

Validation of the Turkish Version of the Facial Nerve Grading System 2.0

Original Investigation

Erdem Mengi¹ , Cüneyt Orhan Kara¹ , Fazıl Necdet Ardiç¹ , Fevzi Barlay¹ , Taylan Çil¹ , Kübra Aktan¹ , Hande Şenol²

¹Department of Otorhinolaryngology, Head and Neck Surgery, Pamukkale University School of Medicine, Denizli, Turkey

²Department of Biostatistics, Pamukkale University, School of Medicine, Denizli, Turkey

Abstract

Objective: We aimed to adapt the Facial Nerve Grading System 2.0 (FNGS 2.0) to Turkish and to investigate the validity and the reliability of the Turkish version.

Methods: The original FNGS 2.0 was translated into Turkish and validated by international standards. Six Turkish physicians, three specialists and three residents, independently rated the videos, two times each, of 40 adult facial palsy patients. Inter-rater and intra-rater reliability were assessed using the intraclass correlation coefficient (ICC) and Cronbach's alpha coefficient. As another indication of reliability, "generalizability" was also evaluated. For all analyses, a p value of <0.05 was considered statistically significant.

Results: ICC and Cronbach's alpha coefficients for the inter-rater reliability of the total score of the FNGS 2.0 were 0.970 and 0.979 for the first assessment, 0.973 and 0.979 for the second assessment, respectively. The intra-rater reliability ICC results for the total score of the FNGS 2.0 were 0.95, 0.976, 0.982, 0.956, 0.96 and 0.931 for the six raters, respectively. The generalizability coefficient was found as G=0.894.

Conclusion: In this study, we adapted the FNGS 2.0 to Turkish, and confirmed its reliability and validity as a facial palsy scale. The Turkish version of the FNGS 2.0 can be safely used to assess.

Keywords: Facial palsy, facial nerve, Facial Nerve Grading System 2.0, House-Brackmann scale

Introduction

Various grading systems have been developed to assess the severity and the progress of patients with peripheral facial paralysis (PFP). Such scales are highly important for the accurate and standard assessment of diagnostic and therapeutic efficacy, while also establishing a common language around the world. These scales are based on the observations of the physician examining the patient and are non-parametric in their approach.

Currently, the most commonly used system is the "House-Brackmann Facial Grading System (HB)", which grades patients on a scale of 1 to 6 based on the degree to which facial function is affected. First described in 1983, the HB scale was recognized as the standard scale for the assessment of facial nerve functions by the American Academy of Otolaryngology-Facial Nerve Disorders (FND)

Committee in 1985 (1). There is no doubt that the HB scale is simple and practical, yet in time the scale has been observed to have some limitations. In particular, its inability to make a regional scoring, and its classification of several facial movements under the same grade, leads to inter-rater grading differences (2-4). These limitations have led to alternative grading systems being developed, such as the Sunnybrook, Yanagihara, and Sydney scales (5-7). To date, as many as 20 grading systems have been described, yet none have come to be used as commonly as the HB scale.

In order to respond to the shortcomings of the HB scale, the FND Committee introduced the "Facial Nerve Grading System 2.0 (FNGS 2.0)" in 2009 as a revision to the scale and recommended the use of this new system for patients with facial paralysis (8). Different from HB, this new scale assesses four facial

ORCID IDs of the authors:

E.M. 0000-0001-8813-2443;
C.O.K. 0000-0003-2219-4283;
F.N.A. 0000-0003-4230-3141;
F.B. 0000-0003-4347-3820;
T.Ç. 0000-0003-1999-4976;
K.A. 0000-0001-8622-7182;
H.Ş. 0000-0001-6395-7924.

Cite this article as: Mengi E, Kara CO, Ardiç FN, Barlay F, Çil T, Aktan K, et al. Validation of the Turkish Version of the Facial Nerve Grading System 2.0. Turk Arch Otorhinolaryngol 2020; 58(2): 106-11.

Corresponding Author:

Erdem Mengi, emengi@hotmail.com

Received Date: 02.02.2020

Accepted Date: 03.04.2020

Content of this journal is licensed under a Creative Commons Attribution 4.0 International License. Available online at www.turkarchotolaryngol.net



regions separately and also assesses the presence of synkinesis; with the total score of the patient being determined by the summing of each score (Table 1). In FNGS 2.0, the total score is graded over 4-24 points, with a higher score indicating greater severity. FNGS 2.0 has gained popularity in recent studies, however, there has, to date, been no study validating the Turkish version of this evaluation system. In the present study, we aimed to adapt the FNGS 2.0 into Turkish and to conduct the validity and reliability studies.

Methods

Ethical committee approval was obtained from the Pamukkale University Clinical Studies Ethics Committee before the research was started (60116787-020/24509). Upon approval, the linguistic validity study for the Turkish version of the FNGS 2.0 was initiated.

Turkish Adaptation

In the initial stage, the original system was translated into Turkish by two researchers, independently of each other. The two researchers then came together and developed the final version of the translation. The Turkish translation was back-translated into English by a third researcher experienced in neurotology. Following this, the original text, the Turkish translation and the English back-translation were examined by the researchers. The differences were discussed, after which the researchers agreed on the final Turkish version to be used. The language validation study was concluded at this point since the system is a technical text. After completing the adaptation stage, the Turkish system was formatted to match the original system (Figure 1).

Selection of Patients

The study included 40 volunteer adult patients diagnosed with PFP, including both those with recently diagnosed PFP and those with chronic PFP. After patients provided their consent for participation, they were video-recorded with a camera in a natural and bright outpatient clinic environment by the same researcher. The video recordings were made using a Sony IILCE-6000A camera. During the video recordings, patients were asked to raise their eyebrows, close their eyes, show their teeth and give a wide smile, as the standard movements of the system. The assessments required for the study were made from these video recordings.

Assessment

To be representative of the potential users of the scale, the raters' group was composed of researchers with different levels of

experience in the field of Ear-Throat-Nose (ENT); namely, of six physicians—three specialists (two professors, one assistant professor) and three residency students. Before the assessment stage, a briefing was held to introduce the FNGS 2.0 scale to the researchers. During the meeting, the raters were trained on how to score the scale with grading exercises carried out using the recordings of five patients, the recordings of whom were subsequently excluded from the study.

During data collection, the raters assessed the patient videos independently. Assessments were made in groups of maximum 10 patients. The raters were allowed to pause and rewatch the patient videos during the assessment session. After initial assessments were completed, the assessments of the FNGS 2.0 were repeated 15 days later for the test-retest process with a randomized order of patients.

Statistical Analysis

The data were analyzed using the IBM Statistical Package for the Social Sciences software version 24.0 (IBM SPSS Corp.; Armonk, NY, USA). Continuous variables were expressed as mean±standard deviation, and as median, minimum-maximum, while categorical variables were expressed as numbers and percentages. Inter-rater reliability and intra-rater reliability were assessed using the intra-class correlation coefficient (ICC) and Cronbach's alpha coefficient. ICC 95% confidence interval (95% CI) was also presented. As another indication of reliability, "generalizability" was also evaluated. For all analyses, a p value of <0.05 was considered statistically significant. The ICC values were interpreted using the accepted criteria of <0.4=poor; 0.4-0.75=fair to good; ≥0.75=excellent. The reliability of the scale was considered good when the Cronbach's alpha coefficient was ≥0.7.

Results

The PFP patients assessed in the study consisted of 25 males (62.5%) and 15 females (37.5%), and the mean age was 45.26±14.08 (min-max: 21-71) years. Among the patients, 27 (67.5%) had acute PFP (≤ 3 months) and 13 (32.5%) had chronic PFP (> 3 months).

For the inter-rater reliability, ICC and Cronbach's alpha coefficients were calculated for both measurements based on the data obtained from the six raters. The results are presented in Table 2. The results of the inter-rater reliability ICC and Cronbach's alpha coefficients were 0.970 and 0.979 at assessment 1, 0.973 and 0.979 at assessment 2 for the total score; and 0.969 and 0.975 at assessment 1 and 0.967 and 0.971 at assessment 2 for the grade, each respectively. It can be concluded from these results that a statistically significant positive strong correlation existed between the raters at both measurements.

For intra-rater reliability, the results of assessments made by the raters at two different time points were compared with each other. Table 3 presents the ICC and Cronbach's alpha coefficients of the total score and grade, which were obtained separately from six raters. The study found an excellent correlation among all

Main Points

- The Facial Nerve Disorders Committee designed the FNGS 2.0 to overcome the limitations of the current facial grading systems.
- The present study adapted the FNGS 2.0 to Turkish and confirmed its reliability and validity as a facial paralysis scale.
- The Turkish version of the FNGS 2.0 can be safely used for the evaluation, follow-up, and reporting of patients with facial paralysis.

Table 1. Facial Nerve Grading System 2.0

Score	Region			
	Brows	Eyes	NLF	Oral
1	Normal	Normal	Normal	Normal
2	Slight weakness >75% of normal	Slight weakness >75% of normal	Slight weakness >75% of normal	Slight weakness >75% of normal
3	Obvious weakness >50% of normal	Obvious weakness >50% of normal	Obvious weakness >50% of normal	Obvious weakness >50% of normal
4	Resting symmetry Asymmetry at rest <50% of normal	Complete closure with mild effort Complete closure with maximal effort Asymmetry at rest <50% of normal	Resting symmetry Asymmetry at rest <50% of normal	Resting symmetry Asymmetry at rest <50% of normal
5	Trace movement	Trace movement	Trace movement	Trace movement
6	No movement	No movement	No movement	No movement
Secondary movement (global assessment)				
Score	Degree of movement			
0	None			
1	Slight synkinesis; minimal contracture			
2	Obvious synkinesis; mild to moderate contracture			
3	Disfiguring synkinesis; severe contracture			
Reporting: sum scores for each region and secondary movement				
Grade	Total score			
I	4			
II	5-9			
III	10-14			
IV	15-19			
V	20-23			
VI	24			

NLF: nasolabial fold

Table 2. Inter-rater reliability results of the Facial Nerve Grading System 2.0

	Assessment 1			Assessment 2		
	ICC	95% CI	Cronbach's alpha	ICC	95% CI	Cronbach's alpha
Brows	0.963	0.939-0.979	0.969	0.965	0.942-0.981	0.972
Eyes	0.975	0.959-0.985	0.977	0.981	0.969-0.989	0.981
Nasolabial fold	0.943	0.905-0.670	0.953	0.951	0.920-0.972	0.957
Oral	0.944	0.903-0.969	0.958	0.943	0.905-0.968	0.954
Secondary movement	0.756	0.616-0.857	0.760	0.823	0.721-0.896	0.828
Score	0.970	0.947-0.984	0.979	0.973	0.954-0.985	0.979
Grade	0.969	0.948-0.983	0.975	0.967	0.946-0.981	0.971

ICC: intraclass correlation coefficient; CI: confidence interval

raters in both the total score- and grade-related results. Based on these results, intra-rater assessments were found to have a statistically significant positive strong correlation.

For generalizability, the percentage of total variance explained by the obtained results were analyzed. The G coefficient calculated from the variance values was 0.894. Such a G coefficient,

Fasiyal Sinir Evreleme Sistemi 2.0				
Skor	Bölge			
	Kaş	Göz	NLO	Ağız
1	Normal	Normal	Normal	Normal
2	Hafif güçsüzlük >%75 normal	Hafif güçsüzlük >%75 normal Hafif eforla tam kapanıyor	Hafif güçsüzlük >%75 normal	Hafif güçsüzlük >%75 normal
3	Belirgin güçsüzlük >%50 normal İstirahat simetrik	Belirgin güçsüzlük >%50 normal Maksimum eforla tam kapanıyor	Belirgin güçsüzlük >%50 normal İstirahat simetrik	Belirgin güçsüzlük >%50 normal İstirahat simetrik
4	İstirahat asimetrik <%50 normal	İstirahat asimetrik <%50 normal Tam olarak kapanmıyor	İstirahat asimetrik <%50 normal	İstirahat asimetrik <%50 normal
5	Belli belirsiz hareket	Belli belirsiz hareket	Belli belirsiz hareket	Belli belirsiz hareket
6	Hareket yok	Hareket yok	Hareket yok	Hareket yok
İkincil hareket (Genel değerlendirme)				
Skor	Hareket derecesi			
0	Yok			
1	Hafif sinkinezi; minimal kontraktür			
2	Belirgin sinkinezi; hafif-orta kontraktür			
3	Yüz şeklini bozan sinkinezi; yaygın kontraktür			
Raporlama: Her bir alan ve ikincil hareket için toplam skor				
Evre	Toplam skor			
I	4			
II	5-9			
III	10-14			
IV	15-19			
V	20-23			
VI	24			
NLO, nazolabial oluk.				

Figure 1. The Turkish version of the Facial Nerve Grading System 2.0

Table 3. Intra-rater reliability results of the Facial Nerve Grading System 2.0

	ICC	95% CI	Cronbach's alpha
Rater 1			
Total score	0.950	0.904-0.973	0.950
Grade	0.911	0.830-0.953	0.915
Rater 2			
Total score	0.976	0.954-0.987	0.975
Grade	0.946	0.897-0.972	0.945
Rater 3			
Total score	0.982	0.965-0.990	0.981
Grade	0.962	0.927-0.980	0.961
Rater 4			
Total score	0.956	0.917-0.977	0.955
Grade	0.920	0.848-0.958	0.919
Rater 5			
Total score	0.960	0.924-0.979	0.962
Grade	0.947	0.900-0.972	0.947
Rater 6			
Total score	0.931	0.866-0.964	0.935
Grade	0.883	0.777-0.938	0.883

ICC: intraclass correlation coefficient; CI: confidence interval

being very close to 1, as the maximum value, indicated that the findings of this study were generalizable and reliable.

Discussion

Validation studies are required when scales to be used in the field of healthcare are translated into another language (9). This process involves three stages, namely, the assessment of the validity of the language, and reliability, and validity studies. The validity of language assessment is conducted to ensure that everybody understands the same thing. Since patients may be examined by different physicians, and treatment outcomes are assessed based on such scales, these scales need to have high reliability. A scale's validity is the degree to which it serves the desired purpose.

FPF is one of the most common patient groups encountered by ENT specialists. Furthermore, the follow-up and treatment of such patients may also be performed by specialists other than ENT, such as those from neurology, pediatrics, and physical medicine or rehabilitation departments. Accordingly, facial paralysis grading systems are significant in terms of creating a common language among different specialties.

An ideal facial paralysis grading system should allow for the regional assessment of the facial functions, make measurements both at rest and in motion, assess facial paralysis sequels, be suitable for clinical use, and have high inter-rater and intra-rater

reliability (10-13). Although the facial nerve grading systems developed to date are successful in assessing the normal functions of the facial nerve and severe paralysis, they may be incapable of assessing the mild functional abnormalities in between. The HB scale, which is the most commonly used scale, is known to have low inter-rater reliability, especially between grades II and IV (2, 7). A multicentered systematic review study by the Sir Charles Bell Society evaluating facial nerve grading systems reported that the most successful scale was the Sunnybrook facial grading system (10). However, the Sunnybrook scale requires too many calculations, restricting its use in practice. It seems unlikely, therefore, that the Sunnybrook scale will replace the HB scale.

The FND Committee designed the FNGS 2.0 to overcome the limitations of the current grading systems. This system enables to assess the eyebrows, the eyes, the nasolabial fold, and the mouth separately, and while the presence of synkinesis is assessed between grade 2 or 3 non-specifically in the HB scale, it is graded between 0 and 3 under a separate category in the FNGS 2.0 scale. The FNGS 2.0 scale aimed to differentiate between HB grade 3 and 4 patients in a more precise way. Lee et al. (14) compared the HB and FNGS 2.0 scales in 60 patients with Bell's palsy and reported the FNGS 2.0 scale to be more successful in grading and establishing a prognosis.

For a facial paralysis scale to have high inter-rater consistency, every rater should reach the same result in an assessment made using a standard facial paralysis grading system. The results from a reliable scale should also be repeatable. In other words, the same physician should obtain similar results from assessments made at different time points. Previous studies conducted using the original FNGS 2.0 have confirmed its reliability (8, 14, 15). The findings of the present study also indicate that the Turkish version of the FNGS 2.0 is reliable in terms of both consistency and repeatability.

Another reliability measure used for scales is generalizability (16). To determine generalizability, all sources of error are examined, and an attempt is made to determine how much the test results overlap with the purpose of the assessment. The G coefficient calculated for the FNGS 2.0 in the present study suggests that the results of the FNGS 2.0 scale are free from potential error factors, meaning that it gives only the results for facial paralysis that is desired to assess.

Conclusion

In conclusion, the present study adapted the FNGS 2.0 to Turkish and confirmed its reliability and validity as a facial paralysis scale. The Turkish version of the FNGS 2.0—produced as an output of this study—can be safely used to assess facial paralysis in patients, to monitor treatment efficacy and to report scientific studies.

Ethics Committee Approval: Ethics committee approval was received for this study from the Clinical Studies Ethics Committee of Pamukkale University (Approval Date: April 4, 2019; Approval Number: 60116787-020/24509).

Informed Consent: Informed consent was obtained from the patients who participated in this study.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept - E.M., C.O.K., F.N.A., F.B., T.Ç., K.A., H.Ş.; Design - E.M., C.O.K.; Supervision - C.O.K., F.N.A.; Materials - E.M., C.O.K., F.N.A., H.Ş.; Data Collection and/or Processing - E.M., C.O.K., F.N.A., T.Ç., K.A.; Analysis and/or Interpretation - E.M., C.O.K., F.N.A., H.Ş.; Literature Search - E.M.; Writing - E.M., C.O.K.; Critical Reviews - E.M., C.O.K., F.N.A., F.B., T.Ç., K.A.

Conflict of Interest: The authors have no conflicts of interest to declare.

Financial Disclosure: The authors declared that this study has received no financial support.

References

1. House JW, Brackmann DE. Facial nerve grading system. *Otolaryngol Head Neck Surg* 1985; 93: 146-7. [\[Crossref\]](#)
2. Rickenmann J, Jaquenod C, Cerenko D, Fisch U. Comparative value of facial nerve grading systems. *Otolaryngol Head Neck Surg* 1997; 117: 322-5. [\[Crossref\]](#)
3. Yen TL, Driscoll CL, Lalwani AK. Significance of House-Brackmann facial nerve grading global score in the setting of differential facial nerve function. *Otol Neurotol* 2003; 24: 118-22. [\[Crossref\]](#)
4. Croxson G, May M, Mester SJ. Grading facial nerve function: House-Brackmann versus Burres-Fisch methods. *Am J Otol* 1990; 11: 240-6.
5. Ross BG, Fradet G, Nedzelski JM. Development of a sensitive clinical facial grading system. *Otolaryngol Head Neck Surg* 1996; 114: 380-6. [\[Crossref\]](#)
6. Yanagihara N. Grading of facial palsy. Fisch U, editor. *Facial nerve surgery: Proceedings of the Third International Symposium on Facial Nerve Surgery*. Zurich, Switzerland: Kugler Medical Publications; 1976.p.533-5.
7. Coulson SE, Croxson GR, Adams RD, O'Dwyer NJ. Reliability of the "Sydney," "Sunnybrook," and "House Brackmann" facial grading systems to assess voluntary movement and synkinesis after facial nerve paralysis. *Otolaryngol Head Neck Surg* 2005; 132: 543-9. [\[Crossref\]](#)
8. Vrabc 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. [\[Crossref\]](#)
9. Karakoç FY, Dönmez L. [Basic principles of scale development.] *Tıp Eğitimi Dönyası* 2014; 40: 39-49. [\[Crossref\]](#)
10. Fattah AY, Gurusinghe AD, 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* 2009; 135: 569-79. [\[Crossref\]](#)
11. 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. [\[Crossref\]](#)
12. Kahn JB, Gliklich RE, Boyev KP, Stewart MG, Metson RB, McKenna MJ. Validation of a patient-graded instrument for facial nerve paralysis: The FaCE scale. *Laryngoscope* 2001; 111: 387-98. [\[Crossref\]](#)
13. Kang TS, Vrabc JT, Giddings N, Terris DJ. Facial nerve grading systems (1985-2002): Beyond the House-Brackmann scale. *Otol Neurotol* 2002; 23: 767-71. [\[Crossref\]](#)

14. Lee HY, Park MS, Byun JY, Chung JH, Na SY, Yeo SG. Agreement between the Facial Nerve Grading System 2.0 and the House-Brackmann grading system in patients with Bell palsy. *Clin Exp Otorhinolaryngol* 2013; 6: 135-9. [\[Crossref\]](#)
15. Henstrom DK, Skilbeck CJ, Weinberg J, Knox C, Cheney ML, Hadlock TA. Good correlation between original and modified House Brackmann facial grading systems. *Laryngoscope* 2011; 121: 47-50. [\[Crossref\]](#)
16. Webb NM, Shavelson RJ. Generalizability theory: overview. Everitt BS, Howell D, editors. *Encyclopedia of statistics in behavioral science*. Chichester: Wiley; 2005.p.599-612. [\[Crossref\]](#)