

ARTICLE ONLINE FIRST

This provisional PDF corresponds to the article as it appeared upon acceptance.

A copyedited and fully formatted version will be made available soon.

The final version may contain major or minor changes.

## Sunnybrook Facial Grading System reliability in subacute stroke patients

Marco TRAMONTANO, Giovanni MORONE, Federica Maria LA GRECA, Valeria MARCHEGIANI, Angela PALOMBA, Marco IOSA, Giulia MUSTO, Marilia SIMONELLI

*European Journal of Physical and Rehabilitation Medicine* 2021 Jun 09

DOI: 10.23736/S1973-9087.21.06629-6

Article type: Original Article

© 2021 EDIZIONI MINERVA MEDICA

Article first published online: June 9, 2021

Manuscript accepted: May 31, 2021

Manuscript revised: May 5, 2021

Manuscript received: October 3, 2020

Subscription: Information about subscribing to Minerva Medica journals is online at:

<http://www.minervamedica.it/en/how-to-order-journals.php>

Reprints and permissions: For information about reprints and permissions send an email to:

[journals.dept@minervamedica.it](mailto:journals.dept@minervamedica.it) - [journals2.dept@minervamedica.it](mailto:journals2.dept@minervamedica.it) - [journals6.dept@minervamedica.it](mailto:journals6.dept@minervamedica.it)

## **Sunnybrook Facial Grading System reliability in subacute stroke patients**

Marco TRAMONTANO<sup>1\*§</sup>, Giovanni MORONE<sup>1§</sup>, Federica Maria LA GRECA<sup>1</sup>, Valeria

MARCHEGIANI<sup>1</sup>, Angela PALOMBA<sup>2</sup>, Marco IOSA<sup>3,1</sup>, Giulia MUSTO<sup>1</sup>, Marilia SIMONELLI<sup>1</sup>

<sup>1</sup> IRCCS Fondazione Santa Lucia, Rome, Italy

<sup>2</sup> Multidisciplinary Department of Medicine for Surgery and Orthodontics, University of Campania Luigi Vanvitelli, Naples, Italy.

<sup>3</sup> Department of Psychology, Sapienza University of Rome, Italy

§The Authors contributed equally.

\*Corresponding Author:

Marco Tramontano

IRCCS Fondazione Santa Lucia

Via Ardeatina 306, 00179 Rome, Italy

Tel +39 0651501420 Fax +39 0651501378

e.mail: m.tramontano@hsantalucia.it

**Running title:** Sunnybrook Facial Grading System and Stroke

**Background:** A lack of evidence on the reliability limits the utility of the Sunnybrook Facial Grading System (SFGS) in individuals affected by stroke both in clinical and research settings.

**Aim:** To examine the test-retest reliability and the inter-rater reliability of the SFGS in patients affected by subacute stroke.

**Design:** A repeated-assessments design (fifteen days apart) was used to examine the test-retest reliability of the SFGS.

**Setting:** Inpatient rehabilitation unit.

**Population:** Thirty-two inpatients with subacute stroke.

**Methods:** Thirty-two inpatients with subacute stroke participated in the test-retest reliability study.

All patients were video-recorded in sitting position and uncovered face for about two minutes, in static position to evaluate symmetry at rest, then assessing symmetry of voluntary movement and synkinesis. For the inter-rater reliability, fifteen experts in neurorehabilitation were selected as raters. The facial function was assessed using the Italian version of the SFGS was used as grading system on a recorded video of each patient. The inter-rater reliability was carried out through the ICC, Intraclass Correlation Coefficient, while the intra-rater reliability was assessed comparing for each rater the scores evaluated at T0 and T1 using Wilcoxon signed-rank test.

**Results:** Evaluations at T0 and T1 showed a significant correlation with a total ICC (T0 + T1) of 0.901 (95% CI 0.852-0.942; p-value <0.001). The ICC of T0 scores is 0.914 (95% CI 0.869-0.950; p-value <0.001), while the ICC of T1 scores is 0.886 (95% CI 0.829-0.995; p-value <0.001). The scores of the "voluntary movement symmetry" area are correlated with ICC of 0.916 (95% CI 0.835-0.958; p-value <0.001). The scores of the "resting symmetry" area show correlation with ICC of 0.567 (95% CI 0.277-0.762; p-value <0.001). The area of "synkinesis" shows a lower correlation with ICC of 0.175 (95% CI -0.180-0.489; p-value <0.001). The scores attributed to the patients in

the two different moments of observation resulted not significantly different from 12 out of 15 raters.

**Conclusions:** The SFGS resulted reliable in its overall score to assess the facial palsy also in patients affected by stroke.

**Clinical rehabilitation impact:** The findings of this study provide empirical evidence of reliable properties of the SFGS for assessing central facial palsy in patients with stroke.

**Keywords:** Sunnybrook, Stroke, Facial Palsy, Rehabilitation, Reliability, Psychometrics

## INTRODUCTION

Central and peripheral facial palsy is the most common disease of the cranial nerves and its incomplete recovery leads to social and psychological consequences [1-3].

Many different etiologies have been linked to facial paralysis (infectious, neurologic, congenital, neoplastic, trauma, systemic), [4] many clinical trials have been performed and have suggested that surgical and nonsurgical interventions can improve peripheral facial palsy [3, 5-7]. Although central facial palsy is a very common impairment after stroke, there are few scales on grading the facial function in central facial palsy in patients affected by central nervous system diseases. [8] One of the possible reasons of this limitation may be the lack of reliability in the functional grading system for central facial palsy. The reliability of all clinician-graded systems for facial palsy is influenced by interrater and intrater variability. The variability is not only based on the raters' subjectivity but also by the patient's comprehension of the mimic movements (often compromised due to neuropsychological disorders, such as apraxia). Facial palsy and asymmetry are common in stroke patients. [9] The incidence of stroke has been calculated per 1000 person-years [10]. Traditional symptoms of stroke include, amongst others, facial weakness [11], with significant effect on swallowing, taste, vision and hearing, and also on verbal and non-verbal communication [12].

Health professionals with varying expertise were recruited to achieve rater-independent assessment of facial function. Multidisciplinary management is necessary, given the variability of symptoms in comorbidity with facial paralysis; to promote a multidisciplinary approach an objective grading system is necessary, which can be used by the whole medical team. Furthermore, an assessment tool is necessary for evaluating and communicating the spontaneous course and the results of medical and surgical treatments of facial palsy [13].

The Sunnybrook Facial Grading System (SFGS) was recently recommended as a validated, reliable, and properly sensitive outcome measure to assess peripheral facial nerve dysfunctions [14-15] and has been recently translated and validated in Italian [16]. It provides a clinical score from 0 to 100, which combines static and dynamic assessments of facial muscles with the degree of synkinesis [16].

This system proved to have high internal consistency, elevated repeatability and good sensitivity to clinical changes [17].

Reliability and agreement are important issues in classification, scale and instrument development, quality assurance, and in the conduct of clinical studies [18]. Results of reliability and agreement studies provide information about the amount of error inherent in any diagnosis, score, or measurement, where the amount of measurement error determines the validity of the study results or scores.

In a neurorehabilitation field, many clinicians and health professionals, possessing a range of clinical background knowledge, degree of expertise, and training routinely assess and manage patients with central facial palsy. Our hypothesis is that the SFGS could be a reliable tool also for patients with central facial palsy.

For this reason, the aim of this study is to assess the inter-rater and intra-rater reliability of central facial palsy in neurorehabilitation setting in patients affected by subacute stroke.

## **MATERIALS AND METHODS**

### **Participants and Raters**

The study protocol was approved by the Local Ethic Committee of IRCCS Fondazione Santa Lucia (FSL). All patients enrolled in the study provided written informed consent. Thirty-two stroke patients were recruited and enrolled using a consecutive sampling approach from the Neurorehabilitation Units of FSL. This phase lasted from December 2014 to September 2016.

The inclusion criteria consisted of proven central facial palsy upon completion of the diagnostics, age  $\geq 18$  years, the capacity to understand movements on request or on imitation, the ability to maintain the sitting position and written informed consent to video recording.

The exclusion criteria were: patients with previous peripheral facial palsy, facial disfigurement, nasogastric tube, cognitive deficits affecting the capacity of patients to understand the task instructions (Mini-Mental State Examination score  $>24$ ); severe unilateral spatial neglect (positive results on the: Letter Cancellation Test. Barrage Test. Sentence Reading Test and Wundt-Jastrow

Area Illusion Test), severe aphasia, and presence of neurological, orthopedic or cardiac comorbidities (all of them clinically evaluated), significant visual acuity impairment caused by cataracts, diabetes, retinopathy and/or glaucoma.

Time course in recovery, gender, or ethnicity were not considered criteria for selection or exclusion.

Demographic and clinical characteristics were reported in Table 1.

Fifteen experts, with at least 5 years of experience in neurorehabilitation, with a specific post graduate education for the assessment tools, were selected as raters. Four physicians (two otorhinolaryngologists, one physiatrist and one neurologist) and eleven health professionals (seven speech therapists, two nurses and two physiotherapists). For the comparison between the different figures, two evaluators were randomly selected for each professional profile.

The study was conducted according to the Guidelines for Reporting Reliability and Agreement Studies (GRRAS) [18].

*Table 1 Demographic and clinical characteristics of the enrolled sample.*

<b>N° patients</b>	32
<b>Mean age</b>	51.3 ± 16.8 years
<b>Gender</b>	19 M
<b>Time since stroke</b>	57.4 ± 17.1 days
<b>Stroke location</b>	44% Left
<b>Stroke type</b>	50% Ischemic

### **Facial Palsy Grading of the Videos**

The facial function was assessed using the Italian version of the SFGS was used as grading system on a recorded video of each patient.

The SFGS provides a score ranging from 0 to 100, where 0 is the absence of movement and 100 the physiological function. The scale is divided into three areas: resting symmetry, symmetry of voluntary movement and synkinesis.

Before starting the rating process, each rater received an hour of training on the grading system using three videos with two neurological patients not involved in the study and one healthy participant.

## **Setting of the Video Recording**

The video recording was performed by two health professionals (speech therapists) not included in the group of raters (Clinical Research Associate: CRA) (Good Clinical Practice (GCP) (D.M. 15 luglio 1997, n. 162).

A SONY HDR-CX 280E camera was used for recording. All patients were video-recorded in sitting position and uncovered face. Each video lasted about two minutes: five seconds in static position to evaluate symmetry at rest (first area); later the patients were asked to perform the five movements to assess symmetry of voluntary movement (second area) and synkinesis (third area). Each movement was repeated three times and maintained for five seconds. The videos were presented on a laptop by the two CRA to the fifteen raters.

The set of all of 32 videos was presented twice (T0 and T1 after fifteen days) and assessed by the raters. The order of the videos was decided by drawing lots. The archive code identifying each specific participant was placed on the source documents by the raters while that videotape was being reviewed.

## **Statistical Analysis**

All statistical analyses were performed using IBM Statistical Package for the Social Sciences (SPSS), version 22. The inter-rater reliability was carried out through the Intraclass Correlation Coefficient ICC (2,1), using a two-way random effects model. All the ICC analyzes were performed, also reporting the 95% confidence interval (CI). Also the intra-rater reliability was assessed using ICC (2,1) comparing for each rater the scores evaluated at T0 and T1. According to previous studies [19,20] the values of ICC from 0.70 to 0.90 should be considered representing a moderate or good reliability (acceptable error), whereas those above 0.90 as strong reliability (minimal error). Standard error of the measurement was computed in accordance with Sainani [21] and reported. The correlation between the scores of each individual evaluation area was studied by combining the scores of the 15



evaluators (resting symmetry, symmetry of voluntary movement, synkinesis). Furthermore, the correlation of the scores assigned by the single operator to each item of the T0 evaluations was verified. For all the analysis, the significant threshold was set at 5%.

## RESULTS

Thirty-two patients (19 men), who met the inclusion criteria, were recorded. The average mean age was  $51.3 \pm 16.8$  years. Forty-four percent of patients had a right facial palsy, the remaining had left facial palsy. The overall average score of SFGS was  $50.3 \pm 26.9$ .

To assess inter-rater reliability, evaluations at T0 and T1 were analyzed, showing a significant correlation. The ICC of T0 scores was 0.914 (95% CI 0.869-0.950; p-value <0.001), while the ICC of T1 scores was 0.886 (95% CI 0.829-0.995; p-value <0.001), and the total ICC (T0 + T1) was 0.901 (95% CI 0.852-0.942; p-value <0.001) (Table 2).

*Table II Intraclass Correlation Coefficient Results computer for Inter-rater reliability, reported with 95% Confidence Interval, p-values, and Standard Error of the Measurement.*

	ICC	95% CI		p-value	SEM
		Lower bound	Upper bound		
<b>T<sub>0</sub></b>	0.914	0.869	0.950	< 0.001	8.44
<b>T<sub>1</sub></b>	0.886	0.829	0.995	< 0.001	10.03
<b>T<sub>0</sub> + T<sub>1</sub></b>	0.901	0.852	0.942	< 0.001	9.20

The scores of the "voluntary movement symmetry" area are correlated with ICC of 0.916 (95% CI 0.835-0.958; p-value <0.001). The scores of the "resting symmetry" area show correlation with ICC of 0.567 (95% CI 0.277-0.762; p-value <0.001). The area of "synkinesis" shows a lower correlation with ICC of 0.175 (95% CI -0.180-0.489; p-value <0.001) (Table 3).

Table III Correlation analysis

Area	ICC	95% CI		p-value
		Lower bound	Upper bound	
<b>1- Symmetry at rest</b>	0.567	0.277	0.762	< 0.001
<b>2- Symmetry of voluntary movement</b>	0.916	0.835	0.958	< 0.001
<b>3- Synkinesis</b>	0.175	-0.180	0.489	< 0.001

The analysis made on the weighted scores of the evaluators divided by profession shows that the highest correlation is between otorhinolaryngologists with an ICC of 0.970 (95% CI 0.939-0.985; p-value <0.001) and the lowest correlation is between nurses with an ICC of 0.855 (95% CI 0.724-0.926; p-value <0.001).

In both speech therapists and physiotherapist groups a strong statistical correlation was confirmed with an ICC, respectively, of 0.927 (95% CI 0.856-0.964; p-value <0.001) and 0.878 (95% CI 0.765-0.939; p-value <0.001).

The correlation between the assessments of the seven speech therapists recruited for the study was also analyzed separately (ICC = 0.934; 95% CI 0.869-0.967; p-value <0.001).

The intra-rater reliability was tested for each raters using ICC for comparing T0 and T1 scores. The scores attributed to the patients in the two different moments of observation resulted with a strong reliability (>0.9) for 13 out of 15 raters, and good (>0.7) for the other two (a nurse and a speech therapist), as shown in Table 4.

Table IV Intra-rater reliability: the Intra-class Correlation Coefficient (ICC), the relevant 95% confidence interval, the standard error of the measurement, and the p-value.

Raters	ICC	95% CI	SEM	p-value
Otorinolaryngol. 1	0.981	0.961 – 0.991	4.36	<0.001
Otorinolaryngol.2	0.978	0.955 – 0.989	3.85	<0.001
Neurologist 1	0.944	0.888 – 0.972	6.27	<0.001
Physiatrist 1	0.943	0.887 – 0.972	6.41	<0.001
Physioterapist 1	0.923	0.849 – 0.962	8.44	<0.001
Physioterapist 2	0.950	0.901 – 0.975	5.36	<0.001
Nurse 1	0.887	0.782 – 0.943	9.57	<0.001
Nurse 2	0.975	0.949 – 0.976	5.21	<0.001
Speech ther. 1	0.942	0.885 – 0.971	5.98	<0.001
Speech ther. 2	0.922	0.847 – 0.961	8.44	<0.001
Speech ther. 3	0.930	0.863 – 0.965	7.91	<0.001
Speech ther. 4	0.963	0.926 – 0.982	6.54	<0.001
Speech ther. 5	0.961	0.921 – 0.981	5.30	<0.001
Speech ther. 6	0.959	0.917 – 0.980	5.98	<0.001
Speech ther. 7	0.732	0.518 – 0.859	14.00	<0.001
Mean values T0 vs. T1	0.998	0.992 – 0.998	1.23	<0.001

## DISCUSSION

The aim of this study was to assess the reliability of SFGS in patients with central facial palsy due to subacute stroke. According to Kanerva [22] and Neely [14], SFGS has been intuitively easy to use, even for naive raters, and it was clinically relevant. Our results show a strong reliability related to the symmetry of voluntary movement, a moderate reliability for symmetry at rest, and weak reliability for the area of synkinesis. The results are due, in accordance with the literature, because the symmetry of voluntary movement item is more simple to assess thanks to the professionals ability to assess muscle contraction [16], whereas the synkinesis was the most difficult item to assess [17].

The results of the synkinesis area could have been influenced by the disagreement between clinical practice and the scale's explanatory criteria. Raters usually consider synkinesis in the injured half-face while in this case the raters were asked to consider synkinesis in the healthy half-face.

A strong reliability was observed among speech therapists between T0 and T1. In general, our research showed a strong reliability between all the evaluations, showing a high repeatability.

According to previous studies [16-18,22], the area of evaluation with the highest degree of reliability is the "symmetry of voluntary movement" despite the wider range of scores attributable.

The reliability of scores related to symmetry assessments at rest and synkinesis, was weaker than that of second area [16-18,19]. For the evaluation of symmetry at rest, only some raters exploited the still image; selecting a frame to exploit the freeze frame could influence the quality of the video [23] also selecting it manually involves a random selection (not repeatable among all raters).

The small differences observed between the same professional categories could be accounted to the different clinical and professional experience of the two compared classes. For example, the otorhinolaryngologist deals with the evaluation of the specific functional deficit in head and neck area. The intra-rater analysis showed that differences between the first and the second observation were not statistically significant, with a strong reliability for 13 out of 15 raters, and good reliability for the other two (a nurse and a speech therapist).

The inter-rater reliability of our study performed on patients with stroke are slightly lower to that computed by the study regarding the Italian validation of SFGS on patients with unilateral peripheral facial palsy (0.901 vs. 0.96, respectively) [16]. In both the studies the lowest ICC was found for the assessment of synkinesis. The intra-rater reliability were similar in the two studies (0.99 vs. 0.98, respectively). Also in our study on patients with stroke, as reported in that study on patients with unilateral peripheral facial palsy [16], the psychometric properties of the Italian version of SFGS have been influenced by the experiential background of the assessors. It suggests that clinicians working with patients with stroke, that could be not expert in the administration of SFGS, should be trained for increasing the reliability of their scoring, and that the overall score should be considered as more reliable than the subscores.

Neither the original article on the SFGS [15] nor that on its Italian validation [16] reported the standard error of the mean. Our study reported a SEM generally ranging from 3.85 to 9.57, (with only one speech therapist showing a higher SEM of 14) a value representing the possible standard deviation of the population of patients with stroke. Considering a mean values of about 50 (as that of our sample), the SEM results lower than 20% of the mean.

## LIMITATIONS

We acknowledge some limitations of the present study; firstly, the random selection of different frame to assess symmetry at rest, as already explained. Furthermore, it must be considered that the evaluation was not carried out in a clinical setting but through a video recording and this could omit some details due to the bias in mediated video analysis [24]. Moreover, the Video modality probably reduced real variability of the approaches to cognitive impairment by side of the every rater; this element can lead to underestimation of implications in terms of method of assessment administration which should be standardized for example through an administration manual or codified instructions. Another limitation is that the sample size of patients and professionals is doubtful according to the COSMIN, despite the used statistics was quite robust against the sample size.

Finally, we did not test the contribute of each item using Item response theory or assessing the influence of specific items on the reliability of the single area, by Fleiss's K or other analysis, because this study was focused on the reliability of SGF between and within clinical staff working with patients with stroke, being the psychometric properties of this scale already evaluated in previous studies [14-16].

## CONCLUSIONS

The SFGS showed an overall inter-rater reliability and an intra-rater reliability strong in 13 professionals out of 15 (and moderate or good in the other 2). The most reliable domain was the symmetry of voluntary movement, whereas symmetry at rest and especially synkinesis showed a

lower reliability, suggesting cautions in the use of these subscores. In conclusion, the SFGS resulted reliable in its overall score to assess the facial palsy also in patients affected by stroke.

## ACKNOWLEDGMENTS

We would thank Giulia Zerbini for the language editing.

## REFERENCES

1. De Swart BJ, Verheij JC, Beurskens CH. Problems with eating and drinking in patients with unilateral peripheral facial paralysis. *Dysphagia* 2003;18:267–273
2. Gilden DH Clinical practice. Bell's Palsy. *N Engl J Med* 2004;351:1323–1331
3. Baricich A, Cabrio C, Paggio R, Cisari C, Aluffi P. Peripheral facial nerve palsy: how effective is rehabilitation?. *Otol Neurotol*. 2012;33(7):1118-1126. doi:10.1097/MAO.0b013e318264270e
4. Melvin T-A N, Limb CJ. Overview of facial paralysis: current concepts. *Facial Plast Surg* 2008; 24:155–16
5. Paolucci T, Cardarola A, Colonnelli P, Ferracuti G, Gonnella R, Murgia M, Santilli V, Paoloni M, Bernetti A, Agostini F, Mangone M. Give me a kiss! An integrative rehabilitative training program with motor imagery and mirror therapy for recovery of facial palsy. *Eur J Phys Rehabil Med*. 2020;56(1):58-67. doi: 10.23736/S1973-9087.19.05757-5.
6. Fu X, Tang L, Wang C, Li M, Wu H, Li J, Ma Q, Yang W. A Network Meta-Analysis to Compare the Efficacy of Steroid and Antiviral Medications for Facial Paralysis from Bell's Palsy. *Pain Physician*. 2018;21(6):559-569. PMID: 30508985.
7. Azizzadeh B, Frisenda JL. Surgical Management of Postparalysis Facial Palsy and Synkinesis. *Otolaryngol Clin North Am*. 2018;51(6):1169-1178. doi: 10.1016/j.otc.2018.07.012.

8. Fattah AY, Gurusinghe AD, Gavilan J, Hadlock TA, Marcus JR, Marres H, Nduka CC, Slattery WH, Snyder-Warwick AK; Sir Charles Bell Society. Facial nerve grading instruments: systematic review of the literature and suggestion for uniformity. *Plast Reconstr Surg.* 2015;135(2):569-79.
9. Schimmel M, Ono T, Lam OL, Müller F. Oro-facial impairment in stroke patients. *J Oral Rehabil.* 2017;44(4):313-326. doi: 10.1111/joor.12486.
10. Feigin VL, Lawes CM, Bennett DA, Anderson CS. Stroke epidemiology: a review of population-based studies of incidence, prevalence, and case-fatality in the late 20th century. *Lancet Neurol* 2003; 2: 43–53
11. Lisabeth LD, Brown DL, Hughes R, Majersik JJ, Morgenstern LB. Acute stroke symptoms, comparing women and men. *Stroke* 2009; 40: 2031–2036
12. Konecny P, Elfmark M, Urbanek K. Facial paresis after stroke and its impact on patients' facial movement and mental status. *J Rehabil Med.* 2011;43(1):73-75. doi:10.2340/16501977-0645
13. House JW. Facial nerve grading systems. *Laryngoscope* 1983;93: 1056–69.
14. Neely JG, Cherian NG, Dickerson CB, Nedzelski JM. Sunnybrook facial grading system: reliability and criteria for grading. *Laryngoscope* 2010;120:1038–1045
15. Ross BG, Fradet G, Nedzelski JM. Development of a sensitive clinical facial grading system. *Otolaryngol Head Neck Surg;* 1996;114(3):380-6.
16. Pavese C, Tinelli C, Furini F, Abbamonte M, Giromini E, Sala V, De Silvestri A, Cecini M, Dalla Toffola E. Validation of the Italian version of the Sunnybrook Facial Grading System. *Neurol Sci,* 2013;34: 457-463.
17. Kayhan FT, Zurakowski D, Rauch SD. Toronto facial grading system: interobserver reliability. *Otolaryngol Head Neck Surg* 2000; 122:212–215
18. Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hrobjartsson A, Robertsg C, Shoukri M, Streiner DL. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *Journal of Clinical Epidemiology* 2011;64: 96-106

19. Hahn EA, Cella D, Chassany O, et al. Precision of health-related quality-of-life data compared with other clinical measures. *Mayo Clin Proc.* 2007;82(10):1244-1254. doi:10.4065/82.10.1244
20. Fayers PM, Machin D. *Quality of Life. The assessment, analysis and reporting of patient-report outcomes.* Wiley Blackwell, Oxford, UK; 2016
21. Sainani KL, Reliability Statistics. *PM R* 2017;9(6):622-628.
22. Kanerva M, Poussa T, Pitkaranta A. Sunnybrook and House-Brackmann facial grading systems: intrarater repeatability and interrater agreement, *Otolaryngol Head Neck Surg.* 2006; 135:856-871.
23. Seshadrinathan K, Soundararajan R, Bovik AC, Cormack LK. Study of subjective and objective quality assessment of video. *IEEE Trans Image Process.* 2010;19(6):1427-41. doi: 10.1109/TIP.2010.2042111.
24. Shapiro-Luft D, Capella JN. Video content in web surveys: effects on selection bias and validity. *Public Opin Q.* 2013;77:936–61

The authors report no involvement in the research by the sponsor that could have influenced the outcome of this work.

The study was Partially supported by the Italian Ministry of Health (Ricerca Corrente)

MS, MT, GM have given substantial contributions to the conception or the design of the manuscript, author FLG, VM, MI, AP to acquisition, analysis and interpretation of the data. All authors have participated to drafting the manuscript, author MT, GM and MS revised it critically. All authors read and approved the final version of the manuscript.

The authors acknowledge all healthcare professionals involved in the present study for their support.



## Supplementary Digital Material

Download supplementary material file: [Eur J Phys Rehabil Med-6629\\_Supplementary Digital Material1\\_V1\\_2021-01-18.pdf](#)