

Original Article

Rating nasolabial appearance on three-dimensional images in cleft lip and palate: a comparison with standard photographs

Adam Stebel*, Dries Desmedt**, Ewald Bronkhorst***,
Mette A Kuijpers** and Piotr S. Fudalej****,*****

*Division of Maxillofacial Surgery, Department of Stomatology and Maxillofacial Surgery, Comenius University, Bratislava, Slovakia, Departments of **Orthodontics and Craniofacial Biology, ***Community and Restorative Dentistry, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands, ****Department of Orthodontics, Palacky University, Olomouc, Czech Republic, and *****Department of Orthodontics and Dentofacial Orthopedics, University of Bern, Bern, Switzerland

Correspondence to: Piotr S. Fudalej, Department of Orthodontics and Dentofacial Orthopedics, University of Bern, Freiburgstrasse 7, 3010 Bern, Switzerland. E-mail: Piotr.Fudalej@zmk.unibe.ch

Summary

Background/Objective: Judgement of nasolabial aesthetics in cleft lip and palate (CLP) is a vital component of assessment of treatment outcome. It is usually performed based on two-dimensional (2D) facial photographs. An increasing use of three-dimensional (3D) imaging warrants an assessment if 3D images can substitute 2D photographs during aesthetic evaluation. The aim of this study was to compare reliability of rating nasolabial appearance on 3D images and standard 2D photographs in prepubertal children.

Methods: Forty subjects (age: 8.8–12) with unilateral CLP treated according to a standardized protocol, who had 2D and 3D facial images were selected. Eight lay raters assessed nasal form, nasal deviation, vermilion border, and nasolabial profile on cropped 2D and 3D images using a 100-mm visual analogue scale (VAS). Additionally, raters answer two questions: 1. Do 2D or 3D images provide more information on nasolabial aesthetics? and 2. Is aesthetic evaluation easier on 2D or 3D images?

Results: Intrarater agreement demonstrated a better reliability of ratings performed on 3D images than 2D images (correlation coefficients for 3D images ranged from 0.733 to 0.857; for 2D images from 0.151 to 0.611). The mean scores showed, however, no difference between 2D and 3D formats ($P > 0.05$). 3D images were regarded more informative than 2D images ($P = 0.001$) but probably more difficult to evaluate ($P = 0.06$).

Limitations: Basal view of the nose was not assessed.

Conclusions: 3D images seem better than 2D images for rating nasolabial aesthetics but raters should familiarize themselves with them prior to rating.

Introduction

Research has shown that physical attractiveness is related with stereotyping. Attractive persons are usually judged more positively than unattractive ones; they are perceived as having more social appeal, more interpersonal competence, and better adjusted than

unattractive individuals. For both adults and children, attractiveness is also strongly related to popularity and to success (1).

Physical attractiveness is influenced by many body characteristics (2) but facial appearance seems to be the most important component of physical attractiveness (3). The search for determinants of facial aesthetics has shown that an average face, i.e. how closely a face

resembles the majority of other faces within a population, symmetry, sexual dimorphism, and quality of skin, i.e. its health and colour, all influence how the face is perceived (4). In subjects with cleft lip and palate (CLP) facial symmetry and the health of skin in the nasolabial region are frequently compromised (5–10). For example, Bugaighis *et al.* (6) found with the use of three-dimensional (3D) imaging that the average CLP face had a wider, more asymmetric nose and mouth, flatter cheeks, and flatter zygomatic regions in comparison with an average non-cleft face. These residual deformities can lead to decreased facial attractiveness and finally to social handicap.

Nowadays, increasing number of cleft centres uses 3D imaging methods for treatment planning and evaluation of their results. The most popular of them—stereophotogrammetry and laser scanning—produce highly accurate and reproducible images, which can be used for documentations and research (11). Three-dimensional technology has also been used for assessment of facial appearance in CLP (12,13). It is unclear if aesthetic outcomes evaluated on two-dimensional (2D) facial photographs, currently the standard stimulus medium, and 3D images are comparable. The only study to date, in which 3D and 2D images were compared as stimulus media (14), demonstrated that 2D colour transparencies and 3D images were equivalent only for assessment of some regions of the face, primarily nose, and midface. Unfortunately, the disagreement between raters during assessment of the upper lip was significant. However, the sample of Al-Omari *et al.* (14) was heterogeneous regarding age and treatment methods. The inclusion of subjects aged 10–30 years treated by various surgeons with different techniques may have increased a diversity of outcomes and it might have led to difficulties in scoring. In addition, a recommendation resulting from the Eurocleft (15) and Americleft (16) studies is that documentation and records should be taken at certain time points and the first most complete records ought to be obtained at age 9–10. Standardized timing of record taking would facilitate research and clinical audit. Therefore, the aim of this study was to compare reliability of rating nasolabial aesthetics on 3D images and standard 2D photographs in subjects aged from 8 to 12 years.

Materials and methods

Ethics

This study respected the Declaration of Helsinki with regard to research in human subjects. A written statement of the Institutional Review Board was obtained stating that this study does not fall within the remit of the Medical Research Involving Human Subjects Act (WMO) because the use of anonymized data gathered during routine patient care is in accordance with Dutch law on medical research. Therefore, this investigation could be carried out without an individual approval by an accredited research ethics committee.

Subjects

The files of the Cleft Palate Craniofacial Unit at the Radboud University Medical Center, Nijmegen, The Netherlands were searched to identify all patients meeting the following inclusion criteria: 1. diagnosis of nonsyndromic unilateral cleft lip and palate (UCLP), ascertained by the clinical geneticist of the team, 2. available standard 2D photographs of the face and 3D stereophotogrammetric images of the face taken at approximately 10 years of age. 2D and 3D images must have been taken no later than within 1 month (between each other).

The group consisted of 40 subjects with a mean age of 10 years (range: 8.8–12 years). All patients were treated in accordance with the standardized protocol used at the Unit—lip closure (Millard

cheiloplasty) was done at age 6–8 months together with a primary nose correction (McComb); soft palate closure was performed at the age of 12–14 months. All surgical procedures were performed by two surgeons. At the time of facial imaging, the included subjects did not have their alveolar bone grafting procedure yet; also, hard palate was not closed.

Three-dimensional photographs of all patients were taken with the same stereophotogrammetrical camera set-up (3dMD face System; 3dMD LLC, Atlanta, Georgia, USA) under standardized conditions. Patients were positioned in the natural head position and asked to keep their eyes open and to relax their facial musculature. All images were taken by an experienced photographer. The photogrammetric system was calibrated every morning.

Assessment of nasolabial aesthetics

Eight junior postgraduate students (four men and four women; age range: 25–31 years) from the Department of Orthodontics and Craniofacial Biology at Radboud UMC, experienced in the use of 3D facial images and their manipulation, who were not familiar with treatment of the cleft deformity were asked to rate nasolabial aesthetics with a visual analogue scale (VAS). The VAS was a 100-mm line with a description ‘least aesthetic’ on the left end (0 mm) and ‘most aesthetic’ on the right end (100 mm). Distance of each rater’s marking from the start of the scale (0 mm—‘least aesthetic’) was measured with a digital caliper and ratings were transformed into continuous variables. Four elements of nasolabial morphology—nasal form, nasal deviation, vermilion border, and nasolabial profile—were assessed on cropped 2D and 3D images (Figures 1 and 2) on the laptops of raters. The images were cropped only after alignment of an interpupillary line as a horizontal line. The 2D images were loaded into PowerPoint for rating, whereas the 3D images were assessed in a 3D viewer (3D-Tool, version 9, 3D-Tool, Weinheim, Germany). The raters could manipulate 3D images in all directions. They scored 2D and 3D images alternately, in a series of 10, i.e. 10 2D images, 10 3D images, 10 2D images, and so on. The order of 2D and 3D nasolabial images was random. There was no time limit for scoring. Prior to rating raters were shown patients’ images to familiarize with a range of aesthetic outcomes in the group (i.e. from most to least aesthetic). Subsequently, raters were told to use the fullest possible spectrum of the VAS scale.

In addition to aesthetic rating, the rating panel was asked to use the VAS scale to answer two questions: 1. Do 2D or 3D images provide more information on nasolabial aesthetics? and 2. Is aesthetic

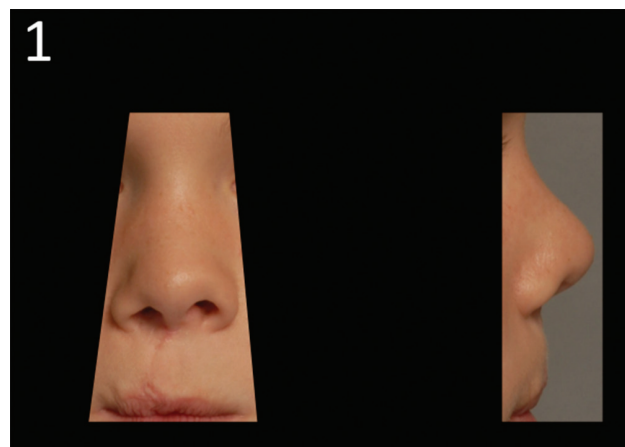


Figure 1. Two-dimensional image of nasolabial area with a random number used for rating.



Figure 2. Three-dimensional image of nasolabial area used for rating.

evaluation easier on 2D or 3D images? If the rater put a mark in the middle of the VAS (corresponding with 50 mm), it meant that neither of the methods was more informative or easier for assessment. Placing the mark closer to the left end of the scale (0 mm) meant that 3D images are more informative and easier to evaluate.

Statistical analysis

In order to assess intrarater reliability of aesthetic ratings, 40 random images (20 2D and 20 3D) were rerated by the panel within 3 weeks. Intrarater reliability was assessed by calculating correlation coefficient, duplicate measurement error (DME), the difference between paired observations and running *t*-tests for paired observations. Inter-rater reliability was evaluated by calculating Cronbach's alpha coefficient.

The relationship between assessments on 2D and 3D images was evaluated by calculating correlation coefficients, the differences between observations and by running *t* tests. The replies to the two additional questions were analysed with one-sample *t* test.

Results

The consistency among raters during the assessment of all four elements of nasolabial morphology was high for both 2D and 3D images (Table 1). The Cronbach's alpha ranged from 0.755 (nasal deviation, 2D) to 0.94 (profile, 2D). Overall rating nasal deviation was relatively most challenging to raters (Cronbach's alpha was lowest for 2D and 3D images), whereas evaluation of the profile view produced most consistent scores.

The assessment of intrarater agreement demonstrated a better reliability of ratings performed on 3D images than 2D images. Table 2 shows that correlation coefficients between duplicate ratings were higher for 3D images than for 2D images. For example, correlation coefficients for 3D images ranged from 0.733 (nasal deviation) to 0.857 (vermillion border) and they ranged for 2D images from 0.151 (nasal deviation) to 0.611 (vermillion border). Moreover,

the DME's for nasolabial components were smaller for 3D images (range from 4.78 to 6.09) than for 2D images (range from 8.63 to 12.96). However, duplicate ratings of three of the four nasolabial components demonstrated statistically significant differences for 3D images. No statistically significant differences were noted for 2D images.

The mean rater scores showed no difference between 2D and 3D formats. The correlation coefficients between ratings on 2D and 3D images ranged from 0.58 (nasal deviation) to 0.767 (profile). The mean differences between ratings on 2D and 3D images were small (2.106 or less) and statistically not significant (Table 3). However, individual scores (See online supplementary table 1) showed that the behaviour of the raters was quite heterogeneous. For example, rater 4 (R4) tended to favour 3D structurally (scores assigned on 3D images were higher than on 2D images for all nasolabial components), while rater 7 (R7) did the opposite (scores assigned on 3D images were lower than on 2D images for all nasolabial components).

Replying to the first question, i.e. which of the two types of images (2D or 3D) are more informative regarding nasolabial appearance, the rating panel pointed out to 3D images as providing significantly more information than 2D images ($P = 0.001$). The answer to the second question, i.e. on which of the images it was easier to evaluate aesthetics, showed a tendency toward considering 2D images as a format easier for aesthetic evaluation ($P = 0.06$)—Table 4.

Discussion

The number of publications listed in PubMed on 3D imaging in CLP patients increases. For example, in years 2006–2009, there were 36 publications with 'cleft' and various forms of '3D' in the title, whereas in the following 4 years the number increased to 70. One can assume that this trend will continue, particularly due to decreasing costs of 3D imaging systems, and therefore it may be expected that the majority of cleft centres will utilize this technology to assess their treatment results, including aesthetics of nasolabial area.

Our findings show that one can confidently use 3D images for rating nasolabial appearance in CLP. We found no difference in the mean scores for any element of nasolabial morphology and correlation coefficients between 2D and 3D scores were high, perhaps with the exception of assessment of nasal deviation for which the coefficient was 0.58. The difficulty to rate nasal deviation might have been related to a restricted view of the midface (both 2D and 3D images were cropped to show only nasolabial area), which might have hindered an establishment of midsagittal reference plane. Moreover, higher intrarater reproducibility for 3D images in comparison to 2D images implies that the former stimulus medium (3D) should be preferred over the latter (2D). The current findings partially disagree with results of Al-Omari *et al.* (14), the only study which directly compared aesthetic assessment made on 2D and 3D formats (13). Al-Omari *et al.* compared three stimulus media used in assessment of facial aesthetics in CLP—clinical, 2D colour transparencies and 3D stereophotographs. They also compared the scores assigned by panels comprising professionals and laypersons. Their main findings were that professional raters showed good reproducibility across the three stimulus media, whereas laypersons demonstrated good reproducibility using 2D and 3D media [the mean kappas ranged from 0.47 (3D, lip) to 0.61 (2D, nose)] but exhibited poor reproducibility for clinical assessment. Furthermore, the 2D and 3D images were equivalent as stimulus medium for rating aesthetics of the nose but not for assessment of the upper lip (the mean kappas for 2D versus 3D evaluation of the nose were 0.60 and 0.47 for professional and

Table 1. Values of Cronbach's alpha for rating of four nasolabial components on two-dimensional (2D) and three-dimensional (3D) images.

Cronbach's alpha	Nasal shape	Nasal deviation	Vermillion border	Profile view
2D	0.869	0.755	0.868	0.940
3D	0.855	0.826	0.869	0.889

Table 2. Intra-rater error during rating of four nasolabial components on two-dimensional (2D) and three-dimensional (3D) images. DME, duplicate measurement error; CI, confidence interval.

	Correlation coefficient	DME	Difference	P value	95% CI
Nasal shape 2D	0.359	10.42	5.45	0.115	[-1.45 to 12.35]
Nasal shape 3D	0.762	6.09	-0.94	0.493	[-3.7 to 1.81]
Nasal deviation 2D	0.151	8.63	7.64	0.011	[1.93 to 13.35]
Nasal deviation 3D	0.733	5.49	-4.43	<0.001	[-6.91 to -1.94]
Vermillion border 2D	0.611	9.58	-0.54	0.86	[-6.89 to 5.8]
Vermillion border 3D	0.857	4.78	2.43	0.028	[0.27 to 4.59]
Profile view 2D	0.305	12.96	1.34	0.747	[-7.24 to 9.92]
Profile view 3D	0.828	5.24	-2.56	0.035	[-4.93 to -0.2]

Table 3. Comparison of mean aesthetic scores for four nasolabial components assessed on two-dimensional (2D) and three-dimensional (3D) images. CI, confidence interval.

	Correlation	Difference between 2D and 3D	P	95% CI of difference
Nasal shape	0.699	-2.106	0.173	[-5.176 to 0.963]
Nasal deviation	0.58	-1.438	0.285	[-4.122 to 1.247]
Vermillion border	0.728	1.416	0.34	[-1.548 to 4.38]
Profile	0.767	-0.081	0.96	[-3.338 to 3.175]

Table 4. Summary of answers of raters to question 1 (Q1, Do 2D or 3D images provide more information on nasolabial aesthetics?) and question 2 (Q2, Is aesthetic evaluation easier on 2D or 3D images?).

	Mean difference*	P value	95% CI
Q1 (included information)	-21.13	0.001	[-30.68 to -11.57]
Q2 (ease of rating)	13.63	0.060	[-0.74 to 27.99]

*Mean difference between averaged score and 50 (the middle of the rating scale corresponding with the answer that both methods are equally informative or equally easy to rate). A minus score implies that 3D images were judged more informative or easier to rate.

lay raters, respectively; the mean kappas for 2D versus 3D evaluation of the upper lip were 0.18 and 0.28 for professional and lay raters, respectively). In contrast, we found that rating aesthetics on 2D images was associated with lower reproducibility than rating on 3D images. There may be several reasons or disagreement between Al-Omari *et al.* and us. First, Al-Omari *et al.* used a five-point aesthetic index developed by Asher-McDade *et al.* (17) whereas we used the VAS scale. A comparison of these two rating scales showed that aesthetic index might be associated with lower reliability than the VAS (18). Secondly, Al-Omari *et al.* assessed the appearance of the whole nose and the upper lip separately, whereas we evaluated aesthetics of nasolabial area as recommended by Asher-McDade *et al.* (17). Particularly, the assessment of the upper lip was different between our studies—Al-Omari *et al.* evaluated the size, smoothness, and visibility of the cleft scar in addition to assessment of the continuity of the vermilion border. In this investigation, we rated only the vermilion border and disregarded appearance of the scar

tissue due to possible problems with good visualization of lip scarring (19). Third, raters participating in the present study were able to manipulate 3D images in all directions and without any time restriction. In contrast, the 3D images shown to Al-Omari *et al.* raters were turned only in horizontal and vertical directions. Fourth, we assessed nasolabial aesthetics in pre-pubertal children treated with a standardized protocol whereas Al-Omari *et al.* assessed the group heterogeneous regarding the age (10–30 years) and treatment methods. Finally, our raters, although not familiar with treatment of CLP, were using 3D images for orthodontic treatment planning for several months prior to rating. It is likely that familiarity with 3D images improved intrarater reproducibility.

We mentioned before that the use of five-point ordinal scale for rating aesthetics in CLP might not be optimal. In our study comparing three various methods of rating nasolabial appearance—five-point scale, VAS, and reference scores—we found largest variability of scores when five-point scale was used (18). The scale with reference photographs was overall most reliable and the reliability of VAS was intermediate. In conclusion of that study we recommended 'the use of reference photographs along with the VAS or numerical (from 0 to 200) semi-continuous scale' to obtain most reliable results. Initially, we planned to use reference images also in this study. However, due to relatively limited number of 8–12-year-old patients with CLP who had 3D stereophotographs of the face available we felt that we would not have been able to select a representative 3D reference image separately for rating boys and girls. As a result, the only modification we made to improve reliability of rating was the use of the 100 mm VAS scale instead of 5 grades originally proposed by Asher-McDade *et al.* (17).

A gold standard in rating nasolabial aesthetics is clinical evaluation. Unfortunately, it is challenging to organize a clinical rating session with participation of several judges. As a result clinical assessments

cannot readily be used in audits or research. Several studies showed, however, that clinical ratings and ratings performed on 2D facial photographs as stimulus medium were comparable (20, 21). Becker et al. (20) found that results from two methods agreed closely—mean coefficient of correlation between methods was 0.73 and mean kappa value was 0.72. Johnson and Sandy (21) found, in turn, that agreement between methods ranged from moderate to good and there was no significant systematic bias. Consequently, facial photographs can be regarded as a ‘second best’ stimulus medium and a comparison of reliability of ratings on 2D photographs and 3D images is justified.

In this investigation, the rating panel consisted of lay judges. Previous studies showed that professional raters, i.e. persons involved in treatment of CLP, might judge nasolabial aesthetics differently than lay raters, i.e. persons neither involved in treatment nor familiar with the cleft deformity (22, 23). As a result there is a controversy regarding ideal panel composition—lay panel versus professional panel—for assessment of aesthetic outcome in CLP. On one hand, professionals see patients with CLP daily and have better appreciation than laymen of the range of possible outcomes in CLP. Also, clinicians usually make a final recommendation to a patient or his/her parent(s) as for the future course of therapy. For example, they may recommend or argue against the revision surgery. On the other, the well-being of the patient is influenced by his/her social environment comprising in the majority of situations persons not familiar with CLP. Their perception of the residual cleft deformity may, therefore, decide about how successful is adjustment of a child with CLP. It seems that arguments both in favour of using professional panels and in favour of using lay panels are equally important and the choice of the type of a rating panel might depend on a research question investigated or practical aspects such as ease of organization of rating session.

A limitation of this study could be that we did not include a basal view, i.e. a view at the nostrils, for aesthetic evaluation. The basal view probably is the best way to assess symmetry of the nostrils. Unfortunately, the nostrils are frequently distorted on 3D images and consequently difficult to judge. On the other hand this view is rarely shown in social circumstances and its appearance may not be important from a patient’s perspective.

In conclusion, our findings seem to favour the use of 3D images above 2D images for rating nasolabial appearance. This is likely so because 3D views of nasolabial area can be rotated to observe the structure more precisely from diverse angles. Potential raters should, however, familiarize themselves with 3D images before rating.

Supplementary material

Supplementary material is available at *European Journal of Orthodontics* online.

References

- Langlois, J.H., Kalakanis, L., Rubenstein, A.J., Larson, A., Hallam, M. and Smoot, M. (2000) Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, 126, 390–423.
- Weeden, J. and Sabini, J. (2005) Physical attractiveness and health in Western societies: a review. *Psychological Bulletin*, 131, 635–653.
- Currie, T.E. and Little, A.C. (2009) The relative importance of the face and body in judgments of human physical attractiveness. *Evolution and Human Behavior*, 30, 409–416.
- Little, A.C., Jones, B.C. and DeBruine, L.M. (2011) Facial attractiveness: evolutionary based research. *Philosophical Transactions of the Royal Society London B: Biological Sciences*, 366, 1638–1659.
- Ayoub, A., Bell, A., Simmons, D., Bowman, A., Brown, D., Lo, T.W. and Xiao, Y. (2011) 3D assessment of lip scarring and residual dysmorphology following surgical repair of cleft lip and palate: a preliminary study. *Cleft Palate Craniofacial Journal*, 48, 379–387.
- Bugaighis, I., Tiddeman, B., Mattick, C.R. and Hobson, R. (2012) 3D comparison of average faces in subjects with oral clefts. *European Journal of Orthodontics*, 36, 365–372.
- Kau, C.H., Medina, L., English, J.D., Xia, J., Gateno, J. and Teichgraber, J. (2011) A comparison between landmark and surface shape measurements in a sample of cleft lip and palate patients after secondary alveolar bone grafting. *Orthodontics (Chic)*, 12, 188–195.
- Ferrario, V.F., Sforza, C., Dellavia, C., Tartaglia, G.M., Sozzi, D. and Caru, A. (2003) A quantitative three-dimensional assessment of abnormal variations in facial soft tissues of adult patients with cleft lip and palate. *Cleft Palate Craniofacial Journal*, 40, 544–549.
- Stauber, I., Vairaktaris, E., Holst, A., Schuster, M., Hirschfelder, U., Neukam, F.W. and Nkenke, E. (2008) Three-dimensional analysis of facial symmetry in cleft lip and palate patients using optical surface data. *Journal of Orofacial Orthopedics*, 69, 268–282.
- Krimmel, M., Kluba, S., Breidt, M., Bacher, M., Dietz, K., Buelthoff, H. and Reinert, S. (2009) Three-dimensional assessment of facial development in children with Pierre Robin sequence. *Journal of Craniofacial Surgery*, 20, 2055–2060.
- Kuijpers, M.A., Chiu, Y.T., Nada, R.M., Carels, C.E. and Fudalej, P.S. (2014) Three-dimensional imaging methods for quantitative analysis of facial soft tissues and skeletal morphology in patients with orofacial clefts: a systematic review. *PLoS One*, 9, e93442.
- Al-Omari, I., Millett, D.T. and Ayoub, A.F. (2005) Methods of assessment of cleft-related facial deformity: a review. *Cleft Palate Craniofacial Journal*, 42, 145–156.
- Mosmuller, D.G., Griot, J.P., Bijnen, C.L. and Niessen, F.B. (2013) Scoring systems of cleft-related facial deformities: a review of literature. *Cleft Palate Craniofacial Journal*, 50, 286–296.
- Al-Omari, I., Millett, D.T., Ayoub, A., Bock, M., Ray, A., Dunaway, D. and Crampin, L. (2003) An appraisal of three methods of rating facial deformity in patients with repaired complete unilateral cleft lip and palate. *Cleft Palate Craniofacial Journal*, 40, 530–537.
- Shaw, W.C., Semb, G., Nelson, P., Brattström, V., Mølsted, K., Prahlandersen, B. and Gundlach, K.K. (2001) The Eurocleft Project 1996–2000. Overview. *Journal of Craniomaxillofacial Surgery*, 29, 131–140.
- Long, R.E., Jr., Hathaway, R., Daskalogiannakis, J., Mercado, A., Russell, K., Cohen, M., Semb, G. and Shaw, W. (2011) The Americleft study: an inter-center study of treatment outcomes for patients with unilateral cleft lip and palate part 1. Principles and study design. *Cleft Palate Craniofacial Journal*, 48, 239–243.
- Asher-McDade, C., Roberts, C., Shaw, W.C. and Gallager, C. (1991) Development of a method for rating nasolabial appearance in patients with clefts of the lip and palate. *Cleft Palate Craniofacial Journal*, 28, 385–390.
- Fudalej, S.A., Desmedt, D., Bronkhorst, E. and Fudalej, P.S. (2015) Comparison of three methods of rating nasolabial appearance in cleft lip and palate. *Cleft Palate Craniofacial Journal* (under revision).
- Asher-McDade, C., Brattstrom, V., Dahl, E., McWilliam, J., Molsted, K., Plint, D.A., Prahlandersen, B., Semb, G., Shaw, W.C. and The, R.P. (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 Craniofacial Journal*, 29, 409–412.
- Becker, M., Svensson, H. and Jacobsson, S. (1998) Clinical examination compared with morphometry of digital photographs for evaluation of repaired cleft lips. *Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery*, 32, 301–306.
- Johnson, N. and Sandy, J. (2003) An aesthetic index for evaluation of cleft repair. *European Journal of Orthodontics*, 25, 243–249.
- Papamanou, D.A., Gkantidis, N., Topouzelis, N. and Christou, P. (2012) Appreciation of cleft lip and palate treatment outcome by professionals and laypeople. *European Journal of Orthodontics*, 34, 553–560.
- Gkantidis, N., Papamanou, D.A., Christou, P. and Topouzelis, N. (2013) Aesthetic outcome of cleft lip and palate treatment. Perceptions of patients, families, and health professionals compared to the general public. *Journal of Craniomaxillofacial Surgery*, 41, e105–e110.