

Validation of a New Graphic Facial Nerve Grading System: FAME Scale

Pawan T. Ojha, Shashank Nagendra¹, Afroz Ansari¹, Nikhil Dhananjay Kadam¹, Ajay Mathur², Neeraja Gopinathan³, Nishu Ojha⁴, Hardik Patel⁵, Akshay Bansode⁵, Orpah Kale⁶, Kishan P. Morwani⁶, Digvijay Jagtap, Sumant Kumar⁷, Vinod Vij⁸

Departments of Neurology, ³Dentistry, ⁴Neurophysiology, ⁵Physiotherapy, ⁶ENT, ⁷Research Co-Ordinator and ⁸Plastic Surgery, Fortis Hiranandani Hospital, Vashi, Navi Mumbai, Maharashtra, ¹Departments of Neurology and ²Dentistry, Grant Medical College and JJ Hospital, Mumbai, Maharashtra, India

Abstract

Background and Objective: Most of the existing qualitative facial nerve grading systems are very subjective while the quantitative grading systems are more complex, require longer data input time and specific software. There is a need for having a scoring system with graphic criteria to improve the subjectivity, reliability and convenience. We aimed to develop and validate such a reliable graphic scale for use in Bell's palsy. **Methods:** Face videos of patients with unilateral facial paralysis were recorded using smartphones and analyzed for six items including five voluntary facial movements apart from complications of facial palsy (synkinesis, hyperkinesis, and contracture). 15 videos were used for pilot study, 75 for the development of scale and 110 for its validation. Each video was rated on two separate occasions by 3 independent raters, a score of 0-4 was assigned to each item using the graphic scoring criteria, and a composite score was obtained (range 0–24). Five disease severity categories: normal (score 0), mild (score 1–6), moderate (score 7-12), severe (score: 13–18) and profound facial weakness (score: 19-24). **Results:** The proposed scale and its component items had high inter-rater and intra-rater reliability (Kappa >0.7). Good correlation (Pearson co-efficient >0.7) was seen among the voluntary movements. The proposed scale is a valid tool to score motor deficits and complications of facial palsy. **Conclusions:** The proposed scale is a valid and reliable graphic scale to describe facial motor dysfunction and its secondary defects.

Keywords: Bell's palsy, facial complications, facial nerve grading system, facial palsy scale

INTRODUCTION

The outcome of Bell's palsy can often be predicted based on the severity at onset and early clinical course.^[1] While mild disease is generally associated with good outcomes, poor recovery and synkinesis are frequently encountered with more severe forms. Reliable assessment of disease severity is vital for monitoring response to treatment.

Many assessment scales proposed during the last few decades being subjective, are therefore limited by the expertise of the professionals administering them. The House–Brackmann facial nerve grading system^[2] is one of the most widely used grading systems, but it has inherent weakness when assessing synkinesis and differential paralysis in various regions of face. Other grading systems were proposed to overcome these limitations e.g., Sunnybrook,^[3] Yanagihara,^[4] MoReSS,^[5] Nottingham scale,^[6] as well as a modified version of House – Brackmann system i.e., Facial Nerve Grading System 2.0 (FNGS2.0).^[7] Most of the grading systems usually lack graphic representations of scoring criteria and scores are assigned either by eye-balling or by rough estimation of percentage for defects in facial movements. To overcome these issues, many quantitative scales have been developed using techniques such as facial marker video analysis, digital facial motion analysis,^[8] video pixel data using artificial neural networks^[9] or automated quantitative grading of facial functions.^[10-17]

e-FACE^[18] is one such popular software-based system for use on i-PAD to score static, dynamic function and disfigurement with slider scale for a total of 16 items in patients with facial palsy. The requirements of specific software, hardware, longer data entry time for these quantitative scales might limit widespread accessibility. Also, in the resource-limited setting, a need for developing a paper-based semi-quantitative graphic scoring system was felt. While being convenient it should also be reliable and valid scale for assessment of unilateral facial palsy and secondary complications such as synkinesis. Towards this end, we propose a new graphic facial nerve grading scale, Facial Motor Evaluation Scale (FAME Scale), that allows reliable assessment of paralysis in different face regions.

Address for correspondence: Dr. Pawan T. Ojha, Fortis Hiranandani Hospital, Mini Seashore Rd, Sector 10, Vashi, Navi Mumbai, Maharashtra – 400 703, India. E-mail: ptojha@yahoo.co.in

Submitted: 19-Jul-2021 **Revised:** 04-Dec-2021 **Accepted:** 07-Dec-2021

Published: 06-Apr-2022

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

DOI: 10.4103/aian.aian_662_21

METHODS

This prospective study was conducted in a tertiary care hospital during the period of July 1, 2012 through June 30, 2017. Ethics approval from the Independent Ethics Committee, Navi Mumbai (approval number is IEC/007/2012) and informed consent from participating subjects for use of their videos or images for the purpose of the study and publishing were obtained prior to enrolment in the study. Patients above the ages of 18 years with unilateral idiopathic facial palsy were recruited, and those with facial scarring were excluded.

The development of concept of the scale involved identifying clinically important parameters such as voluntary facial movements and complications of facial palsy, their scoring criteria and a composite score. A panel of six members consisting of two neurologists, one physical therapist, dentist, ENT specialist and Plastic surgeon was constituted and they evaluated many candidate parameters on a scale of 1 to 3 (1 meaning non-essential, 3 meaning essential and 2 being of doubtful importance) to provide Content validity ratio (CVR). On the basis of CVR being 1, the panel chose in total six items for inclusion in the scale: five voluntary facial movements (eyebrow raising, gentle eye closure, firm eye closure, smile and cheek bulge) and complications of facial palsy (synkinesis, hyperkinesis or contracture). The severity of deficit in each of the voluntary facial movements was graded on a 0-4 points scale representing no, mild, moderate, severe and profound weakness respectively. The scoring criteria were based on specific facial appearances observed during serial visits in patients recovering from facial palsy. Complications were scored on a scale of 0-4 on the basis of severity and distribution of involvement. [Figures 1-3] The graphic representations of scoring criteria were created by first author assisted by a computer graphics designer.

Complications of facial palsy that may be seen sometimes during recovery phase are synkinesis, contracture or hyperkinesis. Facial synkinesis refers to inappropriate and unintended muscle movements in the face with certain voluntary facial expressions like eye movement and mouth movements.

Contracture is characterized by sustained unilateral contraction of the facial muscles associated with mild ipsilateral facial paresis. Hyperkinesis or post-paretic hemifacial spasm is involuntary twitching or spasm of facial muscles on the side of paresis.

Face videos of subjects were recorded and analyzed using smartphones with camera, with attention to adequate lighting and centring of face. 15 videos were analyzed in pilot testing and based on the mean and standard deviation (SD), a sample size of over 63 was arrived at. A larger number of subjects i.e., 75 were used for development phase of the scale. Based on the mean and SD of this group, sample size of 108 was derived for the validation phase. Therefore 110 videos were used for validation of the scale. The videos were rated by 3 independent

raters using the proposed scale, twice at a minimum time gap of 3 weeks between the ratings. 2 raters were neurology residents and third rater a dentist, all with 2 years clinical experience and they were trained for scoring by the first author, a neurologist with more than 20 years of experience. The raters scored each item of scale for a score of 0-4. They were advised to assign higher score in case of a dilemma.

For recording Eyebrow raising (ER), the subject was asked to raise eyebrows fully, with the extent of eyebrow elevation and forehead creases being observed. Grade 3 could be differentiated from Grade 4 by the appearance of slightest elevation of affected eyebrow without any forehead crease. Grade 4 was also assigned when only medial movement (without elevation) of affected eyebrow was caused by pull of the contralateral corrugator muscle. Grade 2 was assigned for moderate elevation of affected eyebrow causing slight forehead crease. Grade 1 was assigned for good elevation of forehead but the forehead crease on the affected side was less than normal.

For recording of Gentle Eye Closure (GC) the subject was asked to close eyes gently as in sleeping and with maximal force for Firm Eye Closure (FC). They were observed for speed, extent of eyelid movements and for palpebral gap. In addition to complete lack of movements, Grade 4 was also assigned when only the upper eyelid showed slight movement due to inhibition of levator palpebrae superioris. In this case, no movement of the lower eyelid was observed. Grade 3 was assigned when both eyelids showed movement but a large oval palpebral gap remained as upper eyelid failed to reach beyond midway. Grade 2 was assigned when upper eyelid reached beyond midway, leaving a small crescentic palpebral gap. Grade 1 Gentle eye closure was assigned when eye closure was complete but slower than fellow eye. Grade 1 of Firm eye closure was assigned when eye closure was complete but weak such that major part of eyelashes remained visible.

Smile (SM): Subject was asked to show maximum teeth and observed for cheek muscle movements, elevation of angle of mouth and symmetry of lip arching. Grade 3 was assigned for slightest cheek movement and grade 2 when it resulted in elevation of angle without visible arching of lips. Grade 1 was assigned when arching of lips appeared to be weak on affected side.

For assessment of Cheek Bulge (CB) subject was asked to forcefully blow air against closed mouth, observed for air leakage from gap between lips and cheek bulge symmetry as well. Grade 4 was assigned when air was noted to leak easily from a wide lip gap. Grade 3 weakness was assigned if air leakage occurred from a small (1-2 mm) gap in lips. Grade 2 was assigned if air leaked from an invisible gap and grade 1 was assigned when poor cheek bulge was seen despite there being no leakage of air.

Complications of Bell's palsy to be scored included synkinesis, hyperkinesis and contracture. Complications were scored in 2 areas of face i.e., around the eye and the mouth, each being assigned 1 point if there were mild and non-disfiguring

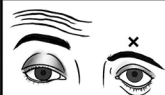

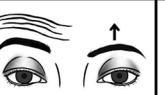
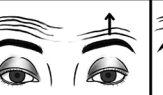




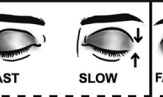


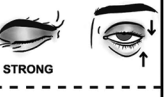

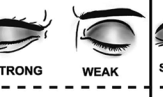




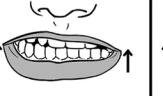





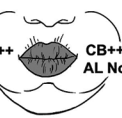
FACIAL MOTOR EVALUATION SCALE: FAME SCALE (Left side is shown weak)					
Items	Grade 4 (total) weakness	Grade 3 (severe)	Grade 2 (moderate)	Grade 1 (mild)	Grade 0 (normal)
Eyebrow raising (ER):	 No elevation ★	 Twitch ★ Minimal elevation on weak side	 Mid-range elevation No forehead crease on weak side	 Fewer creases on weak side. Good elevation	 Normal
Gentle eye closure (GC):	 No movement #	 Oval eye gap # Both eyelids move, upper not reaching beyond midway	 Crescentic eye gap Both eyelids move, upper reaches beyond midway	 Slow, but complete eye closure	 Normal
Firm eye closure (FC):	 No movement #	 Oval eye gap # Both eyelids move, upper not reaching beyond midway	 Crescentic eye gap Both eyelids move, upper reaches beyond midway	 Weak, but complete eye closure	 Normal
Smile (SM): (show maximum teeth)	 No movement	 Cheek twitch But no elevation of mouth angle	 Mouth angle rises No lip arching on weak side	 Reduced lip arch on weak side	 Normal
Cheek Bulge (CB): (blow air against closed mouth)	 Wide lip gap: Total air leak (AL)	 Thin lip gap: Major air leak	 No visible lip gap: Minimal air leak	 Poor cheek bulge No air leak	 Normal
Complication	Severe (disfiguring) complications in both areas	Severe in one area AND mild in the other	Severe in one area OR mild in both areas	Mild (non-disfiguring) complication in one area	None
FAME score (add all items) =					
★ Grade 4: If eyebrow moves inward but not up, pulled by normal side.					
# Grade 4: If only upper but not lower eyelid moves slightly due to levator inhibition.					

Figure 1: Pictorial representation of the FAME scale

complications, or a score of 2 if there were severe and disfiguring complications. The scores of both face areas were added to provide C score (Complication score) ranging from 0-4. The raters also provided detailed description of all type of complications in these two face regions.

A Composite score (FAME score) obtained by adding all the above scores. It ranged from 0-24, higher scores indicating more severe involvement. 5 categories of severity were defined based on FAME score: Category 1 (normal: score 0), category 2 (mild weakness: score 1-6), Category 3 (moderate weakness: score 7-12), Category 4 (severe weakness: score 13-18), Category 5 (profound weakness: score 19-24).

Videos were also scored on FNGS2.0 scale (range 4-24) in all 3 cohorts i.e., pilot study, development and validation groups. Construct validity of the scale was done by comparing FAME score with FNGS2.0 score (concurrent validity). It was also performed using Discriminant validity of FAME scores in all three cohorts. Factor analysis was done by principal component analysis in development cohort.

Those patients whose videos could be recorded at 2 additional visits at monthly intervals were scored to identify a very important aspect i.e., sensitivity to change.

Statistical analysis was done using SPSS 20 software. Intra-rater and Inter-rater reliability was measured by Cohen's kappa. Correlation between non-parametric variables was measured using Pearson's coefficient. Correlation was significant at 0.01 level (1-tailed). In descriptive analysis, tables and percentages were compared using mean, SD and ranges. *P* value of < 0.05 was taken as significant.

RESULTS

Average age of enrolled subjects in development cohort was 38.78 years (range 16–77 years) and 65% were males. 70% subjects were included within 2 weeks of onset of symptoms. It took an average of 35 seconds to record and 56 seconds to analyse the videos. All the three cohorts were normal in distribution as per Kolmogorov-Smirnov testing.

In the pilot study: mean composite score for all raters (FAME score) was 12.66 (SD = 4.98), standard error of mean was 1.28, standard error of measurement 1.07 and minimal detectable change was 0.94. The inter-rater reliability in pilot study was 0.95 (CI: 0.79-1.11).

In the development group: mean composite score for all raters (FAME score) was 10.41 (SD = 4.38), median being 10 (IQR 6,15, range 0-21). The standard error of mean was 0.505, standard error of measurement was 1.31 and minimal detectable change was 1.14. The number of subjects in five severity categories: Category 1 (normal) = 3, Category

2 (mild) = 17, category 3 (moderate) = 28, category 4 (severe) = 20, category 5 (profound) = 7. Median FAME scores for each of these categories were 0, 3, 10, 16, and 21.5 respectively. The three subjects with FAME score of 0 were the ones who had recovered completely during subsequent visits.

Strong and significant correlation (Pearson's coefficient >0.8) was found between all facial movements, except that eyebrow raising movement showed moderate correlation with cheek bulge (Pearson's coefficient 0.672). [Tables 1 and 2] No correlation was seen between voluntary facial motor deficits and complications (M and C scores). [Table 3]. Even though gentle (GC) and firm eye closure (FC) correlated well but their values were different by at least 1 grade in about 40% of subjects in all 3 cohorts. Keeping this in mind, we chose to keep both of these items in the composite score. Factor analysis of all individual facial parameters showed that the voluntary facial movements and complications had separate dimensions.

Excellent inter-rater and intra-rater reliability was found for all facial movements, complications, Kappa being >0.9. Composite score (FAME) too had high inter-rater and intra-rater reliability. The inter-rater reliability was 0.94 (CI: 0.86-1.02). Excellent correlation with FNGS2.0 score was noted (pearson 0.942). The FAME score and all individual parameters excluding Complications had excellent discriminant validity to differentiate between mild and severe degree of facial paresis. Complications could be seen with any grade of weakness.

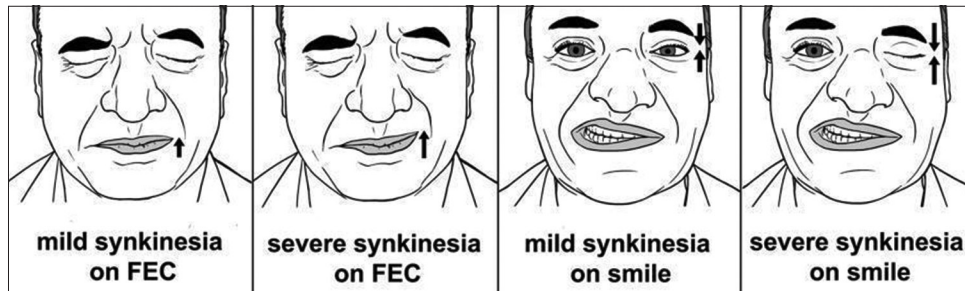


Figure 2: Pictorial representation of complications



Figure 3: Patients graded as per the FAME scale. (a) ER2, (b) ER4, (c) GC2, (d) GC4, (e) FC1, (f) FC2, (g) SM1, (h) SM2, (i) CB3, (j) CB4 [Refer to Figure 1 for grading]

12 subjects in development cohort had complications and they were seen in all severity categories except FAME categories 0 and 1 (mild paresis). Synkinesis were seen in all, contracture in 5 and hyperkinesis only in 1.

11 patients were assessed at 2 additional visits at monthly intervals to determine sensitivity of the scale. The scale was found to be sensitive as clinical recovery correlated with decreasing values of FAME score. [Figure 4]

These results were validated in another cohort of 110 patients. In this group: mean composite score for all raters (FAME score) was 9.6 (SD = 5.82), standard error of mean was 0.555, standard error of measurement was 1.28 and minimal detectable change was 1.12. Excellent correlation (correlation coefficient >0.8) between all voluntary facial movements and a moderate correlation between eyebrow raising and cheek bulge was observed here too. In the validation cohort, 26 subjects had complications. 24 had synkinesis, 8 contracture and 4 hyperkinesis. Excellent inter-rater and intra-rater reliability was found for all facial movements, flaccidity and complications, Kappa being >0.9. [Table 4]. Composite score had high inter-rater and intra-rater reliability in validation cohort (Kappa: 0.76 to 0.80). Excellent correlation with FNGS2.0 score was noted and discriminant validity was also validated in this group as well.

As the descriptive statistics and results of the two study groups were similar, the proposed scale was considered validated.

DISCUSSION

The unique anatomy and clinical evolution of facial nerve injury poses special challenges in grading facial nerve dysfunction. Deficits of muscles of facial expression can affect some regions more than others. The secondary defects or complications arising from aberrant facial nerve regeneration i.e., synkinesis, hyperkinesis, contracture, crocodile tears, abnormalities of taste sensation and hyperacusis also contribute significantly to disability. For this reason, documentation of motor deficits and complications involving different regions of face is crucial.

Many scales measuring facial nerve dysfunction have been proposed over last several decades but most have either not been tested rigorously or do not meet the criteria for an

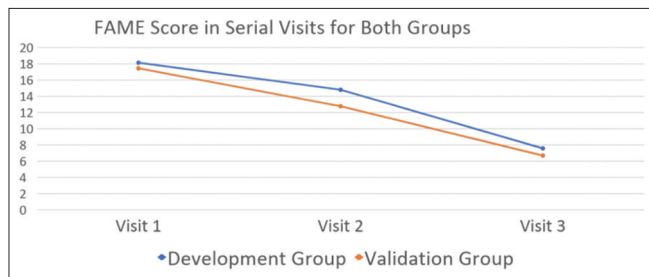


Figure 4: FAME scores for serial visits in Development group and Validation group

ideal scale. Many of these “traditional methods” require no specialized equipment, while the “Computer-based methods,” measure and quantify digital data objectively.

The most commonly used House-Brackmann Scale^[1] is a gross subjective ranked scale in which grade I to VI is assigned for the degree of paralysis and secondary deficit. Some of the limitations are low inter-observer reliability, lack of sensitivity

Table 1: Correlation of all items in Development group

	ER	GC	FC	SM	CB
ER					
Pearson Correlation	1	0.710**	0.800**	0.858**	0.660**
Sig. (1-tailed)		0	0	0	0
n	75	75	75	75	75
GC					
Pearson Correlation	0.710**	1	0.872**	0.749**	0.756**
Sig. (1-tailed)	0		0	0	0
n	75	75	75	75	75
FC					
Pearson Correlation	0.800**	0.872**	1	0.797**	0.757**
Sig. (1-tailed)	0	0		0	0
n	75	75	75	75	75
SM					
Pearson Correlation	0.858**	0.749**	0.797**	1	0.745**
Sig. (1-tailed)	0	0	0		0
n	75	75	75	75	75
CB					
Pearson Correlation	0.660**	0.756**	0.757**	0.745**	1
Sig. (1-tailed)	0	0	0	0	
n	75	75	75	75	75

**Correlation is significant at the 0.01 level (1-tailed)

Table 2: Correlation of all items in Validation group

	ER	GC	FC	SM	CB
ER					
Pearson Correlation	1	0.792**	0.808**	0.874**	0.688**
Sig. (1-tailed)		0	0	0	0
n	110	110	110	110	110
GC					
Pearson Correlation	0.792**	1	0.894**	0.810**	0.749**
Sig. (1-tailed)	0		0	0	0
n	110	110	110	110	110
FC					
Pearson Correlation	0.808**	0.894**	1	0.808**	0.739**
Sig. (1-tailed)	0	0		0	0
n	110	110	110	110	110
SM					
Pearson Correlation	0.874**	0.810**	0.808**	1	0.773**
Sig. (1-tailed)	0	0	0		0
n	110	110	110	110	110
CB					
Pearson Correlation	0.688**	0.749**	0.739**	0.773**	1
Sig. (1-tailed)	0	0	0	0	
n	110	110	110	110	110

**Correlation is significant at the 0.01 level (1-tailed)

Table 3: Correlation of M score with Complications and Flaccidity in both groups

	C	M
C		
Pearson Correlation	1	-0.103
Sig. (1-tailed)		0.189
n	75	75
M		
Pearson Correlation	-0.103	1
Sig. (1-tailed)	0.189	
n	75	75

**Correlation is significant at the 0.01 level (1-tailed)

	C	M
C		
Pearson Correlation	1	-0.038
Sig. (1-tailed)		0.348
n	110	110
M		
Pearson Correlation	-0.038	1
Sig. (1-tailed)	0.348	
n	110	110

*Correlation is significant at the 0.05 level (1-tailed). **Correlation is significant at the 0.01 level (1-tailed)

of region wise scoring and inclusion of synkinesis only at the lowest levels (best recovery) of the scale.^[19] A graphic version of House Brackmann Grading system was proposed, but not tested rigorously.^[20] To address these limitations, a revision Facial Nerve Grading Scale 2.0 was proposed.^[7]

In that system, facial movements are to be assessed in four regions i.e., brow, eyes, nasolabial fold, oral to assign score of 1 to 6 to degree of movement holding a graded scale against the face part being evaluated. Score of 6 is assigned for no movement, 5 for trace movement, 4, 3, 2, 1 respectively for <50%, >50%, >75% and full movement in each region. Ascribing percentages to movements of expression inherently introduces subjectivity in scoring. Secondary defects are scored subjectively on a scale of 0-3 for no, slight, obvious, disfiguring synkinesis, and contracture respectively. The final score ranging between 4 and 24 can be used to stratify patients in I-VI severity grades of House Brackmann Scale. The sensitivity and synkinesis evaluation for Bell's palsy of Facial Nerve Grading Scale 2.0 is better than the House-Brackmann Scale but Kappa for interobserver reliability still varies between 0.39 and 0.63.^[21] Thus, this version too suffers from subjectivity and is quite cumbersome.

Some other popular regional scales include Smith scale,^[22] Burres-Fisch scale,^[23] Nottingham Scale.^[6] The Yanagihara Grading System for Facial Palsy scale assigns scores of 0, 2 or 4 subjectively for each of 10 items.^[4] It also excludes scoring of secondary defects. Sunnybrook Facial Grading Scale is a sensitive scale that evaluates facial symmetry at rest, voluntary facial movements and synkinesis; each being

Table 4: Inter-Rater and Intra-rater Reliability in all three groups

Item	Reliability	15	75	110
ER	Inter-rater 12	0.904	0.948	0.781
ER	inter-rater 23	0.815	0.965	0.781
ER	inter-rater 31	0.904	0.948	0.884
ER	intra-rater 1	1	0.92	0.942
ER	intra-rater 2	0.717	0.861	0.966
ER	intra-rater 3	0.72	0.861	0.965
GC	Inter-rater 12	0.821	0.848	0.894
GC	inter-rater 23	1	1	0.857
GC	inter-rater 31	0.821	0.894	0.917
GC	intra-rater 1	0.821	0.79	0.917
GC	intra-rater 2	0.906	0.912	0.941
GC	intra-rater 3	0.906	0.912	0.976
FC	Inter-rater 12	0.81	0.833	0.854
FC	inter-rater 23	0.717	0.796	0.807
FC	inter-rater 31	0.902	0.852	0.879
FC	intra-rater 1	0.902	0.93	0.951
FC	intra-rater 2	0.717	0.838	0.854
FC	intra-rater 3	1	0.856	0.891
SM	Inter-rater 12	1	0.848	0.875
SM	inter-rater 23	1	0.915	0.852
SM	inter-rater 31	1	0.949	0.977
SM	intra-rater 1	0.912	0.949	0.977
SM	intra-rater 2	0.82	0.813	0.989
SM	intra-rater 3	0.824	0.848	0.977
CB	Inter-rater 12	1	0.898	0.953
CB	inter-rater 23	1	0.965	0.893
CB	inter-rater 31	1	0.965	0.941
CB	intra-rater 1	1	0.965	0.941
CB	intra-rater 2	0.818	0.825	0.929
CB	intra-rater 3	0.907	0.841	0.905
C	Inter-rater 12	0.744	0.965	0.808
C	inter-rater 23	0.911	0.865	0.573
C	inter-rater 31	0.659	0.907	0.641
C	intra-rater 1	0.659	0.955	0.829
C	intra-rater 2	0.911	0.957	0.879
C	intra-rater 3	1	0.91	0.809
FAME	Inter-rater 12		0.992	0.989
	inter-rater 23		0.994	0.986
	inter-rater 31		0.994	0.993
	intra-rater 1		0.995	0.995
	intra-rater 2		0.987	0.995
	intra-rater 3		0.99	0.995

evaluated on point scales, and a composite score (0 to 100) is generated.^[3] It has high intra-observer (ICC 0.83 to 0.98) and interobserver reliability (ICC 0.831 to 0.997).^[21] Main limitations are lack of graphic representations and that it scores only synkinesis among complications. A recent study^[24] proposed a mathematical method that assigns percentages to scored deficits in Facial nerve grading system 2.0.

Among the published computerised systems include Facial Analysis Computerized Evaluation (FACE),^[10] Facial Nerve Function Based on Area Analysis (OSCAR),^[11]

Table 5: Comparison of various major facial scales

Scale	Reliability	Scale details	Limitations/comments.
House Brackmann Score	Mean weighted kappa 0.67.	Global scale, 6 grades.	Although rapid, it is a subjective, discontinuous scale that inadequately considers the regional deficits and secondary defects. Marked intra-observer variation especially in assigning grade 3 and 4.
MoReSS score	ICC of 0.882, 95% confidence interval (0.821, 0.927)	Measures movements, rest symmetry, synkinesis and subjective scoring by patient. Total score is Mo12, Re8, S6, S10.	The scale is complex and subjective as deficits grading based on terms such as mild, moderate, severe.
Nottingham score	Component of variance 7%	Continuous scale, measures movements and secondary defects	Cumbersome to use, requires calculation of percentage deficits and applying complex formula for score.
Yanagihara	Used as a prognostic scale	It measures of 10 separate functions in facial muscles on a scale of 0 to 4.	The scoring is subjective as it relies on vague terms such as mild moderate, severe palsy. Although less accurate but is still accepted as the standard in Japan.
Sunnybrook	Reliability: mean intraclass correlation coefficient 0.63: good	Measures resting symmetry, five voluntary movements and synkinesis. Composite score=Voluntary score-(resting symmetry + synkinesis scores)	Scoring is subjective as it relies on terms such as slight, mild, almost complete. The methods are time-consuming, and the calculations of the linear measurement index are complex.
FNGS 2.0	Intra-class Correlation: 0.843 (CI: 0.701-0.968).	Measures voluntary movements deficits on a scale of 1-6 in 4 face regions and synkinesis on scale of 0-3, added to provide composite score.	Scores assigned on the basis of estimation of percentage deficits (25%, 50%, 75% etc.) in voluntary movements, causing inherent subjectivity. Synkinesis are scored globally and not region-wise.
Facogram	Mean intra class correlation is 0.72	Analysis of video pixel data using computerised artificial neural networks (ANNs).	It does not assess resting symmetry or synkinesis and cannot be used in bilateral weakness or in presence of strabismus, extraocular palsy.
Auto eFACE	The intraclass correlation coefficient is 0.97.	App based scale, for use in electronic gadgets. 16 items scored in a visual analogue scale format to assess resting symmetry, voluntary facial function, synkinesis and provide an overall facial disfigurement score.	Shortcomings of the tool: can be time consuming, depends on availability of an electronic device, lacks utility in patients with bilateral facial paralysis. Requires more upfront "learning" to establish definitions of different levels of function.
FAME (proposed)	Intra-class correlation: 0.876 (CI: 0.778-0.974).	Scores voluntary movements and synkinesis on graphic scale, minimizes subjectivity of scoring.	Difficulty in scoring bilateral palsy. Synkinesis can be scored separately for different face regions.

two-dimensional Peak Motus Motion Measurement System^[12] or three-dimensional measurements proposed by Frey's team.^[13] Some recent attempts to assess facial expression using specialised software include The Facial Action Coding System (FACS)^[14] and Automated Facial Analysis using feature tracking (AFA).^[15-17] They can provide information regarding maximal movement but cannot decipher deficits in facial expression. Secondary defects of facial nerve dysfunction have also not been uniformly incorporated into these systems. One of the recently proposed and popular scale is an electronic, clinician-graded facial function scale (e-FACE). It is a software-based system to assess static, dynamic function and disfigurement in patients with facial palsy using 16 items.^[18] It provides a disfigurement score and a graphic for communication and understanding. The auto-eFACE^[25] compared the clinician-graded e-FACE scale to machine learning-derived automated assessments to achieve conformity in facial palsy assessment and to compare the effectiveness of treatments. However, Auto-e-FACE scores need a machine and learning-based computer software which may not be easily available in rural areas in third world countries. There have been reviews^[26] which assess the clinical parameters involved in grading as well as the objectivity of various scales.

Even though these methods might appear appealing but the need for specialized equipment and software is very likely to preclude widespread access. Subjective methods might therefore be more practical and have larger accessibility. A simple graphic facial scoring system that is based on observable patterns of facial movements might be more convenient and applicable in general.

An ideal facial nerve grading system should be valid, reliable and accurate instrument that can be conveniently used in clinic or research laboratory. There should be good intra-rater reliability between ratings on video and ratings in clinic.^[7,18] It should score static, dynamic features as well as secondary defects in different regions of face. It should also provide graphic representation of the scoring criteria. On the contrary, most of the existing scales have inherent subjectivity in scoring deficits as they rely on vague terms like mild, moderate, severe palsy or assign approximate percentages (25,50,75% etc.) to deficits. Many scales do not score all motor complications in a region wise pattern. Most scales also do not provide graphic illustrations of scoring criteria, leaving it to the rater to interpret the descriptions. Authors felt the need to develop a scale that can be used to score facial expressions and motor complications

in a region wise manner using graphic illustrations that reflect various stages seen during recovery of facial motor deficits. In addition, the composite score needs to be easy to comprehend and derive. Our proposed facial scale meets most of the criteria for an ideal scale. It not only has high inter-rater and intra-rater reliability, but also some notable advantages over other scales. It has been developed on the basis of distinct facial appearances seen during the course of recovery. It scores not only synkinesis but also other motor complications i.e., hyperkinesis and contracture efficiently. The composite score is also easy to calculate and reports deficits in a continuous manner. Graphic scoring criteria will be much useful in reducing the subjectivity in scoring, a problem that is frequently encountered while using most scales as they rely either on vague terms like mild, moderate, severe palsy or assign approximate percentages to facial deficits. A comparison with some of the popular facial nerve grading scales is provided in Table 5.

Our effort to make the scale brief and convenient has introduced some apparent limitations. Our scale and composite score does not include flaccidity even though we validated a score for facial flaccidity. The F score can be used separately to score facial flaccidity if deemed necessary. Our scale does not score non-motor symptoms such as taste, salivation, tearing or subjective feelings. It also does not assess muscles in the regions of nose, chin or platysma, nor does it assess cases of bilateral facial palsy.

CONCLUSION

While the search for an ideal facial nerve grading system continues, our scale offers some distinct advantages. It is a convenient, graphic regional scoring system for use in unilateral Bell's palsy. It is a valid scale with good inter-observer and intra-observer reliability. We propose its routine use in grading motor deficits of unilateral facial palsy.

Key message

The proposed scale is a convenient, valid and reliable graphic facial nerve grading system to assess motor dysfunction and secondary defects of Bell's palsy.

Acknowledgements

Dr Neeta Mehta Shah, Dr Satish Khadilkar, Dr Nirmal Surya, Dr Shripad Pujari, Dr Amit Shah, Dr Kedar Mate, Dr J C Sharma, Saurin Vast, Kajal Wadkar

Declaration of patient consent

The authors certify that they have obtained all appropriate patient consent forms. In the form the patient(s) has/have given his/her/their consent for his/her/their images and other clinical information to be reported in the journal. The patients understand that their names and initials will not be published and due efforts will be made to conceal their identity, but anonymity cannot be guaranteed.

Abbreviations

ER (eyebrow raising), GC (Gentle eye closure), FC (Firm eye closure), SM (smile), CB (cheek bulge).

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

- Baugh RF, Basura GJ, Ishii LE, Schwartz SR, Drumheller CM, Burkholder R, et al. Clinical practice guideline: Bell's palsy executive summary. *Otolaryngol Head Neck Surg* 2013;149:656-63.
- House JW, Brackmann DE. Facial nerve grading system. *Otolaryngol Head Neck Surg* 1985;93:146-7.
- Ross BG, Fradet G, Nedzelski JM. Development of a sensitive clinical facial grading system. *Otolaryngol Head Neck Surg* 1996;114:380-6.
- Hato N, Fujiwara T, Gyo K, Yanagihara N. Yanagihara facial nerve grading system as a prognostic tool in Bell's palsy. *Otol Neurotol* 2014;35:1669-72.
- de Ru JA, Braunius WW, van Benthem PP, Busschers WB, Hordijk GJ. Grading facial nerve function: Why a new grading system, the MoReSS, should be proposed. *Otol Neurotol* 2006;27:1030-6.
- Murty GE, Diver JP, Kelly PJ, O'Donoghue GM, Bradley PJ. The Nottingham system: Objective assessment of facial nerve function in the clinic. *Otolaryngol Head Neck Surg* 1994;110:156-61.
- Vrabec JT, Backous DD, Djalilian HR, Gidley PW, Leonetti JP, Marzo SJ, et al. Facial nerve grading system 2.0. *Otolaryngol Head Neck Surg* 2009;140:445-50.
- Dulguerov P, Marchal F, Wang D, Gysin C. Review of objective topographic facial nerve evaluation methods. *Am J Otol* 1999;20:672-8.
- O'Reilly BF, Soraghan JJ, McGrenary S, He, S. Objective method of assessing and presenting the House-Brackmann and regional grades of facial palsy by production of a facogram. *Otol Neurotol* 2010;31:486-91.
- Neely JG, Cheung JY, Wood M, Byers J, Rogerson A. Computerized quantitative dynamic analysis of facial motion in the paralyzed and synkinetic face. *Am J Otol* 1992;13:97-107.
- Meier-Gallati V, Scriba H, Fisch U. Objective scaling of facial nerve function based on area analysis (OSCAR). *Otolaryngol Head Neck Surg* 1998;118:545-50.
- Linstrom CJ, Silverman CA, Colson D. Facial motion analysis with a video and computer system after treatment of acoustic neuroma. *Otol Neurotol* 2002;23:572-9.
- Frey M, Giovanoli P, Gerber H, Slameczka M, Stussi E. Three-dimensional video analysis of facial movements: A new method to assess the quantity and quality of the smile. *Plast Reconstr Surg* 1999;104:2032-9.
- Ekman, Paul and Wallace V. Friesen. "Facial action coding system: a technique for the measurement of facial movement. Investigator's Guide. Palo Alto, CA: Consulting Psychologists Press; 1978. p. 1-42.
- Cohn JF, Zlochower AJ, Lien J, Kanade T. Automated face analysis by feature point tracking has high concurrent validity with manual FACS coding. *Psychophysiology* 1999;36:35-43.
- Gaber A, Faher MF, Waned MA. Automated grading of facial paralysis using the Kinect v2: A proof of concept study. 2015 International Conference on Virtual Rehabilitation (ICVR). 2015. p. 258-64.
- Gaber A, Taher MF, Abdel Wahed MA. A pilot study on automated quantitative grading of facial functions. *Vibroengineering PROCEDIA. Journal of Vibroengineering* 2020;30:109-15.
- Banks CA, Bhamra PK, Park J, Hadlock CR, Hadlock TA. Clinician-graded electronic facial paralysis assessment: The e-FACE. *Plast Reconstr Surg* 2015;136:223e-30e.
- Brenner MJ, Neely JG. Approaches to grading facial nerve function. *Semin Plast Surg* 2004;18:13-22.
- Lazarini P, Mitre E, Takatu E, Tidei R. Graphic-visual adaptation of House-Brackmann facial nerve grading for peripheral facial palsy. *Clin Otolaryngol* 2006;31:192-7.
- Lee LN, Susarla SM, Hohman MH, Henstrom DK, Cheney ML, Hadlock TA. A comparison of facial nerve grading systems. *Ann Plast Surg* 2013;70:313-6.
- Smith IM, Murray JA, Cull RE, Slattery J. Facial weakness:

- A comparison of clinical and photographic methods of observation. Arch Otolaryngol Head Neck Surg 1991;117:906-9.
23. Burres S, Fisch U. The comparison of facial grading system. Arch Otolaryngol Head Neck Surg 1986;112:755-8.
 24. Bansal M, Shah A, Gosai B, Joshi K, Shah P. A Simple, Objective, and mathematical grading scale for the assessment of facial nerve palsy. Otol Neurotol 2020;41:105-14.
 25. Miller MQ, Hadlock TA, Fortier E, Guarin DL. The Auto-e-FACE: Machine learning-enhanced program yields automated facial palsy assessment tool. Plast Reconstr Surg 2021;147:467-74.
 26. Fattah AY, Gurusinghe ADR, Gavilan J, Hadlock TA, Marcus JR, Marres H, *et al.* Facial nerve grading instruments: Systematic review of the literature and suggestion for uniformity. Plast Reconstr Surg 2015;135:569-79.