JPRS Journal of Plastic and Reconstructive Surgery
https://doi.org/10.53045/jprs.2022-0017
https://j-prs.jp/index.html

Original Research

Mathematical Validation of the Modified Sunnybrook Facial Grading System Using Four-dimensional Imaging

Mahmoud A Alagha[1], Xiangyang Ju[2], Stephen Morley[3] and Ashraf F Ayoub[1]

1) Scottish Craniofacial Research Group, Glasgow Dental Hospital & School, University of Glasgow, Glasgow, United Kingdom
2) Medical Device Unit, Department of Clinical Physics and Bioengineering, NHS Greater Glasgow and Clyde, West Glasgow Ambulatory Hospital, Glasgow, United Kingdom
3) Glasgow University, Canniesburn Plastic Surgery Unit at the Royal Infirmary of Glasgow, NHS Greater Glasgow and Clyde, Glasgow, United Kingdom

Abstract
Objectives**:** Despite the limited validity of the Sunnybrook grading index, it is routinely used for the clinical evaluation of facial palsy. This study aimed to assess the dynamic asymmetry in unilateral facial palsy and mathematically validate a modified version of the Sunnybrook facial grading system.
Methods**:** The Sunnybrook facial grading system was modified to provide more descriptions of the measured parameters of the distorted facial expression in unilateral facial paralysis. This correlation study was conducted on 16 patients with unilateral facial palsy and a matched control group. Three-dimensional video recordings of six facial expressions - rest, maximum smile, cheek puff, lip purse, eyebrow raising, and eye closure - were used for each case in the analysis. Advanced geometric morphometrics were applied to quantify facial asymmetry and morphology throughout the course of each expression. Seven professional assessors graded facial asymmetry for the 16 cases, twice, using the modified Sunnybrook index. Cross-correlations between the objective mathematical measurements and the subjective clinical grades were calculated.
Results**:** The inter- and intra-observer reproducibility of the modified Sunnybrook index was high ($r = -0.8$). Significant positive correlations were detected between the clinical grading of facial palsy and the mathematical measurements at rest, maximum smile, lip purse, and raising of eyebrows. The correlations between the modified Sunnybrook index and mathematical measurements were poor for cheek puff and forceful eye closure.
Conclusions**:** The modified Sunnybrook grading index proved reproducible and mathematically valid for the grading of unilateral facial paralysis in most facial expressions, except for cheek puff and forceful eye closure.

Keywords
4D imaging, facial palsy, dynamic asymmetry, stereophotogrammetry, clinical index

J Plast Recontr Surg Advance Publication
https://doi.org/10.53045/jprs.2022-0017

Introduction

Various methods have been proposed for the analysis of facial palsy, most of which are based on the subjective visual evaluation and inspection of the face[1-11]. The objective assessment of facial morphology and muscle function is crucial to improve the quality of surgical management of facial paralysis[6]. The lack of a universally accepted grading system for facial palsy limits the reliability of evaluating the outcomes of facial reanimation surgery[12-14]. Available knowledge on the dynamics of muscle movements in facial palsy is limited[15]. Fattah et al.[2] evaluated 19 facial nerve grading scales with regard to their reproducibility, inter-observer, and intra-observer variability. The Sunnybrook facial grading system satisfied all the criteria and has been proven to be reproducible[16-18], with high intra- and inter-observer agreements. The authors recommend the Sunnybrook facial grading system as the de facto standard for quantifying abnormal

Corresponding author: Ashraf F Ayoub, ashraf.ayoub@glasgow.ac.uk
Received: June 3, 2022, Accepted: October 25, 2022, Advance Publication by J-STAGE: April 14, 2023
Copyright Ⓒ The Japan Society of Plastic and Reconstructive Surgery
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

facial nerve function; however, its validity for the evaluation of distorted facial muscle movements has not yet been tested.

The literature confirmed that the validity of the subjective evaluation of facial deformity depends on the method of assessment and the experience of the panel of assessors[19-21]. Tan et al.[22] investigated the consistency between two stimulus media for the assessment of facial palsy: the face-to-face examination and the evaluation of video clips. Seven professional assessors evaluated 28 patients with facial palsy using three clinical grading scales: the House-Brackmann scale, Sydney scale, and Sunnybrook index. The assessments were repeated on the two-dimensional (2D) digital videos and recorded on the same day of the initial clinical evaluation. There was limited agreement among assessors in the assessment of facial palsy with the House-Brackmann scale and the Sunnybrook grading system. The study did not provide information on the calibration process and rating protocol of the assessors, but considered the mathematical objective validation.

Banks et al.[23-25] evaluated the facial mimetic function using the electronic, clinician-graded facial function scale (eFACE)[24]. The scoring consistency between the in-person evaluation and the assessment of video recording was investigated. Two surgeons assessed five patients who had various degrees of facial palsy. There was a strong agreement between the two modes of assessment, with a high reproducibility of the video-based eFACE scoring system.

The real-time three-dimensional (3D) imaging of muscle movements has proved to be reliable in recording the dynamics of facial expressions, which facilitates the analysis and quantification of morphological and functional distortions[26]. Stereophotogrammetry has allowed the consistent reliable recording of facial morphology[26-28]. It provides invaluable volumetric and morphological data for analysis and allows the capture of facial movement over time[29].

Recently, a computerized version of the Sunnybrook facial grading scale has been validated using 2D photographs of 30 subjects with unilateral facial palsy. A score of 1-5 quantified the disparity of facial muscle movements between the right and left sides of the face. The movement of automatically identified facial markers on the paralyzed side was compared to that of the unaffected side, and the difference calculated as a percentage. These were correlated to the Sunnybrook scales rated by three clinical assessors. The correlations of the software-derived scores with the clinical scales were limited because of the inconsistency of the assessors. The study did not consider the dynamics of facial expressions; it only measured the differences between the two sides of the face at maximum muscle movements for each expression, which may not show the greatest disparity between the paralytic and non-paralytic sides of the face[30].

The reliability of the Sunnybrook scale was tested in patients with subacute stroke using the video tape of the patients and showed high inter- and intra-rater reproducibility[31]. High reproducibility was also recorded when the grading system was applied in various languages, but these studies and the rest of the literature did not evaluate the mathematical validity of the Sunnybrook facial grading index; they only confirmed its reproducibility among clinical assessors[32-35]. This deficiency inspired our study, which aimed to explore the mathematical validity of the Sunnybrook grading index and to quantify the impairment of facial muscles in unilateral facial paralysis.

Recently, our team applied a linear mixed-effects model to assess the reproducibility of the modified Sunnybrook grading index; the assessors and the repeated assessments were the fixed effects, and the grading parameters and patients were the random effects. The results showed no significant effects of repeated subjective assessments on the grades (p = 1.0; estimated coefficient = 0.00); however, there was a significant effect of assessors on the grades (p = 0.0329; estimated coefficient = −0.02). The results indicated that the modified Sunnybrook (MSB) grading method was reproducible within individual assessors[36].

Therefore, the objective of this study was to assess the mathematical validity of the MSB grading index in quantifying the dynamic dysmorphology of facial expressions in unilateral facial paralysis.

Methods

Clinical grading of facial palsy using the modified Sunnybrook index

The Sunnybrook grading index was modified to allow a realistic validation against the mathematical measurements of facial asymmetry in unilateral facial paralysis. We included a set of specific descriptors to define more comprehensively the abnormality of facial movement of individual facial expressions the MSB index (**Table 1**) graded the facial appearance of the eye, cheek, and mouth regions at rest and during five facial expressions, to grasp the major deficiencies of the muscle activities because of facial paralysis. The snarl movement of the original Sunnybrook facial grading system was replaced with the cheek puff in the modified version. This was based on the study by Spielmann et al.[37], who concluded that there is no long-term evidence of patient benefit from the surgical corrections of the muscles responsible for the snarl movement. The cheek puff is more meaningful as it discloses the activity of the peri-oral muscle group, starting with the contraction of the orbicularis oris, followed by the contraction of the buccinator, risorius, and zygomaticus major and zygomatic minor muscles. The movements of these muscles are impaired because of facial paralysis, resulting in facial asymmetry[38].

The MSB index was piloted against the full range of facial palsy, before it was validated against the mathematical measurements of facial asymmetry in this study. Seven expert assessors viewed the video clips of the four-dimensional (4D) sequence of the five facial expressions of 16 patients, which generated about 1000 3D images per case. Patients provided consent for their photos to be included in publications. The assessors graded facial dysmorphology according to the MSB index, on two occasions, 8 weeks apart.

Table 1. The Modified Sunnybrook Index.

Facial expression	Index parameter	Grades
Rest	MSB 1	Grade: The eye resting asymmetry score compared to the non-affected side Abnormal: narrow/wide/eyelid surgery Normal
	MSB 2	Grade: The nasolabial fold resting asymmetry compared to the non-affected side Absent Altered: less or more pronounced Normal
	MSB 3	Grade: The corner of mouth resting asymmetry compared to the non-affected side Abnormal: corner dropped or pulled up/out Normal
Forehead wrinkling	MSB 4	Grade: The degree of eyebrow movement compared to the non-affected side No movement with no muscle activity No movement with muscle activity Reduced movement with minimal horizontal lines on affected side Reduced movement with horizontal lines on both sides Normal movement
Eye closure	MSB 5	Grade: The degree of eye movement compared to the non-affected side No movement Slight movement with significant gap Moderate movement with small gap Eye closure with forceful contracture Normal movement
Smile	MSB 6	Grade: The degree of smile movement compared to the non-affected side No movement Movement of the cheek without movement of the corner of the mouth Movement of the cheek with some lifting of the corner of the mouth Movement not fully symmetrical, showing teeth on affected side Normal movement
Cheek puff	MSB 7	Grade: The degree of cheek movement compared to the non-affected side No movement Flicker of movement Weak movement with incomplete oral seal Weak movement with complete oral seal Normal movement
Lip purse	MSB 8	Grade: The degree of lip pucker movement compared to the non-affected side No movement Slight movement without memetic lines Movement with philtral deviation Movement without philtral deviation Normal movement

The sample size calculation for the mathematical assessment of the asymmetry was based on the study of Johnston et al.[39]. The expected mean difference in overall landmark positions was 0.49 mm (effect size) and 1 mm standard deviation of the mean.

Capture of facial movements

The Di4D facial imaging capture system, developed by Dimensional Imaging Limited, Hillington, Scotland, UK, was used to record facial movements. The capture station consists of two gray-scale cameras (model avA 1600-65 km/kc, resolution 1600×1200 pixels), a sensor (model KAI-02050; ON Semiconductor, Phoenix, AZ, USA), a color camera (Kodak sensor model KAI-02050, Basler, Germany), and a lighting system (model DIV-401 Diva Lite; Kino Flo Corporation, Burbank, CA, USA). The gray-scale cameras captured the 3D image sequence at a rate of 60 3D facial frames per second, and the color camera captured the surface color and texture. The capture system was connected to a desktop computer to build the sequence of the 3D facial images for each facial expression.

In addition to the facial expression at rest, five facial expressions - maximum smile, lip purse, cheek puff, eyebrow raising, and eye closure - were recorded. Each facial expres-

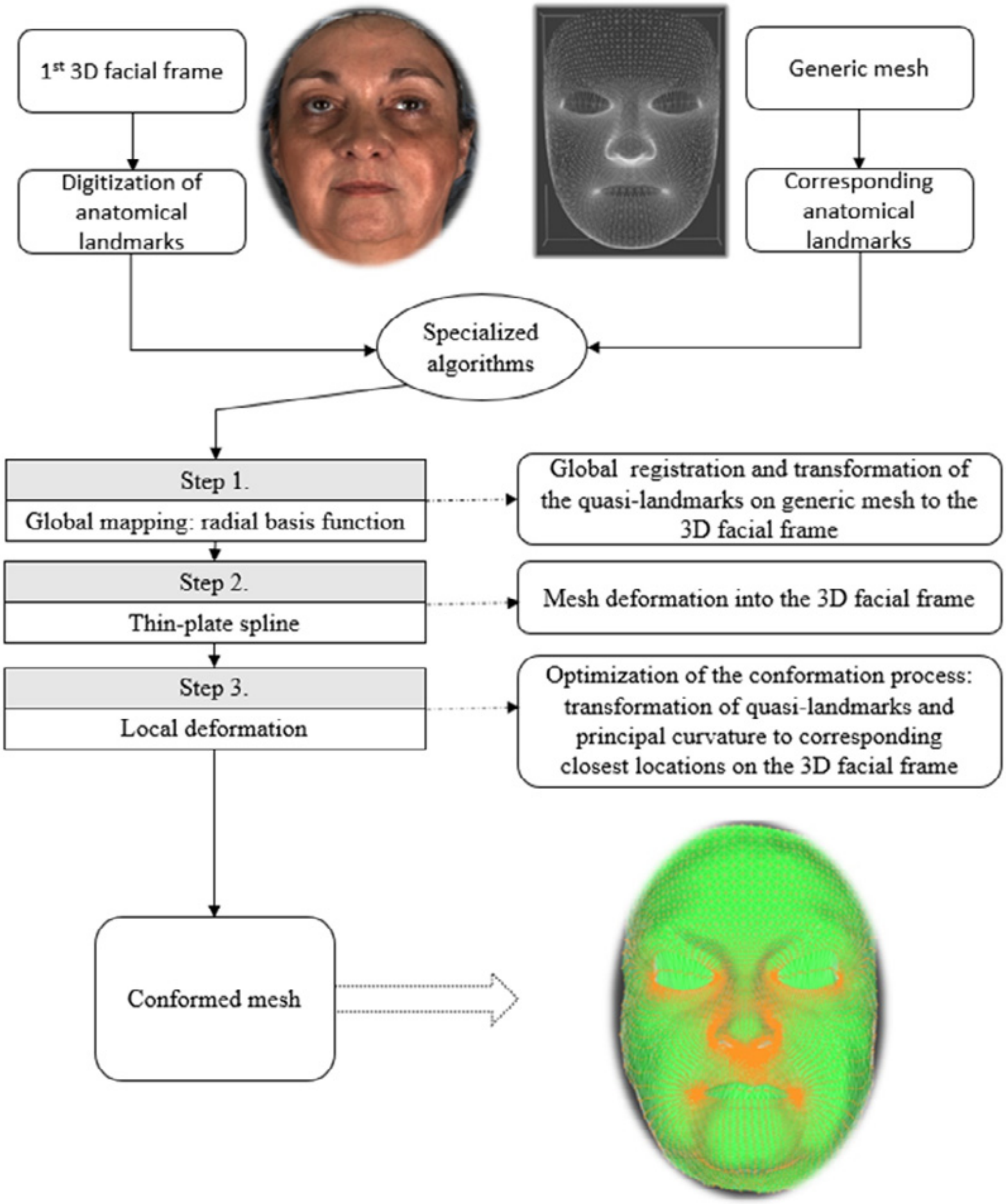

Figure 1. The generic mesh conformation process. Step 1: global mapping; step 2: thin-plate spline; and step 3: local deformation.

sion was recorded over 3-6 seconds, according to an established protocol[40]. This produced a minimum of 180 3D facial frames per expression.

The 3D image processing included the manual digitization of facial anatomical landmarks as well as the building of the dense surface model (DSM). A set of 23 anatomical landmarks were manually digitized on the captured 3D facial model, according to the established protocol[41]. The reproducibility of the digitization of landmarks was investigated and published in our previous study[40]. The 3D coordinates of the landmarks in the X, Y, and Z directions were saved in the "dilm" file format. These were utilized to clone a generic facial template into the individual facial morphology.

A generic facial template composed of 7000 symmetrical and uniformly distributed quasi-landmarks was used to build dense correspondence on the first frame of the sequence of the captured 3D image of each facial expression. The generic mesh was mathematically adapted (conformed) to the 3D image to portray the detailed morphology of the face during each facial expression (**Figure 1**). The mesh conformation process was shown to be accurate to within 0.2-0.7 mm[42]. The conformed meshes provided the mathematical shape for analyses. The accuracy of point tracking of the facial landmarks of the set of 3D images, which were generated during the capture of facial expressions, was within 0.5 mm[43].

The conformed mesh provided a comprehensive full facial representation. The Di4D-View software tracked the changes over time of the 3D facial morphology of each facial expression.

Mathematical analyses of facial asymmetry

Each of the tracked conformed models was aligned on its own mirror image copy using partial Procrustes analysis.

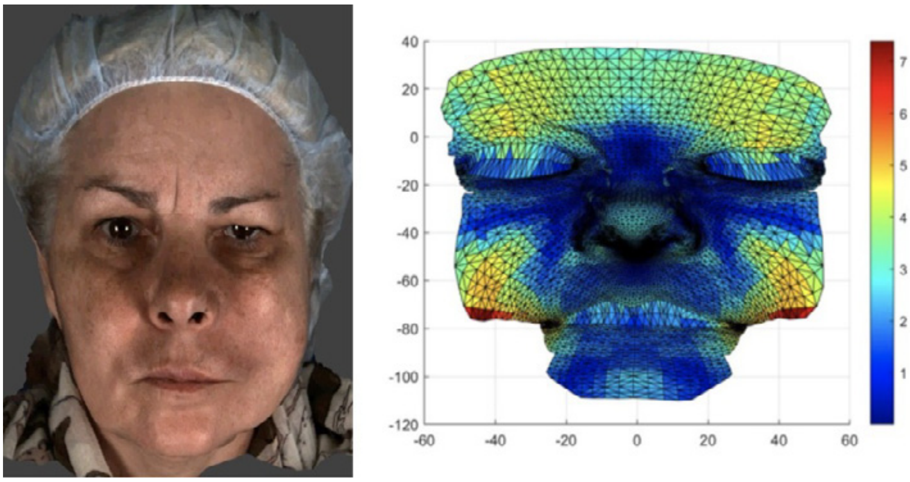

Figure 2. Resting facial asymmetry showing the distance patterns of the minimal facial asymmetry (global asymmetry) at rest. The color code ranges from blue (1) to red (7). The blue color indicates perfect symmetry because of the minimal difference in the mean absolute distance between the vertices of the 3D images and its corresponding reflections. The changing color from deep blue to red indicates an increase in the magnitude of facial asymmetry.

The asymmetry score was calculated based on the average Euclidean distance of the corresponding vertices between the model and its aligned mirror copy. The average distances between the 7000 vertices of the conformed mesh of the 3D facial image and the corresponding vertices of the mirror copy provided a measure of the asymmetry score. In perfect symmetry, the score equals zero.

The DSM of the 3D facial images was subdivided into 10 facial anatomical regions (full face, forehead, eyes, nose, cheeks, nasolabial, upper lip, lower lip, chin, and corner of the mouth) to facilitate the assessment of the regional asymmetry. **Figure 1** explains the steps that were followed.

Dynamic asymmetry scores at each expression per individual were calculated. For patients and controls, the mean asymmetry score for 10 regions of each frame of six expressions was calculated. The minimum, mean, median, and maximum values (pts) of the asymmetry were calculated for the 3D model sequences of each of the 10 individual regions (full face, forehead, eyes, nose, cheek, nasolabial, upper lip, lower lip, chin, and corner of mouth).

For each patient, the Z scores of the minimal, mean, median, and maximum asymmetry values (pts) were calculated for the 10 regions of the six expressions based on the equation:

$$Z = \text{pts} - \text{mean\_}\mathit{asymmetry\_of\_regions\_of\_controls}/\text{std} \quad (1)$$

$$Z = 0.0 \text{ if } (\text{pts} - \text{mean\_asymmetry\_of\_regions\_of\_controls})/\text{std} < 0 \quad (2)$$

The relationship between the clinical grading using the MSB and the mathematical measurements of asymmetry of each expression derived from the 180 3D images of each facial expression was assessed.

Results

No statistically significant differences were detected between the first and second clinical assessments of unilateral facial palsy using the MSB index, with an average intra-rater reliability of $r = 0.8$ and an inter-rater reliability of $r = 0.6$. The landmarking reproducibility assessment[40], using a paired-sample t-test, showed no statistically significant differences between the first and second landmarking sessions ($p < 0.05$), with an absolute distance between the average repeated landmarking of 0.99 mm. p-Values for the mediolateral, vertical, and anteroposterior directions were 0.11, 0.22, and 0.73, respectively.

The minimal value of the facial asymmetries of the rest expression had a higher magnitude of correlation coefficient than that of other values. The MSB correlated reasonably well with the mathematical measurements of facial asymmetry at rest and throughout the course of facial expressions. The clinical grading of the smile was strongly correlated with the objective measurements, followed by lip purse, eye closure, cheek puff, and eyebrow raising.

Assessment of facial asymmetry at rest (Figure 2)

The three parameters of the MSB index at rest showed high correlations with the minimal asymmetry measurements of the cheek (MSB2 $r = -0.76$) and the upper lip (MSB3 $r = -0.69$) but lower correlations with that of the eye region (MSB1 $r = -0.42$; **Table 2**).

Assessment of the smile

Various patterns of facial asymmetry were noted with the maximum smile (**Figure 3**). There were significant correlations of the measured asymmetries and the clinical assessments of the smile at the relevant facial regions (MSB6 $r = -0.67$ at cheek; **Table 3**).

Assessment of lip purse

Strong correlations were detected between the subjective assessments of the MSB index and the average of the mathematical measurements of the asymmetry (r = –0.65 for the upper lip; r = –0.62 for the lower lip; and r = –0.62 for the corner of the mouth regions; **Table 4**).

Assessment of cheek puff

Table 5 shows a general lack of statistically significant correlations between the clinical MSB7 scores and the mathematical measurements. We noted some improvement in facial asymmetry with cheek puff in some cases and some worsening in others (**Figure 4**). Statistically significant correlations were detected between MSB7 and asymmetry scores at the corner of the mouth (coefficient of correlation r = –0.52) and MSB7 showed a significant correlation with asymmetry scores at the cheek region (r = –0.69; **Table 5**).

Assessment of forehead wrinkling "eyebrow raising"

The correlations between the MSB and the mathematical measurements of facial asymmetry associated with this expression were moderate to poor; none was statistically significant (**Table 6**).

Table 2. Correlations between Minimum Mathematical Asymmetry Values in Segmented Facial Regions and MSB Parameters at Rest Expression.

Facial region	Index parameter	Correlation with minimum value of mean asymmetries in expression	p-Value
Full face	MSB 1	–0.45	0.07
	MSB 2	–0.74	0.00
	MSB 3	–0.60	0.01
Forehead	MSB 1	–0.48	0.05
Eyes	MSB 1	–0.42	0.10
Cheek	MSB 2	–0.76	0.00
Nasolabial	MSB 2	–0.76	0.00
Upper lip	MSB 3	–0.69	0.00
Lower lip	MSB 3	–0.47	0.06
Chin	MSB 3	–0.47	0.06
Corner of mouth	MSB 3	–0.60	0.01

Figure 3. Maximum smile. Various patterns of the dynamics of facial asymmetry with the maximum smile of the unilateral facial palsy.

Table 3. Correlations between Mean and Median Mathematical Asymmetry Values and MSB Parameters of Smile.

Facial region	Index parameter	Correlation with mean	p-Value	Correlation with median	p-Value
Full face	MSB 6	–0.66	0.00	–0.66	0.00
Cheek	MSB 6	–0.67	0.00	–0.67	0.00
Nasolabial	MSB 6	–0.64	0.00	–0.64	0.00
Upper lip	MSB 6	–0.66	0.00	–0.67	0.00
Lower lip	MSB 6	–0.61	0.01	–0.61	0.01
Chin	MSB 6	–0.59	0.01	–0.59	0.01
Corner of mouth	MSB 6	–0.60	0.01	–0.61	0.01

Table 4. Correlations between Mean and Median Mathematical Asymmetry Values and MSB Parameters during Lip Purse at Segmented Facial Regions.

Facial region	Index parameter	Correlation with mean	*p*-Value	Correlation with median	*p*-Value
Full face	MSB 8	–0.52	0.03	–0.50	0.04
Cheek	MSB 8	–0.52	0.04	–0.49	0.05
Nasolabial	MSB 8	–0.45	0.07	–0.43	0.09
Upper lip	MSB 8	–0.65	0.00	–0.61	0.01
Lower lip	MSB 8	–0.62	0.01	–0.60	0.01
Chin	MSB 8	–0.43	0.09	–0.46	0.07
Corner of mouth	MSB 8	–0.62	0.01	–0.60	0.01

Table 5. Correlations between Mean and Median Mathematical Asymmetry Values and MSB Parameters of Cheek Puff in Segmented Facial Regions.

Facial region	Index parameter	Correlation with mean	*p*-Value	Correlation with median	*p*-Value
Full face	MSB 7	–0.44	0.08	–0.48	0.05
Cheek	MSB 7	–0.69	0.00	–0.67	0.00
Nasolabial	MSB 7	–0.41	0.11	–0.46	0.07
Upper lip	MSB 7	–0.38	0.14	–0.40	0.12
Lower lip	MSB 7	–0.43	0.10	–0.48	0.05
Chin	MSB 7	–0.37	0.15	–0.36	0.17
Corner of mouth	MSB 7	–0.52	0.03	–0.58	0.01

Assessment of eye closure

The correlations between the MSB and the mathematical measurements of facial asymmetry associated with eye closure were moderate, except that with the forehead asymmetry (MSB5 $r = -0.67$; **Table 7**).

Figure 5 illustrates the color map of the facial movements of facial expressions of each of the five key frames of one of the patients included in this study. The facial movements of the 7000 points of the conformed mesh were tracked throughout the five key frames: at rest; the intermediate between rest and the maximum expression; the maximum expression; the intermediate between the maximum and the end of expression. The green color represents no facial movement and the yellow color represents movements larger than 8 mm of each of the six expressions.

Figure 6 shows the 3D frame of the maximum muscle movements of each facial expression. The blue color represents no muscle movements and the red color represents the maximum movements of the same case. The mean MSB grading scores of the seven assessors for this patient were 1.07, 1.42, 1.00, 3.07, 1.00, 3.44, and 2.64, respectively. **Table 8** shows the Z scores of the asymmetries of all facial expressions, at the nine predefined anatomical regions of this patient.

Discussion

The mathematical evaluation of distorted facial muscle movements provides the ground truth for the objective evaluation of functional deficits[6]. However, this requires a sophisticated 4D imaging system to capture facial expressions and the application of various software packages for 3D measurements and statistical analyses[26].

Currently, there is no universally accepted grading system to quantify facial palsy and measure the impact of various treatment modalities[12-14]. This is particularly important following surgical facial reanimation for the correction of asymmetry and deficient or distorted muscle movements.

To maximize the clinical relevance of the analysis, the 3D facial morphology was divided into anatomical regions, each representing a group of muscles, to quantify the distortion of facial expressions in unilateral facial palsy. This enabled the direct correlation to be assessed between the measured asymmetry and the subjective grading of the distorted muscle movements due to facial palsy. Kim and Oh[44] divided the face into three volumetric proportions, based on four horizontal lines passing through a set of facial landmarks.

Da Costa et al. used the MSB grading system to assess parotidectomy-related facial nerve lesions. They removed the “snarl” expression, performed by the levator labii superioris and levator labii superioris alaeque nasi muscles, and inserting the “show the lower teeth” expression, performed by the

Figure 4. Cheek puff. The top images show facial asymmetry at rest. The bottom images show facial asymmetry at the maximum expression of cheek puff. The two images on the left show that the facial asymmetry is worse at maximum cheek puff. The two images on the right side show that cheek puff improved facial symmetry.

depressor labii inferioris and depressor anguli oris muscles, innervated by the marginal mandibular nerve[45]. There was no inter-rater evaluation or electromyographic examination. These limitations were addressed in our study. We evaluated the mathematical validity of the Sunnybrook grading system, which we modified, and proved its reproducibility in our previous publication[36].

Codari et al.[46] segmented the face according to the regional anatomical innervation of the trigeminal nerve. However, the volumetric analysis of horizontal facial sections does not account for the dissimilarities between the right and left sides of the face and the division of the face into hemifacial thirds based on midline anatomical landmarks may be difficult to identify in severe facial dysmorphology[47].

Two main approaches were considered for the 3D assessment of facial muscle movements: the static and dynamic methods. Gibelli et al.[48] applied surface-based analysis to quantify the shape difference between the facial 3D images at rest and with various maximum expressions. The main drawback of the static approach is that the quantification of asymmetry at the maximum expression does not measure the dynamics of facial muscle movements throughout the course of the expression, from the start to the final rest pose. In our previous study, five key frames of the 180 3D image sequences recorded during each facial expression were considered to describe the dynamics of the facial dysmorphology in unilateral facial palsy[36]. This limited the analysis of facial asymmetry to these key 3D images and facial morphological characteristics between these were not measured. In this study, the mathematical asymmetry scores were derived from the entire 3D image sequence of each expression.

The correlation between the mathematical measurements of the Sunnybrook index and the subjective evaluation of facial palsy has not been fully investigated before. Studies

Table 6. Correlations between Mean and Median Mathematical Asymmetry Values and MSB Parameters during Eyebrow Raising in Segmented Facial Regions.

Facial region	Index parameter	Correlation with mean	*p*-Value	Correlation with median	*p*-Value
Full face	MSB 4	−0.40	0.12	−0.42	0.10
Forehead	MSB 4	−0.41	0.11	−0.43	0.09
Eyes	MSB 4	−0.40	0.12	−0.42	0.10

Table 7. Correlations between Mean and Median Mathematical Asymmetry Values and MSB Parameters in Segmented Facial Regions during Eye Closure.

Facial region	Index parameter	Correlation with mean	*p*-Value	Correlation with median	*p*-Value
Full face	MSB 5	−0.52	0.03	−0.51	0.04
Forehead	MSB 5	−0.67	0.00	−0.65	0.00
Eyes	MSB 5	−0.60	0.01	−0.58	0.01

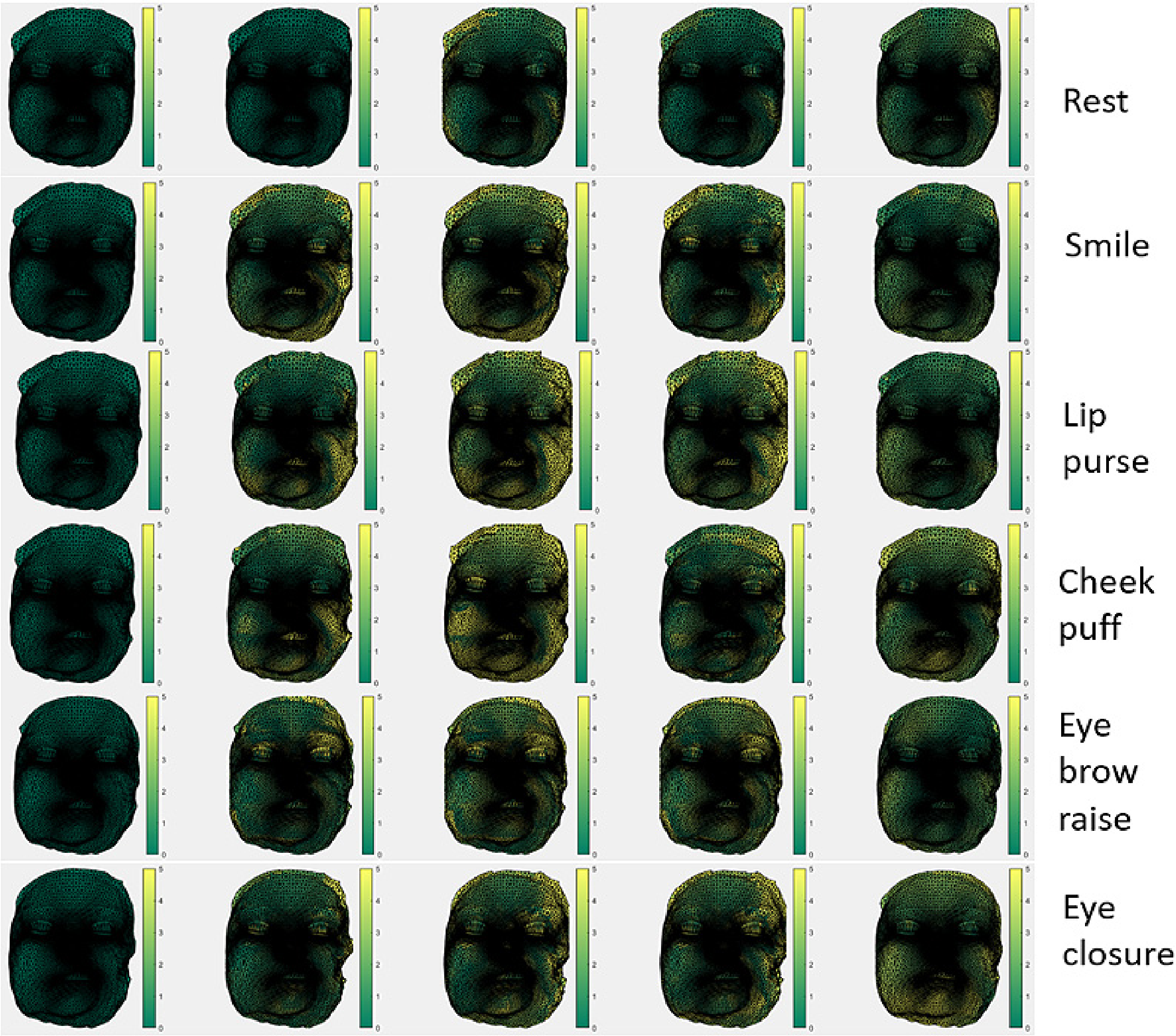

Figure 5. Five key frames of six expressions of the patients included in the study.

were limited to correlation analyses between the subjective clinical assessments of the index and 2D objective measurements. The correlation with the 3D measurements was based on a limited number of facial landmarks[49]. Measurements were obtained from nine facial landmarks only, which may have underrepresented the complexity of facial muscle movements as the landmark-based analysis does not adequately describe the 3D facial dysmorphology. In previous studies, facial muscle movements were measured at the maximum displacement of facial landmarks, which did not consider the entire dynamics of muscle movements. In our study, we noted that the maximum facial expression may reduce the asymmetric facial appearance, so the analysis should be based on the entire course of each expression.

This study introduced, for the first time, the assessment of facial palsy by a panel of expert assessors, based on the evaluation of 4D video recordings of each facial expression. We also conducted a sophisticated objective quantification of the dynamic facial asymmetry to explore the mathematical validity of the subjective assessment. The level of agreement between the mathematical measurements and the subjective assessments varied among facial expressions. The lack of significant correlations between the mathematical measurements and subjective assessments was mainly due to the sample size and the restricted descriptions of some subjective parameters, which might have limited the clinical grading of the expert assessors. The descriptors for grading the eyebrow raising did not fully coincide with the mathematical representation of facial asymmetry. The grading of face function during eyebrow wrinkling may have misled the objective assessment to the forehead region (forehead wrinkling) rather than the eyebrow region (eyebrow raising). Facial expressions took 4-6 seconds, which may be too fast to subjectively quantify the subtle combination of the reduced magnitude and directional asymmetric movements of facial palsy. Kim et al.[50] investigated the impact of timing on the detection and perception of asymmetry of facial muscle movements. They reported a strong inverse correlation (R = 0.82) between the time delay and the grading of naturalness. This raised an intriguing question of which facial features attract the assessor's attention during the assessment of facial muscle movements. The possibility that certain features

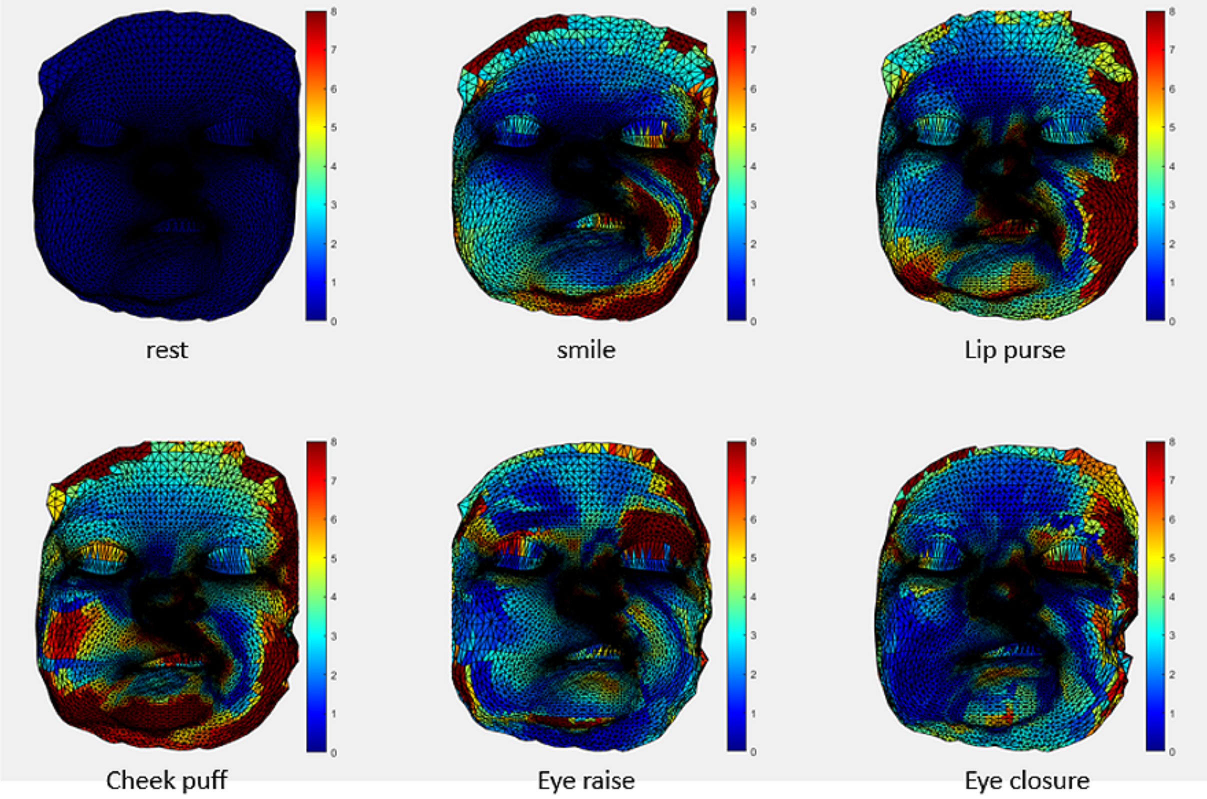

Figure 6. The maximum expression 3D frame of each facial expression. The blue color represents no muscle movement and the red color represents the maximum movement of the same case.

Table 8. Average Facial Asymmetry Measurements of One Patient Included in This Study.

	Full face	Forehead	Eyes	Nose	Cheek	Nasolabial	Upper lip	Lower lip	Chin
Rest	5.83	8.90	4.06	3.01	4.09	4.55	5.83	8.90	4.06
Smile	1.15	1.02	0.81	0.00	1.03	0.53	1.15	1.02	0.81
Cheek puff	2.77	2.11	1.34	1.04	4.28	4.91	2.77	2.11	1.34
Lip purse	0.08	2.81	1.90	0.08	1.57	1.45	0.08	2.81	1.90
Eyebrow raise	4.79	7.28	2.52	2.05	5.14	3.88	4.79	7.28	2.52
Eye closure	3.46	7.10	3.78	2.33	2.92	3.10	3.46	7.10	3.78

of facial asymmetry predominate the decision-making process of grading the asymmetry of muscle movements cannot be ruled out. On the other hand, minor subtle asymmetries, which are readily measurable mathematically, may not be noticed during routine subjective assessments of facial muscle movements. Hence, different facial aesthetic units possess a specific threshold of perception[51].

Poor correlations were noted between the mathematical measurements and the objective assessments around the eye region and cheek. The limited correlation between the MSB and the clinical grading of asymmetry of cheek puff could be attributed to the diverse forms of dynamic dysmorphology among the cases in our study. Interestingly, the dynamic movements of the cheek during puffing reduced the facial asymmetry, which was noted at the rest expression, as demonstrated in **Figure 4**.

The objective quantification of facial dysmorphology provides an accurate measurement of asymmetry at individual facial regions, which underpins the contribution of each group of facial muscles to the measured asymmetry. The mathematical assessment of the dynamics of facial expressions provided further insights into the pattern of facial dysmorphology in unilateral facial palsy. This is particularly important for the evaluation of the impact of surgical rehabilitation and the longitudinal monitoring of facial palsy. This is particularly important to evaluate the improvement in muscle movement following the free muscle flaps for facial reanimation[13]. The accuracy in predicting the required volume of the muscle transfer to the facial region is limited because of the limited validity of objective measurement tools[52].

We would like to highlight that the magnitude, speed, and direction of the asymmetry of the nonverbal expressions were not measured in this study. A larger sample size and the inclusion of patients of diverse ethnic backgrounds would be highly recommended. The application of automated landmarking[53-55], and artificial intelligence to develop an app-based grading of facial muscle movements that could be validated against clinical scores is highly desirable[56]. The development of a more comprehensive grading system of facial muscle movements that provides clinicians with detailed descriptors to quantify muscular dysfunction, including the directionality of asymmetry, should be considered in future studies.

Conclusion

The MSB index proved reproducible and mathematically valid for the grading of unilateral facial paralysis in most facial expressions. The mathematical validity of the MSB grading index for the assessment of facial dysmorphology during cheek puff and forceful eye closure was limited. Future studies may consider the development of a more comprehensive grading system of facial paralysis.

Acknowledgments: The authors would like to thank Mr. A Ray, Mr. D Soutar, Mr. I Mackay, consultant plastic surgeons; Mrs C Harkness, physiotherapist; and Mr. T Gillgrass, consultant orthodontists for their in the subjective analysis of the results. Also thanks to Di4D (Dimensional Imaging, Ltd) for their support.
Author Contributions: The authors contributed equally to the design of the study, data analyses, and writing of the manuscript.
Conflicts of Interest: There are no conflicts of interest.
Ethical Approval: Ethical approval was obtained from the Research Ethics Committee (Reference 17/SC/0541) and the Research and Development National Health Services Greater Glasgow and Clyde Health Board (Reference GN17OD401).
Consent to Participate: The patients/participants provided their written informed consent to participate in this study.
Consent for Publication: The patients/participants provided their written informed consent to participate in this study.

References

1. Niziol R, Henry FP, Leckenby JI, et al. Is there an ideal outcome scoring system for facial reanimation surgery? A review of current methods and suggestions for future publications. J Plast Reconstr Aesthet Surg. 2015 Apr;68(4):447–56.
2. Fattah AY, Gurusinghe ADR, Gavilan J, et al. Facial nerve grading instruments: systematic review of the literature and suggestion for uniformity. Plast Reconstr Surg. 2015 Feb;135(2):569–79.
3. Samsudin WSW, Sundaraj K. Clinical and non-clinical initial assessment of facial nerve paralysis: a qualitative review. Biocybern Biomed Eng. 2014 Jan;34(2):71–8.
4. Samsudin WSW, Sundaraj K. Evaluation and grading systems of facial paralysis for facial rehabilitation. J Phys Ther Sci. 2013 Jan; 25(4):515–9.
5. Lee LN, Susarla SM, Hohman MH, et al. A comparison of facial nerve grading systems. Ann Plast Surg. 2013 Mar;70(3):313–6.
6. Tzou CH, Pona I, Placheta E, et al. Evolution of the 3-dimensional video system for facial motion analysis: ten years' experiences and recent developments. Ann Plast Surg. Aug. 2012 Aug;69(2):173–85.
7. Samsudin WSW, Sundaraj K. Image processing on facial paralysis for facial rehabilitation system: a review. 2012. p. 259–63.
8. Brenner MJ, Neely JG. Approaches to grading facial nerve function. Semin Plast Surg. 2004 Feb;18(1):13–22.
9. Kang TS, Vrabec JT, Giddings N, et al. Facial nerve grading systems (1985-2002): beyond the House-Brackmann scale. Otol Neurotol. 2002 Sep;23(5):767–71.
10. Dulguerov P, Marchal F, Wang D, et al. Review of objective topographic facial nerve evaluation methods. Am J Otol. 1999 Sep;20 (5):672–8.
11. Kleiss IJ, Eviston TJ, Hadlock TA. Quantitative assessment of facial function in patients with peripheral facial palsy: a systematic review. Assess Facial Funct Peripher Facial Palsy Neth. 2015;97–120.
12. Schlosshauer T, Kueenzlen L, Groetsch T, et al. Long-term outcomes of Gillies and McLaughlin's dynamic muscle support in irreversible facial paralysis: a retrospective single-centre study with 25-year follow-up. J Plast Reconstr Aesthet Surg. 2020 Sep;73(9): 1706–16.
13. Roy M, Corkum JP, Shah PS, et al. Effectiveness and safety of the use of gracilis muscle for dynamic smile restoration in facial paralysis: a systematic review and meta-analysis. J Plast Reconstr Aesthet Surg. 2019 Aug;72(8):1254–64.
14. Bos R, Reddy SG, Mommaerts MY. Lengthening temporalis myoplasty versus free muscle transfer with the gracilis flap for long-standing facial paralysis: a systematic review of outcomes. J Craniomaxillofac Surg. 2016 Aug;44(8):940–51.
15. Dong A, Zuo KJ, Papadopoulos-Nydam G, et al. Functional outcomes assessment following free muscle transfer for dynamic reconstruction of facial paralysis: a literature review. J Craniomaxillofac Surg. 2018 May;46(5):875–82.
16. Kanerva M, Poussa T, Pitkäranta A. Sunnybrook and House-Brackmann Facial Grading Systems: intrarater repeatability and interrater agreement. Otolaryngol Head Neck Surg. 2006;135(6): 865–71.
17. Neely JG, Cherian NG, Dickerson CB, et al. Sunnybrook facial grading system: reliability and criteria for grading. Laryngoscope. 2010 Dec;120(5):1038–45.
18. Chong LSH, Eviston TJ, Low TH, et al. Validation of the clinician-graded electronic facial paralysis assessment. Plast Reconstr Surg. 2017 Jul;140(1):159–67.
19. Ritter K, Trotman C-A, Phillips C. Validity of subjective evaluations for the assessment of lip scarring and impairment. Cleft Palate Craniofac J. 2002 Nov;39(6):587–96.
20. Al-Omari I, Millett DT, Ayoub A, et al. An appraisal of three methods of rating facial deformity in patients with repaired complete unilateral cleft lip and palate. Cleft Palate Craniofac J. 2003 Sep;40(5):530–7.
21. Mosmuller DGM, Maal TJ, Prahl C, et al. Comparison of two- and three-dimensional assessment methods of nasolabial appearance in cleft lip and palate patients: do the assessment methods measure the same outcome? J Craniomaxillofac Surg. 2017 Aug; 45(8):1220–6.
22. Tan JR, Coulson S, Keep M. Face-to-face versus video assessment of facial paralysis: implications for telemedicine. J Med Internet Res. 2019 Apr;21(4):e11109.
23. Banks CA, Jowett N, Hadlock TA. Test-retest reliability and agreement between In-Person and video assessment of facial mimetic function using the eFACE facial grading system. JAMA Facial Plast Surg. 2017 May;19(3):206–11.
24. Banks CAMD, Bhama PK, Park J, et al. Clinician-graded electronic facial paralysis assessment: the eFACE. Plast Reconstr Surg. 2015 Aug;136(2):223e–30e.
25. Greene JJ, Guarin DL, Tavares J, et al. The spectrum of facial palsy: the MEEI facial palsy photo and video standard set. Laryngoscope. 2020 Jan;130(1):32–7.

26. Tzou CH, Artner NM, Pona I, et al. Comparison of three-dimensional surface-imaging systems. J Plast Reconstr Aesthet Surg. 2014;67(4):489–97.

27. Kook MS, Jung S, Park HJ, et al. A comparison study of different facial soft tissue analysis methods. J Craniomaxillofac Surg. 2014 Jul;42(5):648–56.

28. Fourie Z, Damstra J, Gerrits PO, et al. Evaluation of anthropometric accuracy and reliability using different three-dimensional scanning systems. Forensic Sci Int. 2011;207(1-3):127–34.

29. Petrides G, Clark JR, Low H, et al. Three-dimensional scanners for soft-tissue facial assessment in clinical practice. J Plast Reconstr Aesthet Surg. 2021 Apr;74(3):605–14.

30. Jirawatnotai S, Jomkoh P, Voravitvet TY, et al. Computerized Sunnybrook facial grading scale (SBface) application for facial paralysis evaluation. Arch Plast Surg. 2021 May;48(3):269–77.

31. Tramontano M, Morone G, LA Greca FM, et al. Sunnybrook Facial Grading System reliability in subacute stroke patients. Eur J Phys Rehabil Med. 2021 Jun;57(5):685–90.

32. Cabrol C, Elarouti L, Montava AL, et al. Sunnybrook facial grading system: intra-rater and inter-rater variabilities. Otol Neurotol. 2021 Aug;42(7):1089–94.

33. Mengi E, Kara CO, Ardiç FN, et al. Validation of the Turkish version of the Sunnybrook facial grading system. Turk J Med Sci. 2020;50(2):478–84.

34. Pavese C, Tinelli C, Furini F, et al. Validation of the Italian version of the Sunnybrook Facial Grading System. Neurol Sci. 2013 Apr;34(4):457–63.

35. Neumann T, Lorenz A, Volk GF, et al. Validation of the German version of the Sunnybrook facial grading system. Laryngorhinootologie. 2017 Mar;96(3):168–74.

36. Alagha MA, Ayoub A, Morley S, et al. Objective grading facial paralysis severity using a dynamic 3D stereo photogrammetry imaging system. Opt Lasers Eng. 2022 Mar;150:16876.

37. Spielmann PM, White PS, Hussain SS. Surgical techniques for the treatment of nasal valve collapse: a systematic review. Laryngoscope. 2009 Jul;119(7):1281–90.

38. Homer N, Fay A. Facial paralysis. Adv Ophthalmol Optom. 2018 Aug;3(1):357–73.

39. Johnston DJ, Millett DT, Ayoub AF, et al. Are facial expressions reproducible? Cleft Palate Craniofac J. 2003 May;40(3):291–6.

40. Alagha MA, Ju X, Morley S, et al. Reproducibility of the dynamics of facial expressions in unilateral facial palsy. Int J Oral Maxillofac Surg. 2018 Feb;47(2):268–75.

41. Gwilliam JR, Cunningham SJ, Hutton T. Reproducibility of soft tissue landmarks on three-dimensional facial scans. Eur J Orthod. 2006 Oct;28(5):408–15.

42. Cheung MY, Almukhtar A, Keeling A, et al. The accuracy of conformation of a generic surface mesh for the analysis of facial soft tissue changes. PLOS ONE. 2016 Apr;11(4):e0152381.

43. Al-Anezi T, Khambay B, Peng MJ, et al. A new method for automatic tracking of facial landmarks in 3D motion captured images (4D). Int J Oral Maxillofac Surg. 2013 Jan;42(1):9–18.

44. Kim MJ, Oh TS. A nasolabial fold reset technique for enhancing midface lifts in facial reanimation: three-dimensional volumetric analysis. J Craniomaxillofac Surg. 2020 Feb;48(2):162–9.

45. da Costa MGST, Marahnao-Filho PAM, Santos IC, et al. Parotidectomy-related facial nerve lesions: proposal for a modified Sunnybrook Facial Grading System. Arq Neuro Psiquiatr. 2019 Jul;77(7):466–9.

46. Codari M, Pucciarelli V, Stangoni F, et al. Facial thirds-based evaluation of facial asymmetry using stereophotogrammetric devices: application to facial palsy subjects. J Craniomaxillofac Surg. 2017 Jan;45(1):76–81.

47. Slice DE. Geometric morphometrics. Annu Rev Anthropol. 2007 Oct;36(1):261–81.

48. Gibelli D, Tarabbia F, Restelli S, et al. Three-dimensional assessment of restored smiling mobility after reanimation of unilateral facial palsy by triple innervation technique. Int J Oral Maxillofac Surg. 2020 Apr;49(4):536–42.

49. Katsumi S, Esaki S, Hattori K, et al. Quantitative analysis of facial palsy using a three-dimensional facial motion measurement system. Auris Nasus Larynx. 2015;42(4):275–83.

50. Kim SW, Heller ES, Hohman MH, et al. Detection and perceptual impact of side-to-side facial movement asymmetry. JAMA Facial Plast Surg. 2013 Nov-Dec;15(6):411–6.

51. Wang TT, Wessels L, Hussain G, et al. Discriminative thresholds in facial asymmetry: a review of the literature. Aesthet Surg J. 2017 Apr;37(4):375–85.

52. Braig D, Bannasch H, Stark GB, et al. Analysis of the ideal muscle weight of gracilis muscle transplants for facial reanimation surgery with regard to the donor nerve and outcome. J Plast Reconstr Aesthet Surg. 2017 Apr;70(4):459–68.

53. Guarin DL, Yunusova Y, Taati B, et al. Toward an automatic system for computer-aided assessment in facial palsy. Facial Plast Surg Aesthet Med. 2020 Feb;22(1):42–9.

54. Mothes O, Modersohn L, Volk GF, et al. Automated objective and marker-free facial grading using photographs of patients with facial palsy. Eur Arch Otorhinolaryngol. 2019 Dec;276(12):3335–43.

55. Lee DY, Kim HS, Kim SY, et al. Comparison between subjective scoring and computer-based asymmetry assessment in facial nerve palsy. J Audiol Otol. 2019 Jan;23(1):53–8.

56. Taeger J, Bischoff S, Hagen R, et al. Utilization of smartphone depth mapping cameras for app-based grading of facial movement disorders: development and feasibility study. JMIR mHealth uHealth. 2021 Jan;9(1):e19346.