Patient- versus physician-reported facial disability in vestibular schwannoma: an international cross-sectional study

Øystein Vesterli Tveiten, MD,1,3 Matthew L. Carlson, MD,5,4 Frederik Goplen, MD, PhD,2 Erling Myrseth, MD, PhD,1 Colin L. W. Driscoll, MD,4,5 Rupavathana Mahesparan, MD, PhD,1 Michael J. Link, MD,4,5 and Morten Lund-Johansen, MD, PhD1,3

Departments of 1Neurosurgery and 2Otorhinolaryngology, Haukeland University Hospital, Bergen; 3Department of Clinical Medicine, University of Bergen, Norway; and Departments of 4Neurologic Surgery and 5Otolaryngology–Head and Neck Surgery, Mayo Clinic School of Medicine, Rochester, Minnesota

OBJECTIVE Patient-reported outcomes are increasingly used in studies of vestibular schwannoma (VS); however, few studies have examined self-evaluated facial nerve function and its relation to physician-reported outcomes. The primary objective of this study was to compare patient self-evaluations of facial disability with physician-evaluated facial nerve status and with self-evaluations of a healthy control group. The second objective was to provide insight into the controversial subject of the optimal initial management of small- and medium-sized VSs; consequently, the authors compared patient-reported facial nerve disability following treatment via observation (OBS), Gamma Knife surgery (GKS), or microsurgery (MS). Lastly, the authors sought to identify risk factors for facial nerve dysfunction following treatment for small- and medium-sized VSs.

METHODS All patients with a VS 3 cm or smaller that was singly treated with OBS, GKS, or MS at either of 2 independent treatment centers between 1998 and 2008 were retrospectively identified. Longitudinal facial nerve measures and clinical data, including facial nerve evaluation according to the House-Brackmann (HB) grading system, were extracted from existing VS databases. Supplementing the objective data were Facial Disability Index (FDI) scores, which were obtained via survey of patients a mean of 7.7 years after initial treatment.

RESULTS The response rate among the 682 eligible patients was 79%; thus, data from a total of 539 patients were analyzed. One hundred forty-eight patients had been managed by OBS, 247 with GKS, and 144 with MS. Patients who underwent microsurgery had larger tumors and were younger than those who underwent OBS or GKS. Overall, facial nerve outcomes were satisfactory following treatment, with more than 90% of patients having HB Grade I function at the last clinical follow-up. Treatment was the major risk factor for facial nerve dysfunction. Almost one-fifth of the patients treated with MS had an objective decline in facial nerve function, whereas only 2% in the GKS group and 0% in the OBS cohort had a decline. The physical subscale of the FDI in the VS patients was highly associated with HB grade; however, the social/well-being subscale of the FDI was not. Thus, any social disability caused by facial palsy was not detectable by use of this questionnaire.

CONCLUSIONS The majority of patients with small- and medium-sized VSs attain excellent long-term facial nerve function and low facial nerve disability regardless of treatment modality. Tumor size and microsurgical treatment are risk factors for facial nerve dysfunction and self-reported disability. The FDI questionnaire is sensitive to the physical but not the social impairment associated with facial dysfunction.

https://thejns.org/doi/abs/10.3171/2016.8.JNS16707

KEY WORDS acoustic neuroma; activities of daily living; dual data collection; cerebellopontine angle tumor; facial paresis; vestibular schwannoma

ABBREVIATIONS AAO-HNS = American Academy of Otolaryngology–Head and Neck Surgery; FDI = Facial Disability Index; GKS = Gamma Knife surgery; HB = House-Brackmann; MS = microsurgery; OBS = observation; QOL = quality of life; VS = vestibular schwannoma.


INCLUDE WHEN CITING Published online December 2, 2016; DOI: 10.3171/2016.8.JNS16707.
Facial nerve paresis is the most visually obvious and disfiguring complication after treatment for vestibular schwannoma (VS), and from the surgeon’s point of view, the preservation of normal facial nerve function after treatment is a primary benchmark of a good outcome. Along with tumor control, facial nerve function is one of the most extensively studied treatment outcomes in the VS literature. However, with few exceptions, studies on facial nerve function after treatment have been completed with either objective physician-assessed parameters or nonvalidated questionnaires. Some of these studies have compared general quality of life (QOL) following treatment—but it is well known that general QOL inventories such as the 36-Item Short Form Health Survey (SF-36) are less sensitive to facial nerve weakness.

Microsurgery (MS) provides a high cure rate when a tumor is completely removed and remains, according to some authors, the preferred treatment modality for small- and medium-sized VSs, especially in younger patients. Given improvements in microsurgical techniques over the past several decades and the introduction of intraoperative electrophysiological monitoring of the facial nerve, the incidence of facial nerve injury associated with VS microsurgery has declined significantly. The majority of contemporary surgical series have reported preserved facial nerve function in over 80% of cases, with smaller tumors associated with the greatest chance of normal, long-term facial function. In the early days of Gamma Knife surgery (GKS), loss of facial nerve function was seen in up to 15% of patients, but modern-day tumor margin doses of 12–13 Gy yield tumor control rates above 90% and facial nerve preservation rates exceeding 95%. Conservative management via clinical and radiological observation (OBS) of small- to medium-sized VSs provides the lowest upfront risk to facial nerve function, but many observed tumors require intervention because of growth. Whether delayed treatment increases the risk of facial paralysis is not known. Most surgical series indicate poorer facial nerve outcomes in larger tumors. Despite comparable results in long-term tumor control among the active treatment options, the initial management strategy in small- and medium-sized VSs remains controversial.

One of the reasons for this disagreement is the poor scientific foundation for decision making; that is, studies of facial nerve outcomes following VS treatment are typically nonsystematic with most cases originating from single institutions, even single surgeons. The literature typically reports the outcomes following one treatment modality; with few exceptions, most studies are not comparative in design. The treating physician evaluates and reports facial nerve status according to the House-Brackmann (HB) facial nerve grading scale. The HB scale is prone to interobserver variability, and grading can be biased and incongruent with patient self-assessment. This is especially true for nervus intermedius dysfunction, as the HB classification does not cover lacrimal dysfunction. Few studies report objective and subjective facial nerve outcomes, and to our knowledge none originate from more than one center.

In the present study, we aimed to analyze and report both patient and physician assessments of long-term facial nerve function in a large multicenter cohort of patients with sporadic small- and medium-sized VSs treated with the contemporary therapeutic options: MS, GKS utilizing the Leksell Gamma Knife (Elekta), or OBS. Longitudinal facial nerve assessment and the results of a composite cross-sectional survey containing a validated facial nerve questionnaire were compared. The study was performed at 2 centers, 1 in the US and 1 in Europe.

Methods

Patients and Reference Population

Appropriate institutional review board approvals were obtained. All patients with small- and medium-sized VSs (≤3 cm in diameter in the cerebellopontine angle) referred to 1 of 2 independent tertiary care neurosurgical centers, 1 in the US and 1 in Norway, between 1998 and 2008 were eligible for this retrospective study. Only patients who had been conservatively managed or singly treated with MS or GKS were included. Patients with neurofibromatosis Type 2, those who had undergone MS or GKS treatment prior to the study period, those treated with more than 1 modality (for example, MS followed by GKS), and subjects younger than 18 years of age were excluded. Six hundred eighty-two eligible patients were identified. A reference population of 103 non-VS controls was recruited from a list obtained from a survey agency (Marketing Systems Group). The VS patients and the controls in the US were compensated with $30 US for study participation. Norwegian patients did not receive any remuneration. All participants provided informed consent. Nonresponders were contacted by telephone or by 1 reminder letter.

Treatment and Grouping

Patients were retrospectively grouped according to treatment: OBS, GKS, or MS. Patients initially managed conservatively and later treated with either MS or GKS were grouped according to intervention, and data recorded at the date of treatment were used as baseline. The date of the initial MRI diagnosis was considered the time of treatment for the OBS cohort. Treatment was individualized based on each patient’s age, health status, and preference as well as tumor characteristics. As a general rule, small- and medium-sized tumors were initially observed and later treated upon proven growth. Microsurgery was commonly avoided in older individuals and in patients with advanced comorbidities. Common reasons for observation were a small tumor size, minimal or no attributable symptoms, and an incidental diagnosis of VS. Patients treated with GKS typically received 12 Gy to the 50% isodose line with high (approximately 95%) tumor coverage and selectivity. Isocenters and dosages were deliberately planned so that the facial nerve (anterior tumor border) and cochlea were spared from higher doses. At both treatment centers, most patients (>90%) underwent MS via the retrosigmoid approach with the goal of complete tumor removal. Microsurgery was routinely performed with intraoperative facial nerve monitoring.

Data Collection and Assessment

Baseline and follow-up clinical, radiological, and de-
mographic data were collected from the respective study center’s VS databases. At baseline and the last clinical follow-up, facial nerve function was assessed according to the HB grading system. An HB grade of I indicates normal facial nerve function, and a maximum grade of VI indicates total facial nerve palsy. In line with previous studies, HB Grades I and II were defined as good outcomes following treatment. All VS patients and healthy controls completed the Facial Disability Index (FDI) questionnaire (see below) a mean of 7.7 years after treatment or diagnosis. The FDI is one of a few validated measures designed to evaluate self-reported disability following facial nerve paresis. In addition to the physical disability, a subjective psychosocial impact of facial nerve dysfunction is also measured by the FDI. The index is composed of 10 Likert-type questions, which generate 2 domain subscales: the physical and the social/well-being function. Items that measure social/well-being function were adapted from other questionnaires such as the Functional Status Questionnaire, the SF-36, and the Beck Depression Inventory. The FDI was validated in a sample of patients with a history of facial neuromotor system disorders and some residual facial neuromuscular function. The 2 scores are calculated on a scale from 0 (worst) to 100 (best) using formulas provided by the creators of the FDI.

**Statistical Analysis**

Continuous features were summarized with means and standard deviations or with means, medians, and interquartile ranges, depending on the data distribution. Comparisons among the 3 treatment groups were performed using the Kruskal-Wallis test. For analyses of means between 2 groups, the Student t-test was used if data were normally distributed (for example, age); for skewed data (for example, FDI), the Mann-Whitney U-test was used. For analyses of categorical values, the chi-square test was used. For statistical analyses and presentation, tumors were grouped according to size: Group 1, intracanalicular tumors; Group 2, 0–9 mm; Group 3, 10–19 mm; and Group 4, 20–30 mm. To assess the relationship between clinical features and facial nerve outcomes according to the HB classification, we dichotomized the final HB grade into either “favorable” (HB Grade I and II) or “unfavorable” (HB ≥ III) and performed a logistic regression analysis that was supplemented with a Spearman’s rank-order correlation. To assess the relationship between FDI and clinical data, linear regression analyses were used in addition to a Spearman’s rank-order correlation test. Two-sided p values less than 0.05 were considered statistically significant. Statistical analyses were performed using SPSS version 22 (SPSS Inc.).

**Results**

Five hundred thirty-nine patients returned a complete data set for analysis, yielding a 79% response rate. One hundred forty-eight patients (27.4%) were managed by OBS, 247 (45.8%) with GKS, and 144 (26.7%) with MS. The mean follow-up time was 7.7 years (Table 1). Nonresponders were slightly older than responders (p = 0.001). The MS-treated patients were significantly younger than OBS- and GKS-treated patients (p < 0.001). The OBS patients had significantly longer follow-ups than the GKS patients (p < 0.001). Tumor size ranging from largest to smallest: MS > GKS > OBS (p < 0.001). There were no differences between the occurrence of pretreatment facial nerve dysfunction across the treatment modalities (HB grade > I; p = 0.593).

<table>
<thead>
<tr>
<th>Variable</th>
<th>All</th>
<th>OBS</th>
<th>GKS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients</td>
<td>539</td>
<td>148</td>
<td>247</td>
<td>144</td>
</tr>
<tr>
<td>Mean age at diagnosis in yrs (SD)*</td>
<td>56.2 (12.3)</td>
<td>58.9 (10.9)</td>
<td>58.2 (12.3)</td>
<td>49.9 (11.4)</td>
</tr>
<tr>
<td>Mean age at survey in yrs (SD)*</td>
<td>63.9 (12.4)</td>
<td>67.2 (10.4)</td>
<td>65.5 (12.4)</td>
<td>57.6 (11.9)</td>
</tr>
<tr>
<td>Mean yrs from diagnosis to survey (SD)†</td>
<td>7.7 (2.4)</td>
<td>8.3 (2.2)</td>
<td>7.3 (2.4)</td>
<td>7.7 (2.5)</td>
</tr>
<tr>
<td>Sex, M/F (%)</td>
<td>304 (56)/235 (44)</td>
<td>88 (60)/60 (40)</td>
<td>119 (48)/128 (52)</td>
<td>88 (61)/56 (39)</td>
</tr>
<tr>
<td>Tumor size according to AAO-HNS group (%)‡</td>
<td>154 (28.6)</td>
<td>86 (58.1)</td>
<td>35 (14.2)</td>
<td>33 (22.9)</td>
</tr>
<tr>
<td>1: intracanalicular</td>
<td>112 (21.0)</td>
<td>30 (20.3)</td>
<td>65 (26.3)</td>
<td>17 (11.8)</td>
</tr>
<tr>
<td>2: 0–9 mm</td>
<td>211 (39.1)</td>
<td>26 (17.6)</td>
<td>128 (51.8)</td>
<td>57 (39.6)</td>
</tr>
<tr>
<td>3: 10–19 mm</td>
<td>57 (10.6)</td>
<td>1 (0.7)</td>
<td>19 (7.7)</td>
<td>37 (25.7)</td>
</tr>
<tr>
<td>4: 20–30 mm</td>
<td>5 (0.9)</td>
<td>5 (3.4)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HB grade (%)§</td>
<td>531 (98.5)</td>
<td>147 (99.3)</td>
<td>243 (98.4)</td>
<td>141 (97.9)</td>
</tr>
<tr>
<td>I</td>
<td>6 (1.1)</td>
<td>0</td>
<td>4 (1.6)</td>
<td>2 (1.4)</td>
</tr>
<tr>
<td>II</td>
<td>2 (0.4)</td>
<td>1 (0.7)</td>
<td>0</td>
<td>1 (0.7)</td>
</tr>
</tbody>
</table>

* The MS-treated patients were significantly younger than OBS- and GKS-treated patients (p < 0.001).
† The OBS patients had significantly longer follow-ups than the GKS patients (p < 0.001).
‡ Tumor size ranging from largest to smallest: MS > GKS > OBS (p < 0.001).
§ There were no differences between the occurrence of pretreatment facial nerve dysfunction across the treatment modalities (HB grade > I; p = 0.593).
0.006); however, there was no difference in the response rate among treatment modalities, and there were no other differences in baseline features between responders and nonresponders, including sex, facial nerve function (HB grade), hearing (AAO-HNS criteria), and tumor size prior to treatment (data not shown).

Baseline Population Data

Microsurgery-treated patients were younger than those who underwent GKS or OBS (p < 0.001), and tumor size was largest in the MS cohort and smallest in the OBS group (p < 0.001). Patients who underwent GKS had shorter follow-ups than those who were observed (7.3 vs 8.3 years, p < 0.001), and those who underwent MS and GKS had comparable follow-ups. The mean age of the 42 male and 61 female control subjects was 55.1 ± 10.7 years, which was significantly younger than the VS patients (p < 0.001). The sex distribution between patients and controls was similar (p = 0.664).

Before treatment, HB grading was normal (Grade I) in 98.5% of patients; 6 patients (1.1%) had partial paralysis (Grade II–V) and 2 (0.4%) had complete facial nerve paralysis (HB Grade VI). There was no difference in the occurrence of pretreatment facial nerve dysfunction among the 3 treatment modalities (p = 0.593), and there was no correlation between the first HB grade and tumor size (p = 0.802) or age at diagnosis (p = 0.453).

Follow-Up Data: Physician’s Assessment

Overall, 92.4% of patients had normal facial nerve function (HB Grade I) at the last clinical follow-up (Table 2); none of the conservatively managed patients experienced an objective decline in facial nerve function, whereas 5 (2%) GKS-treated patients and 28 (19.4%) MS-treated patients had worsened facial nerve function. Impairment was more likely in surgically treated patients (p < 0.001). The difference between OBS and GKS patients did not reach statistical significance (p = 0.07).

A multivariate logistic regression analysis with 1) age at follow-up, 2) sex, 3) tumor size, and 4) treatment modality as predictor variables for an unfavorable facial nerve outcome (HB grade ≥ III) proved significant (p < 0.001). Of the 4 predictor variables, treatment modality and tumor size were statistically significant. Increasing tumor size was associated with an increased likelihood of an unfavorable outcome as only the largest tumor group emerged as a significant predictor (Group 4, p = 0.043); the smaller tumor sizes were not significant predictors. Treatment with GKS (p = 0.046) and MS (p < 0.001) were also significant predictors of an unfavorable outcome, whereas observation was not (p = 0.886). Although treatment was a risk factor for facial nerve dysfunction, there was only a weak, albeit statistically significant, correlation (r = 0.228) between treatment modality and the latest HB grade (p < 0.001).

Follow-Up Data: Patient Self-Assessment

The distribution of both FDI physical and social/well-being scores in 539 patients was skewed toward a favorable outcome (Fig. 1); that is, most patients did not report a profound physical or social/well-being facial disability. This was especially true for the FDI physical subscale for which the mean and median scores were 94.2 and 100, respectively. The mean FDI physical score in VS patients was poorer than that in nontumor controls (Table 3); however, the mean FDI social/well-being score in patients was significantly better than that in the healthy controls (p = 0.017). Similar findings were also found in patients with an HB grade > II. Specifically, the 19 patients with facial nerve dysfunction as defined by an HB grade > II had a poorer FDI physical score than the 103 healthy controls (p < 0.001), but the social/well-being score was similar (p = 0.594).

When comparing FDI physical and social/well-being scores between treatment groups and the control group (Table 3 and Fig. 2), several statistically significant differences were found: MS-treated patients had poorer FDI physical scores than the other 2 treatment groups and the healthy controls (p < 0.001). Patients treated with GKS also had poorer FDI physical scores than the healthy controls (p = 0.017), whereas the OBS group had scores comparable to those in both the GKS-treated patients (p = 1.000) and the control group (p = 0.199). Regarding the social/well-being scores, we found significant differences on a global comparison between treatment groups (p = 0.032) due to significantly better scores in GKS patients compared with the MS group (p = 0.036). There were no differences between OBS patients and the MS group (p = 0.124) or between the OBS and GKS groups (p = 1.000).

Without a previously established cutoff point for the FDI, we defined a score < 90 as “unfavorable” and ≥ 90 as “favorable” and performed a binomial regression analysis to identify possible predictors (age, sex, tumor size, and treatment modality) of an unfavorable outcome. For both FDI physical and social/well-being scores, the model proved significant (p < 0.001). Age and sex did not predict the FDI physical or social/well-being score, but tumor size and treatment modality did (p < 0.001 and p = 0.048, respectively).

Women scored somewhat lower than men on the FDI physical subscale (93.7 vs 94.9, p = 0.02) and social/well-being subscale (77.9 vs 80.0, p = 0.008), even though there were not more women with an unfavorable facial nerve outcome according to HB grade (3.3 vs 3.8%, p = 0.821).

### TABLE 2. Facial nerve function at the last clinical follow-up in 539 patients with small- and medium-sized VSs

<table>
<thead>
<tr>
<th>Variable</th>
<th>All</th>
<th>OBS</th>
<th>GKS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients</td>
<td>539</td>
<td>148</td>
<td>247</td>
<td>144</td>
</tr>
<tr>
<td>HB grade (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>498 (92.4)</td>
<td>147 (99.3)</td>
<td>238 (96.4)</td>
<td>113 (78.5)</td>
</tr>
<tr>
<td>II</td>
<td>22 (4.1)</td>
<td>0</td>
<td>6 (2.4)</td>
<td>16 (11.1)</td>
</tr>
<tr>
<td>III</td>
<td>15 (2.8)</td>
<td>0</td>
<td>3 (1.2)</td>
<td>12 (8.3)</td>
</tr>
<tr>
<td>IV</td>
<td>1 (0.2)</td>
<td>0</td>
<td>0</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td>V</td>
<td>1 (0.2)</td>
<td>0</td>
<td>0</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td>VI</td>
<td>2 (0.4)</td>
<td>1 (0.7)</td>
<td>0</td>
<td>1 (0.7)</td>
</tr>
</tbody>
</table>

* Differences were statistically significant between the OBS patients and the GKS and MS patients (p < 0.001), but the difference between OBS and GKS patients failed to reach statistical significance (p = 0.067).
Age did not correlate with the FDI physical (r = 0, p = 0.993) or social/well-being scores (r = 0.059, p = 0.172). The FDI physical score had a weak yet significant correlation with its social/well-being counterpart (r = 0.363, p < 0.001).

Comparison of Physician and Patient Assessments

Patients reported more physical facial disability with increasing HB grade (Table 4 and Fig. 3 left), and there was a moderate correlation between the 2 measurement scales (r = 0.468, p < 0.001). Patients with HB Grade I reported significantly less facial disability than those with HB Grades II and III, as reflected by the FDI physical score (p < 0.001). Differences between the other groups failed to reach significant levels; for example, we found no difference in physical disability between patients with HB Grade IV and those with Grade VI. It is noteworthy that for HB Grades II and III, the FDI scores were scattered; that is, patients with mild to moderate facial nerve dysfunction evaluated their physical facial disability non-uniformly. The same pattern was not found for the social/well-being subscale score as there were no differences between patients in the different HB groups (p = 0.208), and there was no correlation between HB grade and the FDI social/well-being score (r = -0.039, p = 0.367; Fig. 3 right).

In a univariate linear regression analysis, the final HB grade was a significant predictor of the FDI physical score
Treatment and tumor size were also predictors of a poorer FDI physical score ($p < 0.001$), while age and sex were not ($p = 0.993$ and 0.132, respectively). Neither of the variables above predicted any variance in the FDI social/well-being score.

Patients With HB Grade I Function

The 498 patients with normal facial nerve function (HB Grade I) at the last follow-up reported more physical but less social disability on the FDI than did the healthy controls (95.5 vs 98.2, $p < 0.001$; and 79.11 vs 75.5, $p = 0.006$, respectively).

There were no treatment-related differences in the social/well-being score among the patients with HB Grade I (global comparison, $p = 0.087$), but patients treated with MS reported worse physical disability than the 2 other treatment groups ($p < 0.001$). By analyzing each of the 10 questions on the FDI, we found that the question regarding “excessive tearing and drying of the eye” explained most of the lower scores seen in the MS group (Fig. 4). Specifically, MS-treated patients reported more difficulties than both the OBS- and GKS-managed patients (mean FDI scores for question 4 in OBS, GKS, and MS groups were 4.43, 4.45, and 4.12, respectively, $p < 0.003$; Table 5). One patient did not answer question 4 and was therefore omitted from this comparison. Significant differences on a global comparison were also found for questions 1 and 2 (“difficulties eating” and “difficulties drinking,” $p < 0.030$ and $< 0.033$, respectively)—but the differences among the 3 treatment groups were less than those for excessive tearing and drying of the eye (data not shown).

Discussion

Preservation of facial nerve function remains one of the most heavily studied outcomes of VS treatment and, beyond tumor control, is one of the most important outcomes when treatment options are compared. However, regarding facial nerve outcomes, there is no tradition of using patient-reported outcome measures in the setting of VS, although there are a few exceptions. Parallelizing the results of the current study, a few studies comparing GKS and MS have demonstrated less facial nerve dysfunction with GKS. To our knowledge, our study is

<table>
<thead>
<tr>
<th>TABLE 4. Overview of FDI scores in 539 patients with sporadic VS, grouped by the latest HB grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subscale</td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Physical function</td>
</tr>
<tr>
<td>Social/well-being function</td>
</tr>
</tbody>
</table>
the first to demonstrate less facial nerve dysfunction after both GKS and OBS in a comparison of the 3 prevalent treatment options for small- and medium-sized VSs. When comparing patient- and physician-assessed outcomes, the HB grading system discriminated group differences better than the FDI responses. This finding could indicate that facial nerve dysfunction affects patients to a lesser degree than the treating physician may believe, or that other factors beyond facial nerve capacity influence perceived outcomes of the FDI questionnaire. For example, women reported a higher facial disability score than men, although the HB grade was similar in both groups.

In particular, it was surprising that the social/well-being subscore of the FDI questionnaire was insensitive to facial dysfunction; VS patients as a group had better social/well-being scores than the 103 healthy controls, indicating that the 5 social functioning and well-being questions on the FDI do not relate to objective facial disability. Another explanation is that in the long period (mean 7.7 years) since they acquired facial nerve palsy, patients may have become accustomed to the dysfunction and no longer felt any particular social discomfort or insecurity. It is important to note that all control subjects were US citizens, whereas the study population was international. This potential bias may in theory explain the difference in scores; however, FDI differences were not found between the US and the Norwegian patients (data not shown). Alternatively, the FDI questionnaire in itself may not be sensitive enough to identify differences when groups are arranged by facial function. Notably, the FDI is partly designed from other generic QOL questionnaires, including the SF-36. Only 1 of the social/well-being questions on the FDI specifically inquires about facial function, whereas the remaining 4 questions are general: “feeling calm and peaceful,” “isolation from others,” “irritable towards others,” and “early wake-up/poor nighttime sleep.” It is well documented that general questionnaires such as those on the SF-36 have low sensitivity to facial palsy.44,51 We believe that we have replicated such findings through our use of the FDI; facial nerve function does not seem to impact
general questions on QOL assessments. Interestingly, however, the social/well-being evaluation of the symptom-specific FDI questionnaire demonstrated that the MS cohort had results similar to those of the healthy controls despite the reduced HB grade of the former group. Overall, this finding could indicate that facial weakness alone is not a strong predictor of, for example, posttreatment depression. On the contrary, there was a correlation between HB grade and patient-reported facial physical disability. Given these findings, it appears that the social/well-being scale of the FDI is less valuable than the physical scale when assessing VS outcomes. It is also worth noting that the FDI social/well-being evaluation of the symptom-specific tool demonstrates a moderate, yet significant, correlation with its physical counterpart (data not shown).

Our findings also clearly show that the HB scale is not sensitive enough to distinguish every aspect of facial nerve dysfunction. Facial nerve synkinesis is commonly incorporated into facial nerve assessments—for example, the Sunnybrook Facial Grading System—but other secondary defects (such as lacrimation and gustatory function) are not. While symptoms of nervus intermedius dysfunction are often overlooked, our data demonstrate that excessive tearing and drying of the eye explained most of the increased disability seen in the MS group. Furthermore, because the Sunnybrook system and similar assessment tools remain subjective, they are also prone to the same interobserver variability as the HB classification. There are specific tools for a more in-depth assessment of the facial nerve (for example, the Nottingham grading system); however, none are as universally accepted as the HB grading system. Our findings are important for patient examination and counseling; the treating physician should have secondary facial nerve dysfunction in mind and should ensure that patients are informed about potential long-term facial nerve dysfunction beyond primary motor function. It is notable that the FDI does not assess gustatory function; therefore, it is likely that the FDI is not sensitive to taste impairment and that patients with subtle facial nerve dysfunction also may have such problems—in the same manner that we demonstrate difficulties with tearing and dry eye.

Most of the 539 VS patients had preserved facial nerve function according to both the patient’s and the physician’s evaluations, and our results are within the range of what can be expected following treatment. There were more patients with facial nerve dysfunction in the MS group, and microsurgical treatment as an independent variable was a strong predictor of a poorer outcome. Tumor size is a known predictor of postoperative facial nerve dysfunction, and indeed our surgically managed patients had larger tumors than the 2 other treatment arms. It is also worth noting that we included only singly treated patients in our study, indicating that the degree of tumor resection was extensive and total—with a higher risk of facial palsy. However, our analysis indicates that we cannot attribute the poorer results in the MS group to tumor size alone. Rather, we conclude that the microsurgical treatment of even small- and medium-sized VSs entails an increased risk of facial nerve complications as compared with the other 2 treatment options. Furthermore, we found no difference in facial outcome between the GKS and OBS groups, confirming the general clinical impression that GKS does not cause any long-term clinical change that the patient will recognize.

In closing, we review the primary strengths and limitations of this study. Its main strength is its sample size and high response rate, which ensured that the sample represents the general VS population. To our knowledge, it is the first study comparing patient- and physician-reported facial outcomes of 3 management modalities, as well as outcomes for a control group. The study’s main limitation is the nonrandom treatment allocation and the lack of a true intention-to-treat concept. The choice of treatment is often influenced by factors such as patient age and tumor size and growth, which can produce selection bias and influence outcomes. Moreover, it is worth noting that we excluded patients with recurrent tumors that required additional treatment. Failure rates of VS treatment are low, however, and it is unlikely that the overall conclusions would have significantly changed if patients receiving more than one treatment had been included. Finally, it is possible that the FDI outcome data from nonresponders may have differed from that from responders—that is, nonresponders could be the group with greater social/well-being difficulties. However, given the overall similarities between responders and nonresponders for other features, we find any difference improbable.

### TABLE 5. Differences between treatment groups on question 4 on the FDI: “How much difficulty did you have with your eye tearing excessively or becoming dry?”

<table>
<thead>
<tr>
<th>Score</th>
<th>Response</th>
<th>OBS</th>
<th>%</th>
<th>GKS</th>
<th>%</th>
<th>MS</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Usually did not have tearing because of other reasons</td>
<td>2</td>
<td>1.4</td>
<td>1</td>
<td>0.4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>Usually did not have tearing because of health</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Usually had much difficulty</td>
<td>7</td>
<td>4.8</td>
<td>6</td>
<td>2.5</td>
<td>7</td>
<td>6.2</td>
</tr>
<tr>
<td>3</td>
<td>Usually had some difficulty</td>
<td>12</td>
<td>8.2</td>
<td>35</td>
<td>14.7</td>
<td>25</td>
<td>22.1</td>
</tr>
<tr>
<td>4</td>
<td>Usually had a little difficulty</td>
<td>28</td>
<td>19.2</td>
<td>39</td>
<td>16.5</td>
<td>28</td>
<td>24.8</td>
</tr>
<tr>
<td>5</td>
<td>Usually had no difficulty</td>
<td>97</td>
<td>62.2</td>
<td>157</td>
<td>66.0</td>
<td>53</td>
<td>46.5</td>
</tr>
<tr>
<td>Total no. (overall no. 497)</td>
<td>146</td>
<td>238</td>
<td>113</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All patients had normal facial nerve function (HB Grade I). One patient did not answer this question and was therefore omitted from this comparison.
Conclusions
When managed by an experienced multidisciplinary team, most patients with small- and medium-sized VSs obtain excellent long-term facial nerve function and low facial nerve disability regardless of treatment modality. Tumor size and microsurgical treatment are risk factors for facial nerve dysfunction and self-reported disability. The FDI questionnaire is sensitive to the physical but not the social impairment associated with facial dysfunction. The influence of secondary effects such as lacrimal dysfunction may be underestimated, and future instruments should consider incorporating such assessments.

Acknowledgments
We thank the 2 study nurses Monica K. Finnkirk, RN, and Nicole M. Tombers, RN, for providing clinical care for the patients and for contributing to data collection.

References

Disclosures
Internal departmental funding was used without commercial sponsorship or support.

Author Contributions

Supplemental Information
Previous Presentations
Portions of this manuscript were presented at the 25th Annual Meeting of the North American Skull Base Society, Tampa, Florida, USA (20–25 February 2015), and at the 11th European Skull Base Society Congress (ESBS 2014), Paris, France (26–28 June 2014).

Correspondence
Øystein Vesterli Tveiten, Department of Neurosurgery, Haukeland University Hospital, Jonas Lies vei 65, Bergen N-5021, Norway. email: oystein.tveiten@helse-bergen.no.