Oyo State, <sup>1</sup>Faculty of Dentistry, University of Port Harcourt, Port Harcourt, Rivers State, Nigeria

Address for correspondence:

Dr. A. Olusanya Adeola, Department of Oral and Maxillofacial Surgery, University College Hospital, Ibadan, Oyo State, Nigeria.  
E-mail: defamyl@yahoo.com

## ABSTRACT

**Introduction:** Aesthetic impairment is a major concern for the cleft lip/palate patient. Thus, auditing of postsurgical esthetic outcome needs to be further explored as till date no universally accepted protocol exists. The study objective was to propose a new visual rating chart (VRC) for the aesthetic outcome of cleft lip and palate (CLP) surgery. **Materials and Methods:** In a retrospective review of 200 repaired clefts, the common esthetic deficiencies were identified, categorized and ranked in the order of severity. A chart of the illustrative diagram with textual description of the defects was produced and used as a basis for rating outcome by two groups of raters (familiar raters and recruited raters). Intra- and inter-raters reliability was estimated using Cohen's kappa statistics and intra-class correlation coefficient (ICC). Comparison between mean group coefficient was achieved with Kendall's correlation coefficient of concordance. **Results:** The intra- and inter-rater reliability for familiar raters was found to be strong with kappa values range of 0.80–0.87 ( $P < 0.001$ ). Similarly, inter-raters' reliability by recruited judges was very strong using ICC at both single (0.768) and average measures (0.982). **Conclusion:** The VRC is a reliable tool for assessing the esthetic outcome of CLP repairs.

**Keywords:** Cleft surgery, rating scale, surgical outcome

## INTRODUCTION

Cleft lip and palate (CLP) anomaly is the most common congenital craniofacial anomaly, with a global incidence range of 1:500–1:2703.<sup>[1,2]</sup> The associated aesthetic and functional impairments are major concerns both to the parents and to the individuals affected; especially, those who grow into self-consciousness with an uncorrected defect or a suboptimal repair. Various psychosocial dimensions are affected,<sup>[3-6]</sup> issues like societal rejection and stigmatization, low self-esteem and poor self-acceptance, challenges of peer integration, and functional derangements. These ultimately affect the overall health-related quality of life.<sup>[2,7]</sup> Only an esthetically and functionally acceptable outcome can relieve this huge burden.

However, while functional outcome has become more objectively assessable, acceptable aesthetic outcome is still relative.<sup>[8-10]</sup>

It easily plays upon the variable threshold of satisfaction, which differs between the patients, parents, medical experts and the lay observers in the community.<sup>[10-13]</sup> Hence, there is a continuing search for a standardized and universal tool for aesthetic outcome assessment; such tool as can enhance communication between the novice (i.e., patients, relations, and the general public) and the professionals (i.e., medical experts) thus permitting easy, objective and practical assessment of aesthetic outcome of the

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

**For reprints contact:** reprints@medknow.com

**Cite this article as:** Adeola AO, Oladimeji AA. Developing a visual rating chart for the esthetic outcome of unilateral cleft lip and palate repair. *Ann Maxillofac Surg* 2015;5:55-61.

surgical repair. Desirably, the tool should also be able to simplify documentation, thereby encouraging concision and precision of medical records for better inter-professional communication.

To this end, several instruments have been postulated; some are cumbersome, facility dependent and technically demanding while others are confounded by several intrinsic variables such as preoperative severity of the cleft, surgical technique of repairs, age and gender of the subjects.<sup>[14-19]</sup> We undertook a review of the literature reporting on assessment of aesthetic outcome of unilateral CLP (UCLP) repairs to highlight the current position of the art, and to seek justification for a new rating scale. We then proceeded to develop a new visual rating chart (VRC) for aesthetic outcome of the cleft lip, nose and palate which is simple, practical, easy to apply and comprehensible to both experts and lay judges.

### Review of literature: Current state of the art and justification for a new rating chart

Various authors have deliberated on the importance of outcome assessment in CLP management. While standard yardsticks and parameters are already well established for functional outcome assessment, the same cannot be said for aesthetic outcome of corrective CLP surgeries.<sup>[8-10]</sup> Looking at the literature, the approach to aesthetic outcome assessment has been broadly divided into quantitative and qualitative methods.<sup>[20]</sup> Quantitative methods analyze the extent of facial dysmorphology and disproportion through anthropometric measurements and compare to the noncleft side or the population average.<sup>[20-22]</sup> They are expressed in numerical data, which may be confusing to patients, parents and the lay public. On the other hand, qualitative methods are based on evaluation of overall appearance from a vivid image of the patients to which every observer can readily relate.<sup>[20,23,24]</sup>

Traditionally, quantitative assessments were based on measurements taken from two-dimensional (2D) photographs and radiographs, which were used to calculate the changes achieved after surgery.<sup>[17,25,26]</sup> Currently, various three-dimensional (3D) imaging techniques have been developed to overcome the shortcomings of conventional 2D imaging. These include 3D cephalometry,<sup>[27]</sup> Moire' topography,<sup>[28]</sup> 3D laser scanning,<sup>[29]</sup> 3D optoelectronic digitizers<sup>[30,31]</sup> and 3D stereophotogrammetry.<sup>[31,32]</sup> While these methods have the merits of being authentic and reproducible, their relevance is essentially professional. They are highly technical and facility dependent; therefore they may not be realistically integrated into the everyday evaluation of the cleft patient.

The more versatile methods of esthetic assessment have been the qualitative methods. Various yardsticks have been postulated. Emphasis has been prominently laid on the evaluation of the outcome of UCLP. While some authors focused on the nasal or labial component,<sup>[5,8,13,33]</sup> many others addressed the composite component of naso-labial symmetry.<sup>[17,24,26,34-36]</sup> In a recent systematic review of the outcomes in facial aesthetics in CLP surgery, Sharma *et al.*<sup>[37]</sup> shortlisted 53 articles published in the last 30 years, which were subjected to critical appraisal in line with Preferred Reporting Items for Systematic Reviews and Meta-analysis recommendations. They observed that

direct clinical assessment, clinical photograph evaluation, clinical videographic assessment and 3D evaluation were the common methods of outcome evaluations. In particular, cropped photographs were more noted to be representative than full face, and most techniques were based on a 5-point scale, evolving from the Asher-McDade system. They highlighted the limitations arising from assessments of subjects which were poorly matched for gender, age, preoperative cleft severity and ethnicity. Kuijpers-Jagtman *et al.*<sup>[38]</sup> rightly observed that though the Asher-McDade derived systems had been validly employed in many studies, it remains an absolutely abstract and highly subjective measure. It grades outcome on the 5-point Likert-scale ranging from poor to excellent outcome. They, therefore, conducted a study where they were able to derive reference photographs to describe the different category of assessment on the scale as a vivid image.

The assessment of the intraoral component of the repaired cleft, particularly where a cleft palate is involved has not been a focus in esthetic assessment. The nearest tool is the Great Ormond Street London and Oslo (GOSLON) method which is based on dental arch relationships in patients with repaired CLP.<sup>[39]</sup> This probably suggests an ideology that the intraoral site is not readily in the public domain and, therefore, constitutes no basis for esthetic psychological pressure on the cleft patient.

Many studies have compared esthetic outcome of cleft surgery based on patients' age, gender, surgical technique or preoperative severity of the cleft. While these considerations have a valid scientific basis and are of clinical importance to the surgeon, they bear little relevance to the patients, parents or the lay observer in the community. Since these groups also have a pertinent interest in the aesthetic achievement of a corrective cleft surgery, it is important to develop an outcome assessment tool that de-emphasizes those variables but focus on the residual deficit in repair.

Incidentally, most of the existing tools of assessments and indeed the more popular Asher-McDade system do not highlight the actual anatomical deficiency in each category of the scale rather, non-descriptive qualitative expression like very poor, poor or excellent are used. This method is therefore too subjective and extremely variable between individual assessors and may explain why some authors advocate that mean score of a number of judges rather than individual rating should be adopted.<sup>[3,37,40]</sup>

A more relevant rating scale should be based on realistic weighting of residual anatomic deficiency and the scoring should immediately communicate the type of structural deficit persisting and hint on the type and extent of further surgical correction that might be required. For these reasons a rationale has been provided for the development of a new rating scale, which could enhance better interaction and communication between, lay judges and the clinicians as well as facilitate practical communications between clinicians in terms of documentations and referrals.

In order to develop a new rating tool that captures the aforementioned ideology, our team adopted a scientific methodology, which is further described.

## MATERIALS AND METHODS

The materials for the study are postoperative cleft patients' frontal facial photographs drawn from the archive of the Cleft team. Each photograph had been taken using a standard protocol adopted by the team.
















- Patients are positioned against a white background
- The patient is made to look forward with the canthomeatal line parallel to the horizontal as much as possible
- Photographer stands about 2 feet from the patient (for a frontal view)
- Pictures were taken to capture both ears in equal measures as much as possible.

Photographs are taken using Canon PowerShot A480, 10.0 megapixels with  $\times 3.3$  optical zoom camera by a trained staff.

Two hundred previously repaired unilateral cleft lip and/or palates were reviewed to identify the common residual anatomical deformities associated with the lip, nose, and palate repairs. Those deficiencies were streamlined into 5 categories each for lip, nose and palate. Two photographs were selected by consensus between the authors to illustrate each of the categories for the three anatomical regions (i.e., lip, nose, and palate) and presented to a selected panel of judges comprising 10 adult cleft patients, 10 cleft child parents and 10 medical staff to independently adduce a weighted score of severity to each of the outcomes based on a Visual Analogue Scale (VAS - range: 0–10). The mean VAS score for each item was calculated and used to determine the order of severity of the esthetic outcome of those repairs.

From the 200 postoperative cases of cleft patients reviewed, the authors identified the common anatomical deficiencies in lip repairs as: (1) Mal-alignment of the white roll, (2) Vermilion notching at either or both of the white roll and red line, (3) Muscle bulk deficiency in the line of repair and/or nasal floor dehiscence, (4) Shortened and notched lip involving at least 1/3 of the lip length, with alveolar/tooth exposure but intact nasal floor. Common deficiency of nasal repairs were: (1) Alar base asymmetry; (2) Asymmetry of nasal dome; (3) Flared nostril on the cleft side; (4) Shortened/deviated columella. For the palate, the observations included: (1) Shrunken uvula; (2) Posterior fistula (distal to the incisive foramen); (3) Anterior fistula (proximal to the incisive foramen); (4) Partial/complete bifidization of the soft palate. An illustrative drawing of each of these outcomes was produced with a brief textual description to constitute the VRC [Figure 1]. These were combined into a single chart as five drawings each for the lip, nose and palate and a VRC; a nominal score of 1–5 was attached to each outcome for the different regions (lip, nose and palate). These scores though nominal were rank-ordered based on the severity weighting derived initially from the mean VAS assessment by the mixed panel of judges. Assessors were to indicate the score corresponding to the observed aesthetic deficiency. In cases where multiple deficits are present, the score with the worst aesthetic index was to be adduced, and a plus sign (+) added to indicate that other less severe deficiencies exist.

A set of 21 postoperative photographs of cleft patients was then randomly retrieved from the archive of the cleft team comprising

LIP REPAIRS	NASAL REPAIRS	PALATAL REPAIRS
<p><b>L I</b> - Lip is continuous in bulk, no notching at both vermillion borders i.e the white roll (WR) and the red line (RL).</p>  <p>Score = 1</p>	<p><b>N I</b> - Nostril is not flared, alar base symmetrically positioned and restoration of alar dome.</p>  <p>Score = 1</p>	<p><b>P I</b> - Complete palatal closure with a well formed 'U' shaped uvula.</p>  <p>Score = 1</p>
<p><b>L II</b> - Malalignment of WR, and/or differential lip bulk at surgical site.</p>  <p>Score = 2</p>	<p><b>N II</b> - Nostril not flared, dome restored, alar base asymmetry.</p>  <p>Score = 2</p>	<p><b>P II</b> - Complete palatal closure with clefting, disfigurement or shrinkage of the uvula.</p>  <p>Score = 2</p>
<p><b>L III</b> - Lip notched at either or both borders (i.e. WR and RL) without exposure of underlying teeth or gingiva.</p>  <p>Score = 3</p>	<p><b>N III</b> - Nostril not flared, dome not restored, <math>\pm</math> alar base asymmetry.</p>  <p>Score = 3</p>	<p><b>P III</b> - An isolated hard palatal dehiscence between the incisive foramen and vibrating line.</p>  <p>Score = 3</p>
<p><b>L IV</b> - Isolated lip/nasal floor dehiscence.</p>  <p>Score = 4</p>	<p><b>N IV</b> - Nostril is flared.</p>  <p>Score = 4</p>	<p><b>P IV</b> - An isolated hard palatal dehiscence anterior to the incisive foramen.</p>  <p>Score = 4</p>
<p><b>L V</b> - Notching of lip involving at least 1/3 of lip length but not up to the nasal floor.</p>  <p>Score = 5</p>	<p><b>N V</b> - Shortened columella</p>  <p>Score = 5</p>	<p><b>P V</b> - Dehiscence involving the soft palate.</p>  <p>Score = 5</p>

**Figure 1:** The visual rating chart - A diagrammatic description of the outcome of cleft lip and palate repairs

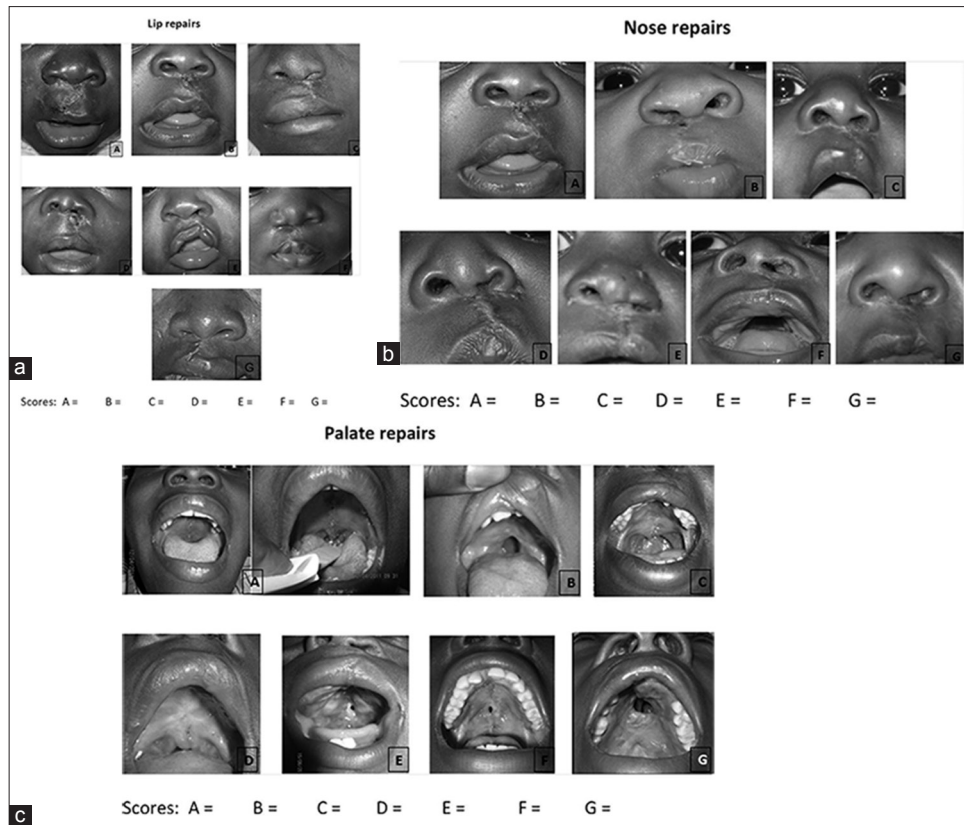
7 each for lip [Figure 2a], nose [Figure 2b] and palate [Figure 2c]. The selection was accomplished by putting the lip, nose and palate pictures into separate stacks after which 7 photographs were randomly (and blindly) selected from each stack. Using the VRC, the authors (subsequently referred to as familiar raters) independently assessed the aesthetic outcome of cases in the selected photographs. The assessment by familiar raters was repeated after a 2 week interval and intra- and inter-observer reliability was estimated.

The same set of photographs was later presented to a panel of 20 recruited judges comprising of members of the cleft teams in two Nigerian teaching hospitals. They were required to score the aesthetic outcome using the VRC. Inter-raters reliability was determined, and the mean correlation coefficient of the familiar raters as a group was later compared with that of the recruited raters as another group.

### Statistical analysis

In order to rank-order the various outcomes for each anatomical region, the mean VAS scores awarded by the mixed panel of judges for each outcome item were statistically correlated, and the spearman's rho coefficients were determined. The results determined the order of presentation of the various outcomes in the VRC.

An inter-rater reliability analysis using the Cohen's kappa statistic was performed to determine internal consistency between the two raters (familiar raters) who developed the VRC and the 95% confidence interval (CI) for the kappa values



**Figure 2:** (a) Postoperative photographs for the lip assessed by all judges (b) postoperative photographs for the nose assessed by all judges (c) postoperative photographs for the palate assessed by all judges

calculated from the formula: Estimate  $\pm$  1.96SE. To exclude the possibility of bias arising from the authors' familiarity with the chart, inter-observer reliability and internal consistency analysis of 20 recruited judges were estimated using intra-class correlation coefficient (ICC) and Cronbach's alpha test. Values were put at 95% CI. Comparison of ICC between the two groups of raters (familiar raters versus recruited raters) was achieved using Kendall's correlation coefficient of concordance.

## RESULTS

### Intra- and inter-observer reliability by familiar raters

The scores given to each item by each of the familiar raters at both the first and the second attempt are presented in Table 1. There is reasonable consistency for each rater and good agreement between raters. The statistical test of agreement is presented in Table 2. The intra-rater reliability for each of the two raters was found to be kappa = 0.87 ( $P < 0.001$ ), 95% CI (0.703, 1.08) and kappa = 0.80 ( $P < 0.001$ ), 95% CI (0.603, 1.01) respectively [Table 2]. Similarly, inter-rater reliability for each set of ratings by the two raters (i.e., initial and 2-week interval ratings) were kappa = 0.81 ( $P < 0.001$ ), 95% CI (0.608, 1.01) and kappa = 0.87 ( $P < 0.001$ ), 95% CI (0.703, 1.08) respectively [Table 2].

### Inter-observer reliability by recruited judges

Table 3 presents the various scores given to each item by the 20 recruited raters. Agreement among the raters assessed by

**Table 1: Tabulation of rating scores by familiar raters**

Item	Familiar rater 1		Familiar rater 2	
	1 <sup>st</sup> attempt	2 <sup>nd</sup> attempt	1 <sup>st</sup> attempt	2 <sup>nd</sup> attempt
Lip-A	1	1	1	1
Lip-B	1	1	1	1
Lip-C	2	2	2	2
Lip-D	4	4	4	4
Lip-E	5	5	5	5
Lip-F	3	3	3	3
Lip-G	2	2	2	2
Nose-A	1	1	1	1
Nose-B	3	2	2	3
Nose-C	1	1	1	3
Nose-D	4	3	3	3
Nose-E	4	4	4	4
Nose-F	4	4	4	4
Nose-G	2	2	3	2
Palate-A	1	1	1	1
Palate-B	5	5	5	5
Palate-C	2	2	2	2
Palate-D	2	2	2	2
Palate-E	3	3	3	3
Palate-F	3	3	3	3
Palate-G	4	4	4	4

intra-class coefficient (ICC) and Cronbach's alpha value showed that single measure (SM) ICC was 0.768, 95% CI (0.648, 0.878) and the average measure (AM) ICC was 0.982, 95% CI (0.969, 0.991) [Table 4] which agrees with the Cronbach's alpha value of 0.982.

**Table 2: Intra- and inter-raters' reliability by familiar raters**

	Value	Asymptotic SE	Approximate t	Approximate significant
<b>Intrarater reliability</b>				
Rater 1	0.873	0.085	7.606	0.000
Rater 2	0.804	0.103	7.024	0.000
<b>Interrater reliability</b>				
Rater 1 versus rater 2 (attempt 1)	0.810	0.101	7.053	0.000
Rater 1 versus rater 2 (attempt 2)	0.873	0.085	7.69	0.000

SE: Standard error

**Table 3: Rating scores by recruited raters**

Item	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Rater 6	Rater 7	Rater 8	Rater 9	Rater 10	Rater 11	Rater 12	Rater 13	Rater 14	Rater 15	Rater 16	Rater 17	Rater 18	Rater 19	Rater 20
Lip-A	1	1	2	1	2	2	1	4	2	2	3	1	1	3	1	2	1	2	3	1
Lip-B	2	1	2	3	1	2	1	4	1	2	3	1	1	1	1	2	2	3	1	1
Lip-C	1	2	1	2	2	1	1	2	2	2	1	2	2	3	2	2	3	2	3	2
Lip-D	4	4	2	3	4	3	4	4	3	2	2	4	4	4	4	4	4	2	4	4
Lip-E	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5	3	3	5	2	5
Lip-F	3	3	3	3	3	4	3	5	3	3	3	3	3	3	2	4	3	3	1	3
Lip-G	2	2	3	2	2	2	3	3	2	2	2	2	2	3	2	2	2	2	2	2
Nose-A	1	1	2	1	1	2	2	1	1	2	1	1	1	1	1	2	1	1	1	1
Nose-B	4	4	3	3	2	3	3	4	2	2	4	2	3	2	2	4	2	4	2	2
Nose-C	1	1	2	3	1	2	1	2	1	2	3	1	3	1	1	3	1	3	1	1
Nose-D	3	2	3	4	2	2	4	4	2	3	4	3	4	3	3	2	3	3	3	3
Nose-E	4	4	4	3	4	2	4	4	3	4	4	4	4	4	4	4	4	4	4	4
Nose-F	4	2	2	4	3	2	4	3	2	4	4	3	3	4	3	4	4	4	4	4
Nose-G	2	2	2	3	1	2	2	2	2	3	3	2	2	2	2	3	2	2	3	2
Palate-A	3	1	1	1	1	3	3	1	1	1	1	1	1	1	2	1	3	1	1	1
Palate-B	5	5	5	5	5	5	5	5	5	3	5	5	5	5	5	2	5	5	5	5
Palate-C	2	2	2	4	2	2	2	2	2	2	2	2	2	2	2	3	2	2	2	2
Palate-D	1	2	2	2	1	2	2	2	2	2	2	1	2	2	1	3	5	2	1	2
Palate-E	3	3	3	3	3	3	3	3	3	3	3	3	3	3	2	3	2	3	3	3
Palate-F	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Palate-G	4	4	4	3	3	3	5	4	4	4	4	4	4	4	4	4	4	4	4	3

**Table 4: Interraters' reliability by recruited judges**

	Coefficient	95% CI	F	df	Significant
Single measure correlation	0.708	Lower: 0.576 Upper: 0.841	56.845	16	0.000
Average measure correlation	0.982	Lower: 0.969 Upper: 0.992	56.845	16	0.000

Cronbach's alpha: 0.982. CI: Confidence interval

**Comparison of correlation coefficient between familiar judges and recruited judges**

The Kendall's correlation coefficient of concordance was calculated, which generated a W value of 0.0365 ( $P > 0.05$ ).

**Inter-rater reliability for separate components of the visual rating chart**

Based on the scores presented in Tables 1 and 3, agreement between raters for each component, that is, lip, nose, and palate was assessed. All raters inter-raters reliability as SM and AM respectively were estimated for each component of the VRC as follows: Lip: SM - 0.699, 95% CI (0.457, 0.935) and AM - 0.982, 95% CI (0.951, 0.997); Nose: SM - 0.717, 95% CI (0.486, 0.928) and AM - 0.973, 95% CI (0.930, 0.995); and Palate: SM - 0.875, 95% CI (0.730, 0.972) and AM - 0.992, 95% CI (0.979, 0.998).

**DISCUSSION**

One of the goals of the treatment of children with UCLP is to improve the esthetic appearance of structures affected by the cleft.<sup>[35]</sup> Unfortunately, repair of the UCLP rarely produces ideal facial aesthetics; patients invariably demonstrate some degree of deformation of the nose, upper lip, palatal vault and dental arches.<sup>[37,41]</sup> As a result, secondary surgery is often required to improve on initial outcome. Incidentally, there is no widely accepted standard rating method to assess facial aesthetics in CLP. While most methods of assessment have been based on clinical assessment by professionals, some authors have advocated for increasing consideration for patients' and parents' perception in deciding further surgical needs.<sup>[5,13,42]</sup> However, studies have shown that subjective perception of patients and parents may sometimes be biased by previous experience of the patients and parents' emotional feelings for the child.<sup>[11-13]</sup> Furthermore, differential satisfaction expression between individuals of different ethnic backgrounds has been reported.<sup>[43]</sup>

Various protocols for clinical assessment of esthetic outcome of cleft surgery, including the ones postulated by Asher-McDade et al.<sup>[40]</sup> Tobiasen et al.,<sup>[44]</sup> Tobiasen and Hiebert,<sup>[23]</sup> and Johnson and Sandy<sup>[45]</sup> have failed to address the extreme subjectivity and non-descriptive nature of the rating scales. Hence, there is still

a need for a tool that can provide uniform communication of the anatomical deficiency of concern between patients, parents, clinicians and even lay observers. This is the rationale for the VRC being proposed in this paper. The VRC makes use of illustrative diagrams with brief textual description to depict the anatomical deficiencies commonly observed after primary cleft corrective surgery. This preliminary experiment strongly supports the reliability of the new tool.

Prior to involving other raters in the experiments, the proponents, in a practice-rating task, used the VRC to assess outcomes of some repaired clefts including lip, nose and palatal outcomes. The intra- and inter-rater reliability based on Cohen's kappa statistics were very strong with kappa values ranging between 0.80 and 0.87 at 95% CIs. In the standard reporting of statistical tests of reliability, it is generally agreed that coefficient values between 0.7 and 1.0 confirms strong reliability. Because there is a chance of bias arising from familiarity of the authors with the chart, a second phase of the experiment was done in which twenty new and blinded raters were recruited and presented with the same set of photographs for assessment based on the VRC. The raters were recruited among medical experts from two teaching hospitals in Nigeria. No further explanations on the chart were given to them. Using intra-class coefficient of correlation and Cronbach's alpha which are standard statistical measures of rating reliability, the results confirm strong internal consistency for the VRC.

The SM coefficient is a measure of the reliability of rating by a single subject while the AM coefficient measures the mean of several judges' ratings of one item. The values were 0.7 and 0.9 respectively suggesting that either way, the chart is strongly reliable; though stronger with the pooled average of many raters. To further check whether familiarity with the rating chart will significantly influence rating, the mean correlation coefficient recorded by the familiar raters on four rating attempts was compared with the mean coefficient of one-time recruited raters, no statistically significant difference were observed. When the reliability of the individual component of the VRC was tested, using the scores of all raters, it was observed that the VRC components were strongly reliable for rating the lip, nose or palate outcome separately. The lip rating was the least strong with a coefficient of 0.699 at 95% CI in SM. The AM for the same component was however very strong with a coefficient of 0.982.

Based on these results, the VRC is now proposed as a new tool for judging the aesthetic outcome of cleft repairs. The strength of this tool is in its ability to immediately express the anatomical deficiency of concern, hence the type and perhaps, the extent of secondary surgery that may be required. The scores adduced are nominal, indicating only the deficit of concern. However, there is an additional value of the VRC, which derives in the fact that the items are rank-ordered in the reverse direction of severity. This implies that score 1 is the best outcome while score 5 is the worst outcome. The order was subjectively derived from the mean VAS score of a panel of judges. Thus, the VRC additionally provides a scale of measuring the aesthetic acceptability or severity of the outcome.

In a previous attempt to create a descriptive tool, Afifi *et al.*<sup>[46]</sup> proposed a visual scoring system - The Pennsylvania Lip and Nose Scoring System, for the postoperative assessment of cleft

lip repair. Their scoring system consists of two sections namely the nose and lip sections, which were subdivided into N1 to N3 and L1 to L3. These were rated as almost perfect with nearly imperceptible asymmetry at conversational distance, moderate deformities with some lip or nasal tip asymmetry at conversational distance and may require corrective surgical procedures and severe deformities with significant lip and or nose deformity that will require a complete revision. Their scoring system is simple but appears lacking in details therefore giving room for ambiguity in selecting cases into the different classes.

The Ascher-McDade system, which is by far most widely used, also does not provide details as to the actual anatomical problems. Kuijpers-Jagtman *et al.*<sup>[38]</sup> provided reference photographs to illustrate in a vivid image what each score on the Ascher-McDade depicts. This in our opinion will strengthen the Ascher-McDade scale, but the lack of descriptive explanation leaves each judge to compare mentally pictures. Also, we are of the opinion that using a particular human face as a reference standard may bias assessment, as judges may confuse similarity unless the outcome is exactly as observed on the standard photograph. Therefore, we adopted an illustrative drawing for the VRC, which immediately registers in the mind of judges that the image is a mere guide.

The shortcomings of the VRC in its present form include the use of a frontal projection to assess a 3D outcome. This may have to be improved upon by producing diagrams in at least 3D. However, this does not seem to have had a significant effect on the reliability of the tool in this study. We believe the descriptive explanation could have partly catered for this demerit. Another shortcoming is the fact that the components are essentially assessed in isolation, those features like nasolabial profile and nasal tip projections may have been underplayed. Also, persisting skeletal discrepancy associated with palatal repairs are not captured by the VRC. We think by combining the VRC with GOSLON yardstick for interarch relationships, a fairly comprehensive assessment of residual aesthetic/anatomic concern of the cleft patient can be made.

The advantages of one standard rating method for facial aesthetics in CLP are numerous including for comparisons between protocols, centers, and individuals and to determine the necessity for secondary surgery and type of surgery that may be required.<sup>[10]</sup> In addition, a standardized esthetic index could be helpful to inform patients about the expected treatment outcome from both primary and secondary repairs. Hence, it is worth all the efforts to derive such a useful tool.

## CONCLUSION

This study proved the VRC as a reliable and descriptive tool for assessing the aesthetic outcome of CLP repairs. In a future endeavor, the authors have proposed to validate the VRC for use by patients, parents and lay observers and to achieve a multicenter comparison of esthetic outcome of CLP surgery based on the VRC.

### Financial support and sponsorship

Nil.

### Conflicts of interest

There are no conflict of interest.

## REFERENCES

- Iregbulem LM. The incidence of cleft lip and palate in Nigeria. *Cleft Palate J* 1982;19:201-5.
- Sinko K, Jagsch R, Precht V, Watzinger F, Hollmann K, Baumann A. Evaluation of esthetic, functional, and quality-of-life outcome in adult cleft lip and palate patients. *Cleft Palate Craniofac J* 2005;42:355-61.
- Van Lierde KM, Dhaeseleer E, Luyten A, Van De Woestijne K, Vermeersch H, Roche N. Parent and child ratings of satisfaction with speech and facial appearance in Flemish pre-pubescent boys and girls with unilateral cleft lip and palate. *Int J Oral Maxillofac Surg* 2012;41:192-9.
- Fudalej P, Katsaros C, Bongaarts C, Dudkiewicz Z, Kuijpers-Jagtman AM. Nasolabial esthetics in children with complete unilateral cleft lip and palate after 1- versus 3-stage treatment protocols. *J Oral Maxillofac Surg* 2009;67:1661-6.
- Gkantidis N, Papamanou DA, Christou P, Topouzelis N. Aesthetic outcome of cleft lip and palate treatment. Perceptions of patients, families, and health professionals compared to the general public. *J Craniomaxillofac Surg* 2013;41:e105-10.
- Vlastos IM, Koudoumnakis E, Houlakis M, Nasika M, Griva M, Stylogianni E. Cleft lip and palate treatment of 530 children over a decade in a single centre. *Int J Pediatr Otorhinolaryngol* 2009;73:993-7.
- Marcusson A, Akerlind I, Paulin G. Quality of life in adults with repaired complete cleft lip and palate. *Cleft Palate Craniofac J* 2001;38:379-85.
- Timoney N, Smith G, Pigott RW. A 20 year audit of nose-tip symmetry in patients with unilateral cleft lip and palate. *Br J Plast Surg* 2001;54:294-8.
- Kim JB, Strike P, Cadier MC. A simple assessment method for auditing multi-centre unilateral cleft lip repairs. *J Plast Reconstr Aesthet Surg* 2011;64:195-200.
- Nollet PJ, Kuijpers-Jagtman AM, Chatzigianni A, Semb G, Shaw WC, Bronkhorst EM, *et al*. Nasolabial appearance in unilateral cleft lip, alveolus and palate: A comparison with Eurocleft. *J Craniomaxillofac Surg* 2007;35:278-86.
- Broder HL, Smith FB, Strauss RP. Effects of visible and invisible orofacial defects on self-perception and adjustment across developmental eras and gender. *Cleft Palate Craniofac J* 1994;31:429-36.
- Broder HL, Smith FB, Strauss RP. Habilitation of patients with clefts: Parent and child ratings of satisfaction with appearance and speech. *Cleft Palate Craniofac J* 1992;29:262-7.
- Pitak-Arnop P, Hemprich A, Dhanuthai K, Yildirim V, Pausch NC. Panel and patient perceptions of nasal aesthetics after secondary cleft rhinoplasty with versus without columellar grafting. *J Craniomaxillofac Surg* 2011;39:319-25.
- Ayoub A, Garrahy A, Millett D, Bowman A, Siebert JP, Miller J, *et al*. Three-dimensional assessment of early surgical outcome in repaired unilateral cleft lip and palate: Part 2. Lip changes. *Cleft Palate Craniofac J* 2011;48:578-83.
- Anastassov Y, Chipkov C. Analysis of nasal and labial deformities in cleft lip, alveolus and palate patients by a new rating scale: Preliminary report. *J Craniomaxillofac Surg* 2003;31:299-303.
- McKearney RM, Williams JV, Mercer NS. Quantitative computer-based assessment of lip symmetry following cleft lip repair. *Cleft Palate Craniofac J* 2013;50:138-43.
- van Loon B, Maal TJ, Plooij JM, Ingels KJ, Borstlap WA, Kuijpers-Jagtman AM, *et al*. 3D Stereophotogrammetric assessment of pre- and postoperative volumetric changes in the cleft lip and palate nose. *Int J Oral Maxillofac Surg* 2010;39:534-40.
- Lindsay WK, Farkas LG. The use of anthropometry in assessing the cleft-lip nose. *Plast Reconstr Surg* 1972;49:286-93.
- Williams HB. A method of assessing cleft lip repairs: Comparison of LeMesurier and Millard techniques. *Plast Reconstr Surg* 1968;41:103-7.
- He X, Shi B, Kamdar M, Zheng Q, Li S, Wang Y. Development of a method for rating nasal appearance after cleft lip repair. *J Plast Reconstr Aesthet Surg* 2009;62:1437-41.
- Whittle J. Preoperative anthropometric analysis of the cleft child's face: A comparison between groups. *Int J Surg* 2004;2:91-5.
- Farkas LG, Hajnis K, Posnick JC. Anthropometric and anthroposcopic findings of the nasal and facial region in cleft patients before and after primary lip and palate repair. *Cleft Palate Craniofac J* 1993;30:1-12.
- Tobiasen JM, Hiebert JM. Facial Impairment Scales for Clefts. *Plast Reconstr Surg* 1994;93:31-41.
- Asher-McDade C, Roberts C, Shaw WC, Gallager C. Development of a method for rating nasolabial appearance in patients with clefts of the lip and palate. *Cleft Palate Craniofac J* 1991;28:385-90.
- Quintero JC, Trosien A, Hatcher D, Kapila S. Craniofacial imaging in orthodontics: Historical perspective, current status, and future developments. *Angle Orthod* 1999;69:491-506.
- Kim SK, Cha BH, Lee KC, Park JM. Primary correction of unilateral cleft lip nasal deformity in Asian patients: Anthropometric evaluation. *Plast Reconstr Surg* 2004;114:1373-81.
- Swennen GR, Schutyser F, Barth EL, De Groeve P, De Mey A. A new method of 3-D cephalometry Part I: The anatomic Cartesian 3-D reference system. *J Craniofac Surg* 2006;17:314-25.
- Takasaki H. Moire topography. *Appl Opt* 1970;9:1467-72.
- Blais F. Review of 20 years of range sensor development. *J Electron Imaging* 2004;13:231-43.
- Honrado CB, Larrabee WF Jr. Update in three-dimensional imaging in facial plastic surgery. *Curr Opin Otolaryngol Head Neck Surg* 2004;12:327-31.
- Kau CH, Richmond S, Incrapera A, English J, Xia JJ. Three-dimensional surface acquisition systems for the study of facial morphology and their application to maxillofacial surgery. *Int J Med Robot* 2007;3:97-110.
- Ras F, Habets LL, van Ginkel FC, Prahl-Andersen B. Three-dimensional evaluation of facial asymmetry in cleft lip and palate. *Cleft Palate Craniofac J* 1994;31:116-21.
- Sándor GK, Ylikontiola LP. Patient evaluation of outcomes of external rhinoplasty for unilateral cleft lip and palate. *Int J Oral Maxillofac Surg* 2006;35:407-11.
- Vairaktaris E, Neukam FW, Nkenke E. Nasolabial symmetry following Tension-Randall lip repair: A three-dimensional approach in 10-year-old patients with unilateral clefts of lip, alveolus and palate. *J Craniomaxillofac Surg* 2006;34:253-62.
- Fudalej P, Katsaros C, Hozyasz K, Borstlap WA, Kuijpers-Jagtman AM. Nasolabial symmetry and aesthetics in children with complete unilateral cleft lip and palate. *Br J Oral Maxillofac Surg* 2012;50:621-5.
- Kane AA, Pilgram TK, Moshiri M, Marsh JL. Long-term outcome of cleft lip nasal reconstruction in childhood. *Plast Reconstr Surg* 2000;105:1600-8.
- 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;65:1233-45.
- Kuijpers-Jagtman AM, Nollet PJ, Semb G, Bronkhorst EM, Shaw WC, Katsaros C. Reference photographs for nasolabial appearance rating in unilateral cleft lip and palate. *J Craniofac Surg* 2009;20 Suppl 2:1683-6.
- Mars M, Plint DA, Houston WJ, Bergland O, Semb G. The Goslon Yardstick: A new system of assessing dental arch relationships in children with unilateral clefts of the lip and palate. *Cleft Palate J* 1987;24:314-22.
- Asher-McDade C, Brattström V, Dahl E, McWilliam J, Mølsted K, Plint DA, *et al*. 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;29:409-12.
- Landsberger P, Proff P, Dietze S, Hoffmann A, Kaduk W, Meyer FU, *et al*. Evaluation of patient satisfaction after therapy of unilateral clefts of lip, alveolus and palate. *J Craniomaxillofac Surg* 2006;34 Suppl 2:31-3.
- Reekie T. The effect of South Asian ethnicity on satisfaction with primary cleft lip and or palate repair. *J Plast Reconstr Aesthet Surg* 2011;64:189-94.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
- Tobiasen JM, Hiebert JM, Boraz RA. Development of scales of severity of facial cleft impairment. *Cleft Palate Craniofac J* 1991;28:419-24.
- Johnson N, Sandy J. An aesthetic index for evaluation of cleft repair. *Eur J Orthod* 2003;25:243-9.
- Affif A, Smith D, Deleyiannis FW, Jiang S, Mooney M, Vecchione L, *et al*. PLAN Score: Pennsylvania Lip and Nose Scoring System for Cleft Lip and Nose Deformities. Presented at the Annual Meeting American Cleft Palate-Craniofacial Association, Broomfield, Colorado, April 27, 2007 and the University of Pittsburgh Fourth Annual Plastic Surgery Resident Research Day, Pittsburgh, PA, May 30, 2007.