

Agreement between the Facial Nerve Grading System 2.0 and the House-Brackmann Grading System in Patients with Bell Palsy

Ho Yun Lee · Moon Suh Park · Jae Yong Byun · Ji Hyun Chung · Se Young Na · Seung Geun Yeo

Department of Otorhinolaryngology-Head and Neck Surgery, Kyung Hee University School of Medicine, Seoul, Korea

Objectives. We have analyzed the correlation between the House-Brackmann (HB) scale and Facial Nerve Grading System 2.0 (FNGS 2.0) in patients with Bell palsy, and evaluated the usefulness of the new grading system.

Methods. Sixty patients diagnosed with Bell palsy from May 2009 to December 2010 were evaluated using the HB scale and FNGS 2.0 scale during their initial visit, and after 3 and 6 weeks and 3 months.

Results. The overall intraclass correlation coefficient (ICC) was 0.908 ($P=0.000$) and the Spearman correlation coefficient (SCC) was 0.912 ($P<0.05$). ICC and SCC displayed differences over time, being 0.604 and 0.626, respectively, at first visit; 0.834 and 0.843, respectively, after 3 weeks; 0.844 and 0.848, respectively, after 6 weeks; and 0.808 and 0.793, respectively, after 3 months. There was a significant difference in full recovery, depending on the scale used (HB, $P=0.000$; FNGS 2.0, $P<0.05$). The exact agreements between regional assessment and FNGS 2.0 for the mouth, eyes, and brow were 72%, 63%, and 52%, respectively.

Conclusion. FNGS 2.0 shows moderate agreement with HB grading. Regional assessment, rather than HB grading, yields stricter evaluation, resulting in better prognosis and determination of grade.

Keywords. House-Brackmann scale, Facial Nerve Grading System 2.0, Bell palsy

INTRODUCTION

The House-Brackmann (HB) scale, which was first introduced in 1983 and selected as the standard by the Facial Nerve Disorders (FND) Committee in 1985, provides a gross impression by classifying facial motor function into 6 grades when diagnosing a patient with facial palsy. Due to the convenience and simplicity of the HB scale, it remains the most widely used facial nerve grading system [1].

The HB grading system, however, has various shortcomings, including its inability to accurately evaluate synkinesis and contracture. Thus, this scale cannot be used for systematic regional assessment and is limited in determining prognosis. Alternate grading systems include the “Yanagihara” and “Sunnybrook” scales [2,3]. In 2009, the FND Committee introduced the Facial Nerve Grading System 2.0 (FNGS 2.0), which compensates for the limitations of HB grading while retaining its simplicity [4]. The FNGS 2.0 determines the final grade by adding regional assessments of the brow, eye, nasolabial fold, and oral regions to the score assessing the impact of secondary movement. To date, however, few studies have compared these two grading systems in real patients and confirmed whether FNGS 2.0 has a significant advantage over the existing HB scale. We therefore analyzed the rate of agreement of the two scales and confirmed the properties and usefulness of FNGS 2.0.

• Received March 14, 2012
 Revised October 5, 2012
 Accepted October 8, 2012

• Corresponding author: **Seung Geun Yeo**
 Department of Otorhinolaryngology-Head and Neck Surgery, Kyung Hee University School of Medicine, 26 Kyungheedaero, Dongdaemun-gu, Seoul 130-701, Korea
 Tel: +82-2-958-8474, Fax: +82-2-958-8470
 E-mail: yeo2park@yahoo.co.kr

Copyright © 2013 by Korean Society of Otorhinolaryngology-Head and Neck Surgery.

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

MATERIALS AND METHODS

We prospectively evaluated patients who visited our hospital between May 2009 and December 2010 within 7 days of onset of unilateral facial palsy. Patients were excluded if they 1) presented with Bell palsy more than 1 week after onset; 2) were suspected of having Varicella zoster virus infection, based on physical and serologic examinations; 3) had a history of trauma or otologic surgery; 4) had other types of neurologic deficits; 5) had recurrent facial palsy; or 6) had a psychiatric disease. This study was approved by the Ethical Committee of Kyung Hee University Hospital, and all patients provided written informed consent.

All 60 included patients were treated with oral methylprednisolone for 12 days (64 mg/day for 4 days, 48 mg/day for 2 days, 40 mg/day for 2 days, 20 mg/day for 2 days, and 8 mg/day for 2 days). None of the patients received an antiviral agent. All patients were hospitalized for 7 days and followed-up as outpatients at 3 weeks, 6 weeks, and 3 months. During hospitalization, each patient was graded every 12 hours (twice per day) by both the HB scale and FNGS 2.0 grading system. Patients were also evaluated by both scales at each follow-up time point, with each patient evaluated at least twice for both systems by an otolaryngologist who understood both grading systems well. Patients were re-evaluated if there was any difference in the results of each grading system, and the final result was documented.

Agreement between the grading systems and their evaluation of patient prognosis were calculated by intraclass correlation coefficient (ICC), Spearman correlation analysis (SCC), and overall percentage agreement. Grade I at 3 months was considered the standard for complete recovery for evaluation of prognosis. Chi squared analysis was used to determine any difference between the two scales in judging recovery. Finally, the difference was confirmed by comparing the final results of the two grading

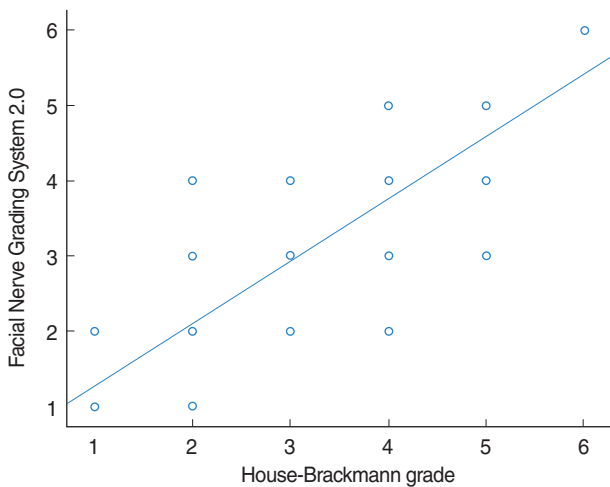


Fig. 1. Correlation between House-Brackmann grading Nerve Grading System 2.0 after complete evaluations

systems and the results of regional assessments of the brow, eye, nasolabial fold, and oral region. All statistical analyses were performed using SPSS ver. 18.0 (SPSS Inc., Chicago, IL, USA), and a $P < 0.05$ was considered to indicate statistical significance.

RESULTS

Subjects

The 60 patients with unilateral facial palsy comprised 28 males and 32 females, of mean age 43.0 ± 15.4 years (range, 15 to 69 years). Of these, 31 patients had Bell palsy on the right side and 29 on the left side. Mean time from occurrence to treatment was 2.87 ± 1.49 days.

Agreement between the HB and FNGS 2.0 grading systems

The correlation curve of the two systems is depicted in Fig. 1. The ICC between the two scales was 0.908, with agreement gradually increasing over time, being 0.604 at the initial visit; 0.834 and 0.844 after 3 and 6 weeks, respectively; and 0.808 after 3 months (Table 1). The SCC was very high, 0.912. The average match rate for the HB scale and FNGS 2.0 was 70.8%, 65.0% initially, 66.7% at 3 weeks, 71.7% at 6 weeks, and 80.0% at 3 months, showing that agreement between the two grading systems increased over time (Table 2).

Use of the two scales in patient prognosis

Recovery status was evaluated at 3 months, with grade I of both grading systems defined as complete recovery. We found that 43 patients (71.7%) achieved complete recovery on the HB grading system, compared with 38 (63.3%) for FNGS 2.0, a significant difference by chi squared analysis ($P < 0.05$).

Five patients were judged to have fully recovered by obtaining grade I in HB grading, but they were graded II on FNGS 2.0. Among them, one patient scored 2 points in the brow region, and the other four scored 2 points in the oral region in FNGS 2.0.

In contrast, one patient showed full recovery on the FNGS 2.0 but incomplete recovery (grade II) on HB grading; this patient, however, could not be further evaluated, since there were no additional records.

Table 1. Agreement using intraclass correlation coefficient (ICC) and Spearman correlation coefficient (SCC)

| | Overall | Initial visit | After 3 weeks | After 6 weeks | After 3 months |
|-----|---------|---------------|---------------|---------------|----------------|
| ICC | 0.908 | 0.604 | 0.834 | 0.844 | 0.808 |
| SCC | 0.912 | 0.626 | 0.843 | 0.848 | 0.793 |

All of the results of this table showed statistical significance ($P = 0.000, < 0.05$).

Table 2. Percentage of agreement between House-Brackmann scale and Facial Nerve Grading System 2.0

| | Overall (n=240) | Initial visit (n=60) | After 3 weeks (n=60) | After 6 weeks (n=60) | After 3 months (n=60) |
|-------------|--------------------|-------------------------|----------------------------|----------------------------|-----------------------------|
| % Agreement | 170 (70.8) | 39 (65) | 40 (66.7) | 43 (71.7) | 48 (80) |
| Av HB | 2.81 | 3.97 | 3.75 | 2.08 | 1.45 |
| I | 64 (26.7) | 0 (0.0) | 0 (0.0) | 21 (35.0) | 43 (71.7) |
| II | 40 (16.7) | 3 (5.0) | 8 (13.3) | 21 (35.0) | 8 (13.3) |
| III | 49 (20.4) | 15 (25.0) | 16 (26.7) | 10 (16.7) | 8 (13.3) |
| IV | 52 (21.7) | 23 (38.3) | 20 (33.3) | 8 (13.3) | 1 (1.7) |
| V | 34 (14.2) | 19 (31.7) | 15 (25.0) | 0 (0.0) | 0 (0.0) |
| VI | 1 (0.4) | 0 (0.0) | 1 (1.7) | 0 (0.0) | 0 (0.0) |
| Av FNGS 2.0 | 2.77 | 3.88 | 3.62 | 2.13 | 1.45 |
| I | 53 (22.1) | 0 (0.0) | 1 (1.7) | 14 (23.3) | 38 (63.3) |
| II | 54 (22.5) | 3 (5.0) | 5 (8.3) | 29 (48.3) | 17 (28.3) |
| III | 51 (21.3) | 13 (21.7) | 21 (35.0) | 12 (20.0) | 5 (8.3) |
| IV | 60 (25.0) | 32 (53.3) | 23 (38.3) | 5 (8.3) | 0 (0.0) |
| V | 21 (8.8) | 12 (20.0) | 9 (15.0) | 0 (0.0) | 0 (0.0) |
| VI | 1 (0.4) | 0 (0.0) | 1 (1.7) | 0 (0.0) | 0 (0.0) |
| > 1 grade | 70 (29.2) | 21 (35.0) | 20 (23.3) | 17 (28.3) | 12 (20.0) |

Values are presented as number (%).

% Agreement, percentage of exact agreement; Av HB, average of grade using House-Brackmann scale; Av FNGS 2.0, average of grading using Facial Nerve Grading System 2.0; > 1 grade, number of cases higher than one grade difference between two scales.

Comparison between regional assessments and final grade

The extent of matching of the HB and FNGS 2.0 systems was further evaluated by directly comparing the results of final grading and regional assessments (Table 3). On the HB scale, identical regional assessment results and final grade were observed for 67% of eye, 60% of oral, and 55% of brow regions. On the FNGS 2.0, regional assessments and final grade coincided for 72% of oral, 63% of eye, and 52% of brow regions.

DISCUSSION

The HB grading system, which was chosen as the standard by the FND Committee of the American Academy in 1985, remains the most widely used facial nerve grading system [1]. The motor function of the facial nerve can be rapidly and conveniently categorized into six HB grades. However, this system provides only gross impressions, thus limiting its usefulness. HB grading has several shortcomings, including: 1) regional assessment results are not properly reflected in the evaluation of overall grade of paralysis; 2) it can be difficult to evaluate patients with differing extent of weakness in each branch; 3) the HB system is not sensitive to changes in paralysis status; and 4) it can be difficult to compare clinically different recoveries [2,5-7]. To minimize these shortcomings, studies have explored the Yanagihara and Sunnybrook grading systems.

Introduced in 1976, the Yanagihara system is a regional scale

Table 3. Percentage of agreement between regional assessment and two scales (n=240)

| Grading system | Brow | Eye | Nasolabial fold | Oral |
|------------------------------------|------------|------------|--------------------|------------|
| House-Brackmann | 132 (55.0) | 161 (67.1) | 10 (43.8) | 144 (60.0) |
| Facial Nerve Grading System 2.0 | 124 (51.7) | 150 (62.5) | 104 (43.3) | 172 (71.7) |

Values are presented as number (%).

most frequently used in Japan. This system evaluates movements of 10 facial muscles, assigning each a score of 0-4 points, resulting in a maximum score of 40 points. When compared with HB grading, the kappa value was 0.64, indicating moderate agreement [6]. Scores of 0-6, 8-14, 16-22, 24-30, 32-38, and 40 points on the Yanagihara system correspond to grades VI, V, IV, III, II, and I, respectively, of HB grading [8]. In addition to not being widely used outside Japan, this system is too difficult for convenient use as the evaluation criteria are rather complex.

The Sunnybrook system grades paralysis by evaluating symmetry at rest and during voluntary movements, and, following a series of calculations, is used to gauge synkinesis on a 100-point scale. This system is sensitive in assessing changes in facial recovery [2] and has been reported highly reliable, with intra- and inter-rater reliability similar for beginners and experts [9]. However, it has been difficult for this system to replace HB grading.

FNGS 2.0, first introduced in 2009, was designed to overcome the limitations of existing grading systems. Although unquantifiable sensory deficits cannot be evaluated, detailed regional assessments of the brow, eye, nasolabial fold, and oral regions are possible, as are assessments of movement at any point of paralysis. In addition, the HB grading system ambiguously scored synkinesis as grade II or III, whereas the FNGS 2.0 system categorizes synkinesis in a separate category and scaled 0-3 depending on its severity, as well as evaluating final facial paralysis. The FNGS 2.0 also was designed to more clearly categorize grades difficult to distinguish in actual use, such as HB grades III and IV [4].

We found that the rate of complete recovery was lower on the FNGS 2.0 (63%) than on the HB grading system (71%). Moreover, among patients with incomplete recovery, grades II and III were distributed differently on the two scales. These differences are likely due to the more detailed analysis on the FNGS 2.0, including regional assessment and evaluation of synkinesis. The FNGS 2.0 can be considered superior to HB grading, due to its stricter evaluations. Thus, the FNGS 2.0 is not only more useful in distinguishing between grades III and IV, but in differentiating low grade paralysis such as grades I-III. Proper assessment of complete recovery requires a stricter assessment system. This is especially applicable when evaluating, for example, the effects of antiviral agents on prognosis in patients with Bell palsy.

The reasons for the disagreement over grade I outcomes between the two scales remains unclear. In theory, normal is nor-

mal no matter what scale is used. We assumed that the difference in grade I may reflect inter-examiner variation. Because every known scale has subjective characteristics, examiners are unable to fully assess facial function [4]. Although this possibility cannot be excluded, we believe that the principal reason for the difference between the scales is the ambiguity of HB grading in most cases, preventing accurate analysis without strict regional assessment [4].

When we compared the results of regional assessment with the final grade, it differed from results observed using existing "regional" HB grading [5]. In that study, the final grade match was highest for the eyes (61%), followed by the mid-face (40%) and mouth (32%). In comparison, we found that the highest HB grade match was for the eyes (67%), whereas the grade match on the FNGS 2.0 was higher for the mouth (72%) than for the eyes (63%).

Previously, evaluators focused primarily on whether the eye closed during different degrees of paralysis of each branch of the facial nerve [5]. Although not identical, as the evaluation criteria differ for the "regional" grading system and FNGS 2.0, the scope of interest has widened beyond the eyes. Our more systematic regional assessment confirmed that the FNGS 2.0 has the advantage of assessing the various paralysis states of each nerve, and of these states being reflected in the final grade.

This study had several limitations. First, all patients were evaluated on both systems by a single examiner. Had patients been evaluated by two or more examiners, we could have calculated interobserver differences, enhancing the value of our results. Second, the final follow-up was performed 3 months after treatment; however, this period is too short to make decisions on facial paralysis. We had hypothesized that patients with Bell's "paresis" would completely recover within 3 months [10], and we therefore regarded 3 months as the minimum period for determining patient prognosis. Since the main objective of this study was to assess the difference between the two facial nerve grading systems, the follow-up period itself was not a huge limitation. In other studies, many observers watch videos of patients [4,9,11,12]. In addition, a longer follow-up period increases the likelihood that some patients will drop out of the study, especially since patients may go elsewhere for a second opinion or seek alternative treatments. These problems may be overcome and long term results obtained by a multi-center study with larger numbers of patients.

Another limitation of this study was that we did not evaluate patients whose paralysis was due to other causes such as external injury or surgery. However, the purpose of this study was to compare the two grading systems in patients with unilateral Bell palsy. Future studies are needed to evaluate these grading systems in patients whose paralysis was due to other causes.

The future use of the FNGS 2.0 to grade patients with facial paralysis may result in a more objective evaluation of patient condition and recovery, both by regional assessment and by de-

termination of level of synkinesis. Regional assessments using the HB grading system may enhance communication between researchers [5].

Due to its simplicity and convenience, the HB grading system will be difficult to replace. However, the time required for FNGS 2.0 grading is not much longer. This system, which can overcome the shortcomings of the existing system may be a good tool for the more accurate evaluation of patients.

In conclusion, FNGS 2.0 showed moderate agreement with HB grading. The regional assessment capacity of the FNGS 2.0 yields stricter and more evaluations, enhancing the ability to assess patient grade and prognosis.

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENTS

This research was supported by the Kyung Hee University Research Fund in 2011 (KHU 2011-1098).

REFERENCES

- House JW, Brackmann DE. Facial nerve grading system. *Otolaryngol Head Neck Surg.* 1985 Apr;93(2):146-7.
- Ross BG, Fradet G, Nedzelski JM. Development of a sensitive clinical facial grading system. *Otolaryngol Head Neck Surg.* 1996 Mar;114(3):380-6.
- Yanagihara N. Grading of facial palsy. In: Fisch U, editor. *Facial nerve surgery.* Birmingham, AL: Aesculapius Publishing Co.; 1977. p. 533-5.
- 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 Apr;140(4):445-50.
- 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 Jan;24(1):118-22.
- Berg T, Jonsson L, Engström M. Agreement between the Sunnybrook, House-Brackmann, and Yanagihara facial nerve grading systems in Bell's palsy. *Otol Neurotol.* 2004 Nov;25(6):1020-6.
- Croxson G, May M, Mester SJ. Grading facial nerve function: House-Brackmann versus Burres-Fisch methods. *Am J Otol.* 1990 Jul;11(4):240-6.
- Satoh Y, Kanzaki J, Yoshihara S. A comparison and conversion table of 'the House-Brackmann facial nerve grading system' and 'the Yanagihara grading system'. *Auris Nasus Larynx.* 2000 Jul;27(3):207-12.
- Hu WL, Ross B, Nedzelski J. Reliability of the Sunnybrook Facial Grading System by novice users. *J Otolaryngol.* 2001 Aug;30(4):208-11.
- Linder TE, Abdelkafy W, Cavero-Vanek S. The management of peripheral facial nerve palsy: "paresis" versus "paralysis" and sources of ambiguity in study designs. *Otol Neurotol.* 2010 Feb;31(2):319-27.

11. 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 Apr;132(4):543-9.
12. Kanerva M, Poussa T, Pitkaranta A. Sunnybrook and House-Brackmann Facial Grading Systems: intrarater repeatability and interrater agreement. *Otolaryngol Head Neck Surg.* 2006 Dec;135(6):865-71.