Otolaryngology–Head and Neck Surgery (2009) 140, 76-81

ORIGINAL RESEARCH–FACIAL PLASTIC AND RECONSTRUCTIVE SURGERY

Assessment of the reproducibility of facial expressions with 3-D stereophotogrammetry

Adam R. Sawyer, BSc, DOHNS, MRCS, MSc,
Marlene See, MBBS, MRCS, MSc, and **Charles Nduka, MD, FRCS,**
West Sussex, England, United Kingdom

OBJECTIVE: There is little research into analyzing whether facial expression changes are secondary to actual treatment or normal day-to-day variation. This study aims to ascertain whether nine-facial expressions were reproducible.

SUBJECTS AND METHODS: Thirty-nine white subjects performed nine-facial expressions, captured by three-dimensional stereophotogrammetry in three sessions. After initial capture (session 1), each expression was repeated after 15 minutes (session 2) and then 4 weeks (session 3) after the initial session. Statistical analysis was performed on the mean variability of facial landmarks between session 1 and 2 and session 1 and 3.

RESULTS: Repose was the most reproducible expression. The least reproducible was "blow-out-the-cheeks." Analysis between session 1 and 2 showed no significant differences in expression reproducibility. Analysis between session 1 and 3 showed significant differences for the "smile-with-lips-open" and "blow-out-the-cheeks" expressions.

CONCLUSION: Facial expressions are reproducible in a 15-minute period. There are significant differences in the ability to reproduce facial expressions 4 weeks apart for "smile-with-lips-open" and "blow-out-the-cheeks."

Physicians have tried to devise a simple, user-friendly, quantitative system for the assessment of facial paralysis and subsequent recovery in Bell's palsy and other facial neuromuscular disorders. Grading facial function is necessary for evaluating and communicating response course and results of medical and surgical treatments of facial palsy.[1,2] Many clinicians use subjective scoring systems such as the House-Brackmann Facial Grading System (HBFGS) scale or Sunnybrook Facial Grading System (SFGS),[3] as there is not an agreed on objective evaluation system. The advantage of objective facial nerve evaluation methods is that this will reduce errors by avoiding observer bias inherent to subjective grading systems.

Several investigators have examined facial movement,[4,5] but few have looked at the reproducibility of facial expressions. To determine the degree of abnormality from facial measurements, it is important to know what degree of variability is normal. There has been little research to determine whether facial expression changes are secondary to treatment and how much is due to normal day-to-day and month-to-month variation.

The evaluation of facial expressions and their reproducibility has suffered from the lack of an adequate method of recording and analysis. Previously, authors have used photographs with the disadvantage of assessing a three-dimensional object by two-dimensional means and thus give no account of anterior-posterior changes.[6] Regardless of the techniques used, previous investigations of facial expressions have used small numbers of patients,[7] or have not divulged variables such as race[7] and sex.[8] In other research, actual expressions performed and instructions used to evoke the expressions are not fully described.[8]

Burres et al[9] studied the facial movements of 30 subjects with integrated electromyography but the subject was required to maintain expressions that maximum intensity for 30 seconds. This is likely to lead to patient fatigue and therefore a loss of accuracy. Wood et al[10] investigated 11 normal subjects and analyzed the variability of facial movement with video microscope but only evaluated day-to-day variability with two facial expressions. Johnston et al,[11] in a well-designed study, investigated facial expression reproducibility more comprehensively with stereophotogrammetry and found this method to be accurate at looking at facial landmarks. However, this study only looked at five expressions in the lower half of the face only.

To assess facial expressions accurately we have used a three-dimensional stereophotogrammetry camera system that is the latest tool for evaluating facial muscle movements. It is a rapid, noninvasive means of quantitatively recording facial expressions and has been found to be very satisfactory in the assessment of the face.[12]

To the best of our knowledge, the reproducibility of facial expressions has not been objectively measured on all

Received May 14, 2008; revised July 15, 2008; accepted September 4, 2008.

0194-5998/$36.00
doi:10.1016/j.otohns.2008.09.007

areas of the face to allow assessment of all motor branches of the facial nerve. The purpose of the study was to ascertain whether nine facial expressions are reproducible in normal white subjects with the use of three-dimensional stereophotogrammetry. The second part of the study was to investigate whether there is any sexual dimorphism in the reproducibility of facial expression.

METHODS

Local ethics committee approval was obtained before commencement of the study. All subjects had full dentition, normal maxillary/mandibular relationships, and no previous history of any facial nerve paralysis or cleft lip or palate. Written and verbal information was given to all subjects and written consent was obtained before commencing the study.

Thirty-nine white volunteers (21 male, 18 female) average age 33 years (range, 21 to 36 years) were asked to perform a series of facial expressions. These expressions were captured with a 3-D stereophotogrammetry camera system. Each subject was imaged with the VECTRA-3D 2 dual module system (Canfield Scientific, Inc, Fairfield, NJ), which uses stereophotogrammetry (Fig 1). The VECTRA system we used was supplied by Surface Imaging International Ltd, UK.

The 3-D stereophotogrammetry system integrates two pods, each with three cameras; on either side two monochrome cameras are synchronized to capture images illuminated with integral projectors. The camera system required calibrating each day before capturing facial data. The 3-D facial model that is generated can be analyzed with VAM (visualization, analysis, measurement) version 2.8.2 (Canfield Scientific Inc) application software. As it is a digital facial model, one is able to rotate, pan, or zoom into the images, as well as view multiple surfaces simultaneously to facilitate analysis.

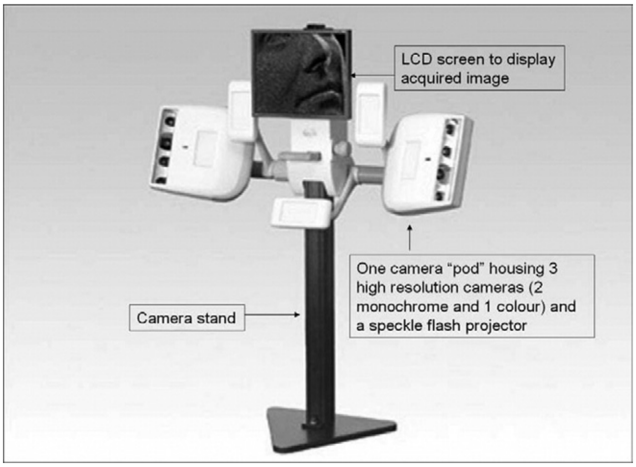

Figure 1 Stereophotogrammetry system VECTRA-3D dual module system.

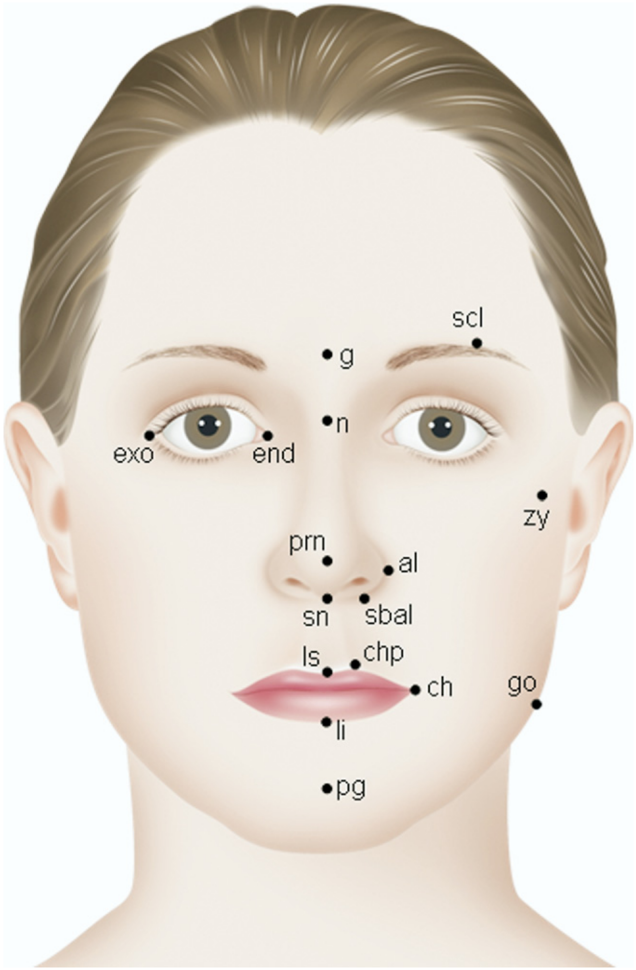

Figure 2 Landmarks used in this study: nasion (n), glabella (g), right and left endocanthion (enR and enL), pronasale (prn), right and left exocanthion (exoR and exoL), labiale inferioris (li), labiale superius (ls), alare (alR and alL), subnasale (sn), right and left subalare (sbalR and sbalL), right and left superciliare (sclR and sclL), right and left christa philtri (chpR and chpL), pogonion (pg), right and left zygion (zyR and zyL), right and left cheilion (chR and chL), and right and left gonion (goR and goL).

The sequence of facial expressions was practiced at least twice by the subjects before collection of the data. Images were able to be captured within 2 milliseconds. All of their recordings were completed and images analyzed by a single investigator. Each landmark was positioned according to those described by Farkas.[13] The 25 landmarks that were used are given in Figure 2.

The subjects were asked to perform maximal or widest facial expressions in response to both spoken and visual cue cards provided by the first author in the following order: repose (resting facial position); raised eyebrows; close eyes as in sleep; close eyes as tight as possible; smile as wide as possible with lips closed; smile as widely as possible with lips open; purse your lips; show lower teeth (platysma contraction); blow out the cheeks.

To investigate the reproducibility of facial expressions, each subject had each facial expression captured by 3-D

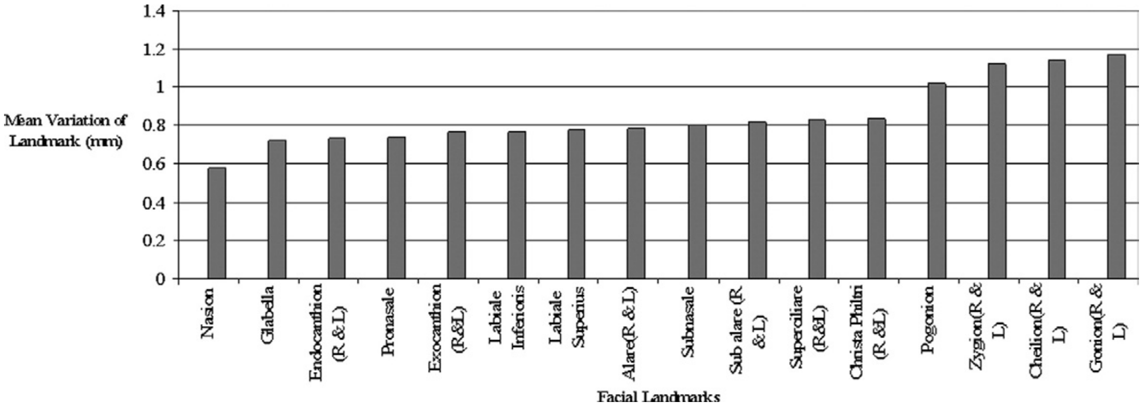

Figure 3 Mean variation in landmark position reproducibility for all subjects, sessions, and expressions.

stereophotogrammetry three times at each session. Session 1, initial capture; session 2, 15 minutes after initial capture; session 3, 1 month after initial capture.

A comparison was made between the expressions in session 1 and 2 to investigate intrasessional reproducibility. A further comparison was made between session 1 and 3 to investigate intersessional reproducibility. Statistical analysis was performed with analysis of variance (ANOVA) on mean discrepancy of facial landmark movement. A significance level of $P = 0.05$ was set.

RESULTS

All subjects were able to complete all facial expressions except for two who were unable to perform "show me your lower teeth" (platysma contraction).

The mean variation in landmark position for all subjects over all sessions is shown in Figure 3. The x-axis shows each of the facial landmarks and these have been ranked in increasing order of their variability. This shows that the least variable facial landmark is the nasion and the most variable facial landmark is the gonion.

The mean variation to each individual expression between the first and second session is shown in Figure 4. This graph shows that "repose" is the most reproducible facial expression followed by "eyes as in sleep," "raised eyebrows," "eyes tight," "pursed lip," "show me your lower teeth" (platysma contraction), "smile with lip closed," "smile with lips open." The least reproducible expression was "blow out the cheeks." There was little difference in facial expression reproducibility between session 1 and session 2 to all expressions; however, ANOVA on the mean discrepancies between the "smile with lips open" and "blow out the cheeks" expressions approached significance ($P = 0.08$ and $P = 0.07$).

The mean variability for each individual expression between the first and third sessions is shown in Figure 5. Again "repose" is the most reproducible facial expression. There was little difference in facial expression reproducibility between session 1 and session 3 to all expressions. However, ANOVA on the mean discrepancies between sessions 1 and 3 for the "smile with lips open" and "blow out the cheeks" expressions was significantly different ($P = 0.03$ and $P = 0.03$, respectively).

Figure 6 shows the mean variability that compares males and females with all facial landmarks for each expression.

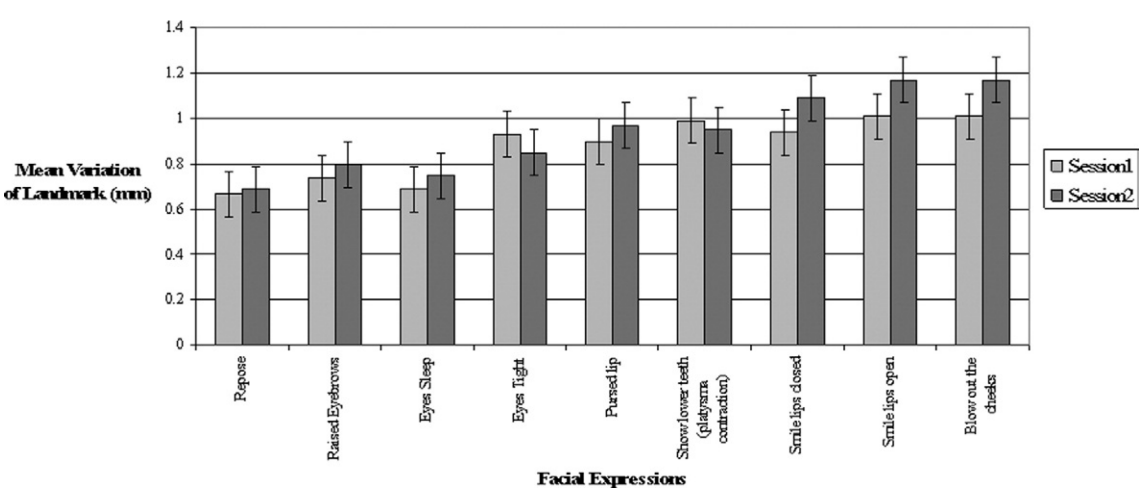

Figure 4 Mean variation for all landmarks to each expression for sessions 1 and 2.

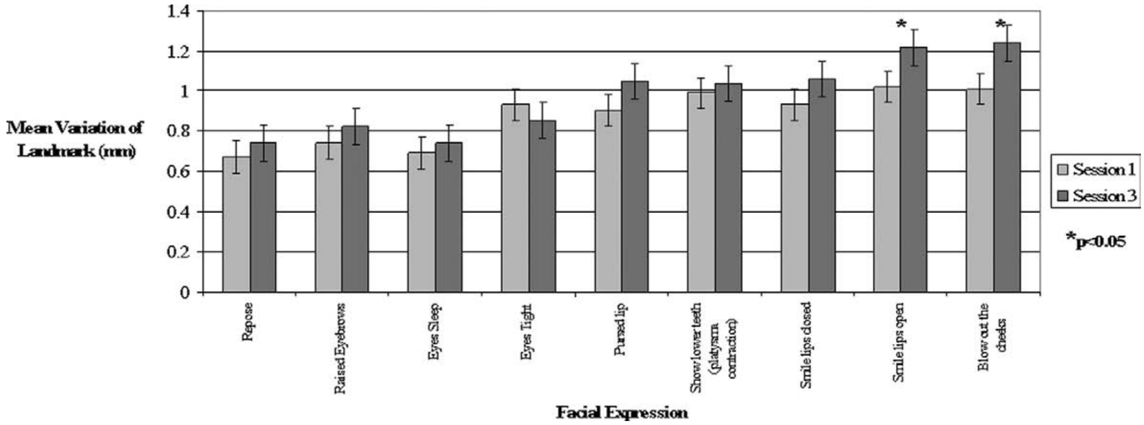

Figure 5 Mean variation for all landmarks to each expression for sessions 1 and 3.

Men showed less reproducibility of facial expressions than women except for "repose" and "pursed lip" expressions: however, this was not statistically significant ($P = 0.2$ and $P = 0.3$). Males showed significant differences compared with females in the reproducibility of "raised eyebrows" and "blowing out the cheeks," both of which were significantly different than females ($P = 0.04$ and $P = 0.03$).

When the left and right sides of face with subjects for all expressions were compared, we could find no significant difference (Student t test, $P = 0.46$) in facial symmetry.

DISCUSSION

Landmarks demonstrated their greatest reproducibility in the position of repose. Relatively fixed points such as the nasion and the glabella are easy to locate and do not move significantly from repose to a facial expression. Points that showed the greatest variability were the cheilion, gonion, and zygion. This is consistent with other author's findings that reproducibility of specific landmarks such as gonion and zygion is difficult, as locating these points required palpation on the face before capturing the image.[14] The cheilion has the greatest movement particularly with the smile with lips open, and therefore this represents a greater variation in discrepancy. This result is consistent with Trotman et al,[7] who also demonstrated that circum-oral markers had the largest movements.

The primary aim of this study was to evaluate the reproducibility of facial expression. Our results show that "repose" is the most reproducible facial expression followed by eyes closed as in sleep, raised eyebrows, eyes tight, pursed lip, exposing the lower teeth (platysma contraction), smile with lips closed, and smile with lips open in decreasing order. The least reproducible expression was "blow out the cheeks." It is likely that subjects when performing "blow out the cheeks" may create variable amounts of expression of the cheeks thus increasing its variability. One would anticipate that repose would be the most reproducible facial expression because it does not require the subject to create any facial movement and is merely the resting position of their face. These results are in harmony with the work from Johnston et al.[11]

Previous studies of reproducibility of facial expressions have only analyzed a limited number of expressions.[10,11,15] Johnston et al[11] investigated facial expression reproducibility, but this study only looked at five expressions in the lower half of the face. In order to test the facial nerve, we

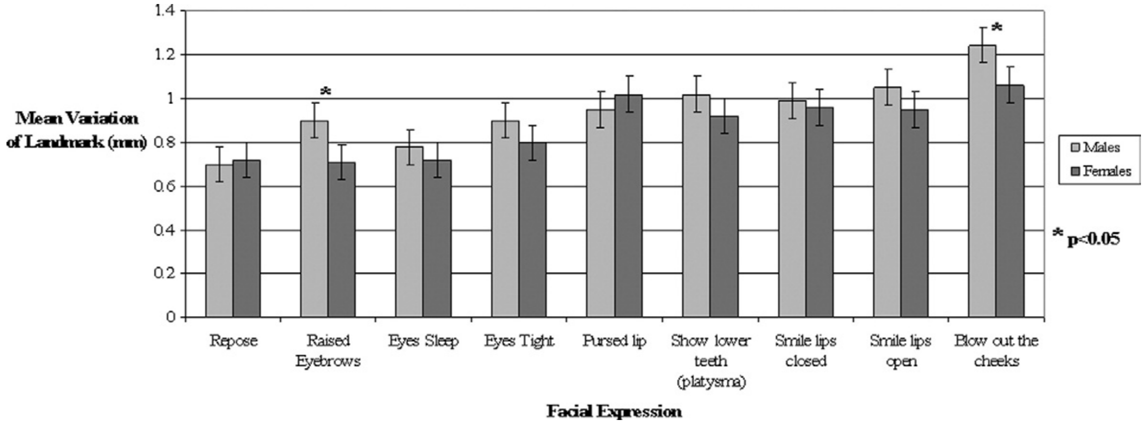

Figure 6 Mean variation for all landmarks of each expression comparing males and females.

used a variety of facial expressions that allows us to assess function of all branches of facial nerve (temporal, zygomatic, buccal, marginal mandibular, and cervical) and subsequent facial movement.

We specifically asked subjects to perform maximal contractions when performing facial expressions; for instance, when performing smile with lips open they were specifically asked to create the widest smile possible. If this is not clearly stated to the subject, it is likely that the subject will produce variability in facial expression. The results obtained by Burres[9] also agree that maximal facial expression should be performed because the measurements for gentle eye closure had a higher variability than those for tight eye closure. Zachrisson et al[16] used natural smile rather than maximal smile and this again could have created variability in these published results. In order to obtain accurate standardization of facial movements, they should always be performed to their maximal extent for as short a period as possible.

Interestingly, two subjects were unable to perform the facial expression “show me your lower teeth,” which is created by the contraction of platysma despite many attempts at practising the expression. It may be important to not include this expression when creating a facial grading system, as not all of the population may be able to perform this. All other facial expressions were easily produced by all subjects.

There was no significant difference between the first and second session reproducibility (albeit “smile with lips open” and “blow out the cheeks” approached significance). This shows that facial expressions are highly reproducible within a 10 to 20 minute period. These results concur closely with the previously reported analysis by Johnston et al.[11]

Few previous studies have assessed reproducibility of expressions. Wood et al[10] showed that the average test retest variability was 4 percent and 5 percent, respectively, day to day variability 5 percent and 6 percent. They also compared both sides of the face and found their side to side ability to be 6 percent and 14 percent, respectively. Intersubject variability was found to be the greatest at 25 percent and 23 percent, respectively. Similarly Dulguerov et al[15] also showed that in healthy subjects more than 80 percent of total variation in facial expression was accounted for by intersubject variability. Our findings confirm these results that intrasubject variability was low compared with intersubject differences.

When comparing the first and third session reproducibility, there was a significant difference in the “smile with lips open” and “blow out the cheeks” expressions. All of the other expressions showed no significant difference between the first and third sessions. Although there was significant difference in the “smile with lips open” and “blow out the cheeks” between sessions, this only amounted to a difference of 0.21 and 0.23 mm, respectively. These differences are clearly negligible in terms of clinical importance and would of course be undetectable with the use of a subjective facial paralysis grading scale.

While subjects were performing “smile with lips open” and “blow out the cheeks,” several facial landmarks move significantly, particularly in the periorbital region of the face (with little movement in the upper regions of the face). These periorbital region landmarks demonstrated consistently large movements during smile, which is likely to be as a result of the fact that these landmarks are moving in all three planes, horizontal, vertical, and dorso-ventral, and this increases the potential variability in the landmark. When performing eye closure and raised eyebrows, facial expressions, the upper and midfacial landmarks underwent the greatest amount of movement.

One potential area for error in this study is that the investigator had to place the landmarks on the face for each session, and it is impossible to place the points exactly in the same position. However, the VAM software used in this study allows the investigator to zoom in on specific areas and can in fact allow very accurate placement of landmark points. Furthermore, the program allows the investigator to rotate the 3-D facial image, and this can improve the accuracy of placing lateral landmarks, such as the zygion. Any ideal system that assesses facial movement should avoid touching or marking the patient’s face. Although not specifically investigated, we found the VECTRA to be a quick, noninvasive tool to investigate facial distances, volumes (eg, lower lip volume), surface areas, and surface distances. In essence, this system provides the surgeon or researcher with all the necessary data required to assess the recovery of facial paralysis. The VECTRA 2 is available for US $45,000 (UK £30,000) and is user friendly (requires no technician).

Females on average showed greater facial expression reproducibility in all expressions except for repose and pursed lip. It is interesting to note that Johnston et al[11] also observed that men could also reproduce the lip purse more accurately than women. In this study, the difference between males and females approached statistical significance ($P = 0.1$).

Males showed significant differences compared with females in the reproducibility of “raised eyebrows” and “blow out the cheeks.” It is hard to clearly explain the difference in the lack of reproducibility of these two facial expressions in males. However, males do show more excursion of facial landmarks, which may have contributed to the significant results.

CONCLUSION

Facial expressions are highly reproducible within a 10 to 20 minute period. Although there were some significant differences when comparing the reproducibility of expressions one month apart, these would be considered clinically insignificant. Males on average showed less facial expression reproducibility than females. We have also found that 3-D stereophotogrammetry is a useful technique for looking at facial motion and is a promising research tool for further investigation of facial motion.

AUTHOR INFORMATION

From the Plastic and Reconstructive Surgery Department, Queen Victoria Hospital NHS Foundation Trust, West Sussex, England.

Corresponding author: Adam R. Sawyer, BSc, DOHNS, MRCS, MSc, 11 Thurlow Close, Chapel Break, Norwich, Norfolk, England NR5 9NR.

E-mail address: drasawyer@hotmail.com.

AUTHOR CONTRIBUTION

Adam R. Sawyer, study design, data collection, writer; **Marlene See,** study design, writer; **Charles Nduka,** study design, equipment supply, writer.

FINANCIAL DISCLOSURE

None.

REFERENCES

1. House JW. Facial nerve grading systems. Laryngoscope 1983;93(8): 1056–69.
2. House JW, Brackmann DE. Facial nerve grading system. Otolaryngol Head Neck Surg 1985;93(2):146–7.
3. Ross BG, Fradet G, Nedzelski JM. Development of a sensitive clinical facial grading system. Otolaryngol Head Neck Surg 1996;114(3): 380–6.
4. Bajaj-Luthra A, Mueller T, Johnson PC. Quantitative analysis of facial motion components: anatomic and nonanatomic motion in normal persons and in patients with complete facial paralysis. Plast Reconstr Surg 1997;99(7):1894–1902; discussion 1903–4.
5. Paletz JL, Manktelow RT, Chaban R. The shape of a normal smile: implications for facial paralysis reconstruction. Plast Reconstr Surg 1994;93(4):784–789; discussion 790–1.
6. Ackerman JL, Ackerman MB, Brensinger CM, et al. A morphometric analysis of the posed smile. Clin Orthod Res 1998;1(1):2–11.
7. Trotman CA, Faraway JJ, Silvester KT, et al. Sensitivity of a method for the analysis of facial mobility. I. Vector of displacement. Cleft Palate Craniofac J 1998;35(2):132–41.
8. Frey M, Jenny A, Giovanoli P, et al. Development of a new documentation system for facial movements as a basis for the international registry for neuromuscular reconstruction in the face. Plast Reconstr Surg 1994;93(7):1334–49.
9. Burres SA. Facial biomechanics: the standards of normal. Laryngoscope 1985;95(6):708–14.
10. Wood DA, Hughes GB, Secic M, et al. Objective measurement of normal facial movement with video microscaling. Am J Otol 1994; 15(1):61–5.
11. Johnston DJ, Millett DT, Ayoub AF, et al. Are facial expressions reproducible? Cleft Palate Craniofac J 2003;40(3):291–6.
12. Ghoddousi H, Edler R, Haers P, et al. Comparison of three methods of facial measurement. Int J Oral Maxillofac Surg 2007;36(3):250–8.
13. Farkas L. Examination. In: Farkas LG, ed. Anthropometry of the head and face, 2nd ed. New York: Raven Press; 1994.
14. Ayoub A, Garrahy A, Hood C, et al. Validation of a vision-based, three-dimensional facial imaging system. Cleft Palate Craniofac J 2003;40(5):523–9.
15. Dulguerov P, Wang D, Perneger TV, et al. Videomimicography: the standards of normal revised. Arch Otolaryngol Head Neck Surg 2003; 129(9):960–5.
16. Zachrisson B. Esthetic factors involved in anterior tooth display and the smile: vertical dimension. J Clin Orthod 1998;32:432–45.