Long-term quality of life in patients with vestibular schwannoma: an international multicenter cross-sectional study comparing microsurgery, stereotactic radiosurgery, observation, and nontumor controls

Matthew L. Carlson, MD,1 Oystein Vesterli Tveiten, MD,3 Colin L. Driscoll, MD,1,2 Frederik K. Goplen, MD, PhD,4 Brian A. Neff, MD,1 Bruce E. Pollock, MD,2 Nicole M. Tombers, RN,1 Marina L. Castner, RN,2 Monica K. Finnkirk, RN,3 Erling Myrseth, MD, PhD,3 Paal-Henning Pedersen, MD, PhD,3,5 Morten Lund-Johansen, MD, PhD,3,5 and Michael J. Link, MD,1,2

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

OBJECT The optimal treatment for sporadic vestibular schwannoma (VS) is highly controversial. To date, the majority of studies comparing treatment modalities have focused on a narrow scope of technical outcomes including facial function, hearing status, and tumor control. Very few publications have investigated health-related quality of life (HRQOL) differences between individual treatment groups, and none have used a disease-specific HRQOL instrument.

METHODS All patients with sporadic small- to medium-sized VSs who underwent primary microsurgery, stereotactic radiosurgery (SRS), or observation between 1998 and 2008 were identified. Subjects were surveyed via postal questionnaire using the 36-Item Short Form Health Survey (SF-36), the 10-item Patient-Reported Outcomes Measurement Information System short form (PROMIS-10), the Glasgow Benefit Inventory (GBI), and the Penn Acoustic Neuroma Quality-of-Life (PANQOL) scale. Additionally, a pool of general population adults was surveyed, providing a nontumor control group for comparison.

RESULTS A total of 642 respondents were analyzed. The overall response rate for patients with VS was 79%, and the mean time interval between treatment and survey was 7.7 years. Using multivariate regression, there were no statistically significant differences between management groups with respect to the PROMIS-10 physical or mental health dimensions, the SF-36 Physical or Mental Component Summary scores, or the PANQOL general, anxiety, hearing, or energy subdomains. Patients who underwent SRS or observation reported a better total PANQOL score and higher PANQOL facial, balance, and pain subdomain scores than the microsurgical cohort (p < 0.02). The differences in scores between the nontumor control group and patients with VS were greater than differences observed between individual treatment groups for the majority of measures.

CONCLUSIONS The differences in HRQOL outcomes following SRS, observation, and microsurgery for VS are small. Notably, the diagnosis of VS rather than treatment strategy most significantly impacts quality of life. Understanding that a large number of VSs do not grow following discovery, and that intervention does not confer a long-term HRQOL advantage, small- and medium-sized VS should be initially observed, while intervention should be reserved for patients with unequivocal tumor growth or intractable symptoms that are amenable to treatment. Future studies assessing HRQOL in VS patients should prioritize use of validated disease-specific measures, such as the PANQOL, given the significant limitations of generic instruments in distinguishing between treatment groups and tumor versus nontumor subjects.

http://thejns.org/doi/abs/10.3171/2014.11.JNS14594

KEY WORDS quality of life; microsurgery; stereotactic radiosurgery; Gamma Knife radiosurgery; acoustic neuraoma; vestibular schwannoma; cerebellopontine angle

ABBREVIATIONS AAO-HNS = American Academy of Otolaryngology-Head and Neck Surgery; CPA = cerebellopontine angle; GBI = Glasgow Benefit Inventory; HRQOL = health-related quality of life; IQR = interquartile range; MCID = minimum clinically important difference; PANQOL = Penn Acoustic Neuroma Quality-of-Life; PROMIS-10 = 10-item Patient-Reported Outcomes Measurement Information System; SF-36 = 36-item Short Form Health Survey; SRS = stereotactic radiosurgery; VS = vestibular schwannoma.


INCLUDE WHEN CITING Published online January 2, 2015; DOI: 10.3171/2014.11.JNS14594.

DISCLOSURE The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper. Internal departmental funding was used without commercial sponsorship or support.
VESTIBULAR schwannomas (VSs) are benign, slow-growing tumors arising from the eighth cranial nerve and constitute 8% of all intracranial neoplasms and 90% of cerebellopontine angle (CPA) lesions. Increasing patient access to MRI and the implementation of screening protocols for patients with asymmetrical hearing loss have led to a significant increase in the number of VSs diagnosed each year, with a higher proportion of patients presenting with small tumors and minimal symptoms. Contemporary management strategies include microsurgical resection, stereotactic radiosurgery (SRS) or stereotactic radiotherapy, and observation with serial imaging. Owing to concurrent advancements in surgical technique and radiation delivery strategies, disease-specific mortality now approaches zero and the majority of patients benefit from durable tumor control whether managed using SRS or microsurgery. In addition, recent detailed natural history studies have revealed that the majority of small VSs do not grow over many years of follow-up, leading to an ever-increasing number of patients being managed conservatively, without intervention.

To date, the great majority of published studies comparing VS management strategies have concentrated only on physician-reported objective outcome measures, including facial nerve function, hearing status, tumor control, and other cranial nerve–related morbidity and/or mortality. Within the last 20 years there has been growing interest within the medical community to develop methods for quantifying patient-perceived outcomes, understanding there is often disparity in what we as health care providers prioritize and what patients value. Previous attempts to ascertain which treatment strategy ultimately provides the best quality of life for patients with small- to medium-sized VSs have been severely hampered by small numbers, short follow-up, potentially important selection biases, lack of a disease-specific quality of life analysis, and no nontumor control group for comparison. After an extensive review of the recent literature, Gauden and colleagues concluded that since most studies have relied on the 36-Item Short Form Health Survey (SF-36) and have reported outcomes in disparate ways, no strong conclusion can be drawn regarding the merits of one treatment over another, and the existing literature suffers significantly from monomethod biases. The primary objective of the current study is to assess long-term health-related quality of life (HRQOL) in a large cohort of patients following observation, microsurgery, and SRS for small- to medium-sized VS using 3 multipurpose HRQOL questionnaires and 1 recently validated disease-specific instrument.

Methods

Study Cohort

After obtaining institutional review board approval, we performed a cross-sectional study via postal questionnaire at 2 tertiary academic referral centers, one located in the United States and one in Western Europe. All patients with smaller than 3-cm sporadic VSs who underwent microsurgery, SRS, or observation between 1998 and 2008 were identified. Subjects with neurofibromatosis Type 2, those who received prior SRS or microsurgery, patients with tumors larger than 3 cm, and subjects younger than 18 years were excluded from study. Baseline clinical data were gathered for all patients and included sex, age at treatment, pretreatment hearing capacity, baseline facial nerve function, and tumor size. For the purpose of analysis, time of treatment for the observation cohort was considered the date of initial diagnosis. Tumor size and audiometric data were classified according to the 1995 American Academy of Otolaryngology-Head and Neck Surgery (AAO-HNS) Committee reporting guidelines. Facial nerve function was recorded using the House-Brackmann grading scale.

In our practice, most VSs smaller than 2 cm in maximum posterior fossa diameter are initially observed until serial MRI reveals evidence of tumor growth, understanding that a significant number of VSs remain dormant for extended periods of time. Because currently there is no compelling evidence supporting one therapy over another, we do not follow a rigid management algorithm, and treatment is highly individualized. Following definite radiological progression, in the majority of cases the treatment decision is based on tumor characteristics, symptoms, age, health, and patient preference. Patients with minimal symptoms and progressive growth are more commonly treated with SRS, while those with intractable symptoms that are related to mass effect are more likely to undergo microsurgery. Following this paradigm, subjects who are older, those with advanced medical comorbidities, and/or those with minimal symptoms are more likely to receive SRS, while younger fit patients and those with severe symptoms are more likely to undergo microsurgery. Over the past 15 years there has been a near-even distribution of management between microsurgery and SRS, with the large majority of patients being evaluated by a neurotologist-neurosurgeon team.

HRQOL Assessment

Those meeting initial inclusion criteria were surveyed using a panel of standardized HRQOL questionnaires as outlined below. Global HRQOL status was evaluated using the SF-36 and the 10-item Patient-Reported Outcomes Measurement Information System (PROMIS-10) questionnaires. The SF-36 is a generic, 36-item instrument that assesses HRQOL status across 8 scales including physical functioning, role limitations secondary to physical problems, bodily pain, general health, vitality, social functioning, role limitations secondary to emotional problems, and mental health. Scores for each domain are calculated based on the weighted sum of corresponding responses and are adapted to a 0- to 100-point system, with higher scores indicating more favorable HRQOL. Finally, individual scales are combined to form physical and mental health component summary scores. Similar to the SF-36, the PROMIS-10 provides mental and physical health summary scores. Each item contains 5 responses with the exception of an 11-point pain scale; a higher score once again indicates better HRQOL.

The Glasgow Benefit Inventory (GBI) is an 18-point postintervention questionnaire used to assess change in health status as a result of treatment, developed especially for otolaryngological interventions. Each item contains 5 responses ranging from a large deterioration in HRQOL to
a large improvement. A total score is generated from general, social support, and physical health subscale scores. The GBI total and subdomain scores range from −100 to +100 with a score of 0 indicating no change in health status as a result of treatment, and a score of +100 denoting a maximum improvement.

Disease-specific HRQOL was assessed using the Penn Acoustic Neuroma Quality-of-Life (PANQOL) scale. This questionnaire has been validated previously and was shown to discriminate VS patients from nontumor controls better than the SF-36. The PANQOL consists of 26 questions with responses ranging from 1 to 5, with 1 denoting strong disagreement through 5 indicating strong agreement. Domain scores for facial function, balance, hearing, pain, anxiety, energy, and general health are obtained by averaging the responses of items assigned to the domain. A total score is calculated as the equal average of the 7 domain scores. As such, the domain scores and the total score could range from 0 to 100, with higher scores indicating better HRQOL.

A pool of general population adults was also surveyed using the SF-36, PROMIS-10, and PANQOL, providing a group of nontumor controls for comparison. The GBI was not completed by control patients since this questionnaire requires a comparison of before and after an event or treatment, and therefore cannot be applied to this population. A list of names and contact information for non-VS controls was purchased from a survey agency, Marketing Systems Group. All US patients with VSs and control subjects were given $30 (US dollars) remuneration upon survey completion. One follow-up reminder was sent to all initial nonresponders.

Statistical Analysis
Continuous features were summarized with means, medians, and interquartile ranges (IQRs); categorical features were summarized with frequency counts and percentages. Comparisons of baseline demographics, clinical outcomes, and HRQOL outcomes among patients in the 3 management groups were evaluated using the Kruskal-Wallis tests for continuous and ordinal features, while chi-square and Fisher exact tests were used to analyze categorical variables. If the p value from a global test for differences among the SRS, observation, and microsurgery groups was < 0.10, pairwise comparisons between SRS and observation, between SRS and microsurgery, and between observation and microsurgery were evaluated using Wilcoxon rank-sum, chi-square, and Fisher exact tests. Comparisons of HRQOL outcomes among groups of interest were also evaluated in a multivariable setting after adjusting for differences in baseline patient features using linear and logistic regression models. Statistical analyses were performed using the SAS software package (SAS Institute). All tests were 2-sided, and p values < 0.05 were considered statistically significant.

Results
Baseline Population Data
A total of 642 consecutive respondents were analyzed, including 144 microsurgery, 247 SRS, and 148 observation patients in addition to 103 nontumor controls. Overall, 79% of VS patients who were successfully contacted agreed to participate and returned a completed questionnaire set. The response rate was not statistically significantly different between treatment groups. Furthermore, aside from older age in the nonresponder population (p = 0.006), differences in baseline features including sex, tumor size, pre- and posttreatment hearing, and pre- and posttreatment facial nerve function were not different between responders and nonresponders (p > 0.05). A comparison of baseline demographics and clinical outcomes among patients who underwent SRS, microsurgery, and observation is summarized in Table 1. Patients who were treated with microsurgery were significantly younger than those who received SRS or observation (p < 0.001). Patients who underwent microsurgery had larger tumors compared to the SRS (p = 0.040) and observation (p < 0.001) cohorts. Tumors in the radiosurgery group were likewise significantly larger compared to those in the observation arm (p < 0.001). Overall, the mean time interval between treatment and survey was 7.7 years; patients receiving SRS had the shortest duration of follow-up compared with patients who were observed (7.3 vs 8.3 years; p < 0.001).

Audiometric and Facial Nerve Outcomes
Prior to treatment, 48% of patients who underwent SRS, 60% of patients who underwent microsurgery, and 63% of patients who were observed had ipsilateral serviceable hearing (AAO-HNS Class A or B). Those who underwent observation and microsurgery were more likely to have serviceable hearing at the time of diagnosis than subjects treated with SRS (p = 0.025). At last follow-up, 68% of observed patients who had serviceable hearing at presentation, 40% of those who underwent SRS, and 14% of subjects who were treated with microsurgery maintained serviceable hearing following management (p < 0.001). There were no differences between groups with respect to pretreatment facial nerve function. At last follow-up, 99% of observed patients, 98% of those who underwent SRS, and 90% of patients who received microsurgery had good facial nerve function (House-Brackmann Grades I and II) (p < 0.001).

HRQOL Outcomes
A comparison of HRQOL outcomes among patients who underwent SRS, observation, and microsurgery is summarized in Table 2 and Fig. 1. After adjusting for baseline differences of age, tumor size, initial hearing status, and time from initial evaluation, there was no statistically significant difference found between treatment groups with the PROMIS physical, PROMIS mental, PANQOL general, PANQOL hearing, PANQOL energy, or SF-36 Physical or Mental Component Summary scores using multivariable regression models. After multivariable adjustment, pairwise comparisons of the PANQOL facial, PANQOL balance, PANQOL pain, PANQOL total, and GBI physical scores showed no statistically significant difference between the SRS and observation groups; however, both had statistically significantly better scores than the microsurgery arm. When analyzing the GBI total and
general outcome scores, patients receiving SRS had the highest score, followed by observation and microsurgery. GBI physical scores showed no statistical difference between the SRS and observation groups, and both had better scores than the microsurgery arm. When evaluating the GBI social assessment, those who received microsurgery had the highest average score followed by SRS and observation.

Multivariate comparisons of HRQOL outcomes of non-tumor controls and patients who underwent SRS,观察或显微手术。

<table>
<thead>
<tr>
<th>Feature</th>
<th>SRS (n = 247)</th>
<th>Observation (n = 148)</th>
<th>Microsurgery (n = 144)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at diagnosis in yrs</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mean</td>
<td>58</td>
<td>59</td>
<td>50</td>
<td>0.51</td>
</tr>
<tr>
<td>Median (IQR)</td>
<td>59 (49–67)</td>
<td>59 (52–67)</td>
<td>51 (42–59)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age at survey in yrs</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mean</td>
<td>66</td>
<td>67</td>
<td>58</td>
<td>0.13</td>
</tr>
<tr>
<td>Median (IQR)</td>
<td>65 (56–74)</td>
<td>67 (61–74)</td>
<td>59 (50–67)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Diagnosis to survey in yrs</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mean</td>
<td>7.3</td>
<td>8.3</td>
<td>7.7</td>
<td>0.14</td>
</tr>
<tr>
<td>Median (IQR)</td>
<td>6.8 (5.1–9.0)</td>
<td>8.2 (7.0–9.5)</td>
<td>7.7 (5.7–9.6)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>128 (52)</td>
<td>88 (59)</td>
<td>88 (61)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>119 (48)</td>
<td>60 (41)</td>
<td>56 (29)</td>
<td></td>
</tr>
<tr>
<td>Pretreatment hearing (n = 521)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>61 (25)</td>
<td>45 (32)</td>
<td>52 (37)</td>
<td>0.015</td>
</tr>
<tr>
<td>B</td>
<td>53 (22)</td>
<td>45 (32)</td>
<td>31 (22)</td>
<td>0.012</td>
</tr>
<tr>
<td>C</td>
<td>28 (12)</td>
<td>11 (8)</td>
<td>9 (6)</td>
<td>0.021</td>
</tr>
<tr>
<td>D</td>
<td>98 (41)</td>
<td>41 (29)</td>
<td>47 (34)</td>
<td>0.96</td>
</tr>
<tr>
<td>Most recent hearing (n = 470)</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>A</td>
<td>14 (7)</td>
<td>28 (22)</td>
<td>8 (6)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>B</td>
<td>25 (12)</td>
<td>30 (23)</td>
<td>3 (2)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>C</td>
<td>22 (11)</td>
<td>14 (11)</td>
<td>0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>D</td>
<td>143 (70)</td>
<td>57 (44)</td>
<td>126 (92)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Pretreatment facial function (n = 539)</td>
<td></td>
<td></td>
<td></td>
<td>0.60</td>
</tr>
<tr>
<td>I</td>
<td>243 (98)</td>
<td>147 (99)</td>
<td>141 (98)</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>4 (2)</td>
<td>0</td>
<td>2 (1)</td>
<td></td>
</tr>
<tr>
<td>VI</td>
<td>0</td>
<td>1 (1)</td>
<td>1 (1)</td>
<td></td>
</tr>
<tr>
<td>Most recent facial function (n = 536)</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>I</td>
<td>236 (96)</td>
<td>147 (99)</td>
<td>112 (78)</td>
<td>0.070</td>
</tr>
<tr>
<td>II</td>
<td>6 (2)</td>
<td>0</td>
<td>16 (11)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>III</td>
<td>3 (1)</td>
<td>0</td>
<td>12 (8)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>IV</td>
<td>0</td>
<td>0</td>
<td>1 (1)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>V</td>
<td>0</td>
<td>0</td>
<td>1 (1)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>VI</td>
<td>0</td>
<td>1 (1)</td>
<td>1 (1)</td>
<td></td>
</tr>
<tr>
<td>Pretreatment tumor size (n = 534)</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Intracanalicular tumor</td>
<td>35 (14)</td>
<td>86 (60)</td>
<td>33 (23)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>0–9 mm, CPA component</td>
<td>65 (26)</td>
<td>30 (21)</td>
<td>17 (12)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>10–19 mm, CPA component</td>
<td>128 (52)</td>
<td>26 (18)</td>
<td>57 (40)</td>
<td>0.040</td>
</tr>
<tr>
<td>20–30 mm, CPA component</td>
<td>19 (8)</td>
<td>1 (1)</td>
<td>37 (26)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

MS = microsurgery; Ob = observation.
* Values are presented as the number of patients (%) unless stated otherwise; p values in boldface are statistically significant.
† AAO-HNS hearing class.
‡ House-Brackmann facial nerve grading scale.
tion, and microsurgery are shown in Table 3. Additionally, a comparison between nontumor controls and the collective population of patients with VS is provided. Of note, after adjusting for age, the PROMIS physical, and the SF-36 Physical and Mental Component Summary scores did not show any statistically significant differences between nontumor control subjects and any of the individual management arms. Compared with the nontumor control group,
Previous studies evaluating HRQOL in VS have carried several notable limitations including low patient numbers, lack of a normal non-VS control population, exclusion of one or more treatment groups, and short-term follow-up and have therefore been criticized for being methodologically weak. To date, only 3 prospective, nonrandomized studies have been published comparing HRQOL outcomes following treatment, and currently no Level 1 evidence exists. It is unlikely that a prospective randomized trial comparing treatment outcomes will ever materialize given the significant number of obstacles, including patient recruitment in a relatively uncommon condition, the number of subjects required to detect clinically meaningful differences, and significant practice disparities among many major centers, making multicenter collaborations difficult.

**Discussion**

Over the last century, the treatment of VS has undergone tremendous refinement resulting from improvements in surgical technique, advancements in radiation delivery, and a better understanding of the natural history of disease. We have now reached a relative plateau with several important objective clinical outcome measures. Namely, more than 90% of patients receive durable tumor control and satisfactory facial nerve function, and disease-specific mortality remains exceedingly low whether treated with microsurgery or SRS. While microsurgical resection remains the preferred treatment for tumors larger than 3 cm, management practices for small- to medium-sized tumors vary significantly between treatment centers, reflecting the relative lack of evidence supporting one therapy over another. Despite the large number of studies evaluating VS outcomes, there remains no consensus regarding optimal management.

Unfortunately, it is not uncommon that we, as health care providers, encounter a patient dissatisfied with his or her outcome despite having what we perceive to be an excellent treatment result. Over the past several decades there has been increasing importance placed on patient-centered outcomes, realizing that clinicians commonly underestimate patient impairment, and there is frequent disparity between what health care providers value and what patients prioritize. Additionally, treating physicians may not fully appreciate the psychological impact of the diagnosis of a brain tumor, even a benign one, and the significant impact it may have on HRQOL even prior to embarking on a plan of care.

Outcome assessment tools developed by the health care industry often evaluate a narrow range of technical outcomes and frequently fail to provide insight into how the disease itself or the outward effects of treatment influence overall patient well-being. With an increasing prevalence of small tumors in minimally symptomatic patients it is paramount that we carefully examine both traditional clinical measures and HRQOL data when determining best practices.

Previous studies evaluating HRQOL in VS have carried several notable limitations including low patient numbers, lack of a normal non-VS control population, exclusion of one or more treatment groups, and short-term follow-up and have therefore been criticized for being methodologically weak. To date, only 3 prospective, nonrandomized studies have been published comparing HRQOL outcomes following treatment, and currently no Level 1 evidence exists. It is unlikely that a prospective randomized trial comparing treatment outcomes will ever materialize given the significant number of obstacles, including patient recruitment in a relatively uncommon condition, the number of subjects required to detect clinically meaningful differences, and significant practice disparities among many major centers, making multicenter collaborations difficult.
In fact, Myrseth et al. had to abandon their initial plan to randomize enrollment, as patients were unwilling to relinquish their treatment to chance.22

Pollock et al. reported the first prospective nonrandomized study in 2006 including 82 patients with primary sporadic, smaller than 3-cm VSs who underwent microsurgery or SRS.25 The mean duration of follow-up was 42 months, and the Health Status Questionnaire (modification of the SF-36) was used to assess longitudinal change in HRQOL. The SRS cohort did not experience any deterioration in HRQOL, while the microsurgical arm saw an early decline in the standardized physical component score at 3 months following treatment, which subsequently returned to near baseline. At last follow-up the net change from baseline for the physical and mental summary scores for both treatment groups were very similar; there was a net loss of 3.1 and 3.6 points in the standardized physical component score and a net improvement of 4.5 and 3.7 points in the standardized mental component score for the SRS and microsurgery groups, respectively.

Subsequently, in 2009 Myrseth and colleagues published the second prospective nonrandomized study comparing SRS and microsurgery in a cohort of 91 Norwegian patients using the SF-36 and GBI to evaluate longitudinal changes in HRQOL at baseline, 1 year, and 2 years following treatment.22 At last follow-up, there were no statistical differences between treatment groups on individual or summary scores for the SF-36. Similarly, using the GBI, there were no differences between groups at 1 year after treatment. However, at 2 years there was a statistically significant decline in the GBI total score, and the general and physical subdomain scores for the microsurgery arm. At last follow-up, there were no statistically significant differences in job disability or sick leave status between groups.

Di Maio et al. published the most recent prospective study including 47 patients who underwent observation, 48 who received linear accelerator–based radiotherapy, and 97 who underwent microsurgical resection of VS smaller than 3 cm in size.7 At a mean follow-up of 31.8 months, there were no statistically significant differences between groups on either the SF-36 physical or mental health summary score; all 3 groups were within approximately 4 points of baseline values. Notably, HRQOL remained unchanged for the observation and radiotherapy cohorts throughout follow-up, while the surgical group experienced a significant improvement in mental health at 24 months, which subsequently returned to baseline at last follow-up.

In the current study, patients who underwent microsurgery had statistically significantly worse outcomes in their PANQOL total score as well as the subdomains of facial function, balance, and pain compared with subjects treated with SRS or observation. Considering that the incidence of objective facial nerve dysfunction following microsurgery was approximately 10% greater than SRS or observation, it is not surprising that these patients had a statistically worse facial PANQOL score. Furthermore, as primary VSs are rarely associated with trigeminal neuralgia or other pain syndromes and since prolonged headache following surgery is a known potential complication, it is not unexpected that the microsurgery cohort had an average poorer score than the SRS or observation groups in the pain subdomain as well. Several studies have demonstrated that ongoing dizziness, compared with other variables such as facial nerve dysfunction and hearing loss, has the most profound effect on HRQOL. It is interesting that the microsurgery cohort scored poorer in the PANQOL balance subdomain since it is commonly held that tumor resection is advantageous for treatment of patients who are experiencing debilitating vertigo.10 This seemingly paradoxical observation can be explained by the differences in dizzy symptoms patients experience following microsurgery, SRS, and observation as well as treatment selection bias. Microsurgery nearly always imparts a complete ipsilateral peripheral vestibular paresis, often resulting in acute postoperative vertigo; however, over a period of several weeks or months most subjects achieve central compensation with a significant improvement in symptoms. In contrast, patients who are observed or receive SRS often have residual ipsilateral vestibular function, which reduces the risk of mild symptoms associated with unilateral vestibular hypofunction (e.g., transient imbalance with quick head turns), but with a functioning labyrinth, patients may still experience incapacitating vertigo. Since the prevalence of long-term disabling vertigo is low, by sheer number, the milder symptoms of unilateral vestibular hypofunction seen in the surgical cohort dominates the PANQOL balance score rather than the few outlier patients with debilitating vertigo following SRS or observation. This may also be explained by treatment selection bias. In our practices, patients with growing tumors and incapacitating vertigo are more likely to receive microsurgery than SRS or observation. This strategy may inherently select for those that are more sensitive to dizziness, which will likely negatively influence balance scores following treatment.

There are several important considerations that should be taken into account when interpreting VS HRQOL data. First, it is generally held that microsurgery imparts upfront morbidity with subsequent improvement over time, while the untoward effects of radiation and observation are often not evident for many years following treatment; therefore, long-term analysis is critical to ensure fair comparison. To illustrate this point, Pollock et al. demonstrated that at 3 months after treatment, patients who underwent surgery had a statistically significant decline in several of the SF-36 domains, including physical functioning, role-physical, energy/fatigue, and overall physical.25 However, at last follow-up only a decline in bodily pain persisted. Furthermore, since most patients are diagnosed in their 4th to 5th decade of life and are expected to live 20 or more years beyond diagnosis, late HRQOL data should be prioritized over transitory HRQOL changes found early after treatment.

One other important consideration when reviewing HRQOL data is whether small statistically significant differences between groups are clinically meaningful. That is to say, is a several point discrepancy even perceptible to the patient or important to the clinician? For example, although statistically significant, the difference between observation and SRS on the PANQOL total score was only 2 points on a 100-point scale, and there were only 5 points separating SRS and microsurgery. The minimum clinically important difference (MCID) was developed to
assist in answering this complex but important question. The MCID represents the minimum change in score that a patient would perceive as an improvement over baseline for a particular condition. While the MCID of the SF-36 has not been specifically evaluated for patients with VS, the MCID for many other studied disease conditions is commonly greater than 5 points for the physical and mental health summary scores. Notably, none of the 3 prospective studies discussed earlier exceeded a mean 5 point change in either score at last follow-up.

Until very recently, one significant obstacle to HRQOL assessment has been the lack of a validated disease-specific instrument. By default, the SF-36 has become the most widely used instrument for assessing HRQOL outcomes within the VS literature. Importantly, the SF-36, GBI, and PROMIS-10 are multipurpose instruments and therefore are heavily influenced by patient comorbidity and may lack the sensitivity required to detect subtle changes after treatment or small differences between management arms. In 2010 Shaffer and colleagues introduced the disease-specific PANQOL scale, a 26-item questionnaire evaluating facial function, balance, hearing, pain, anxiety, energy, and general health. Initial evaluation yielded robust face, content, criterion, and construct validity with high test-retest reliability. The current study corroborates previous data demonstrating that the PANQOL is a more responsive instrument to disease- and treatment-associated changes than multipurpose measures such as the SF-36 or PROMIS-10. Specifically, after multivariable adjustment, the latter 2 instruments did not demonstrate any statistically significant differences between patients receiving SRS, observation, or microsurgery. Moreover, in almost every case the SF-36 and PROMIS-10 failed to detect differences between nontumor controls and patients with VS, while the PANQOL did so in almost every subdomain score. Moving forward, we believe that the PANQOL scale should supplant the SF-36 as the default instrument for VS HRQOL assessment. The next important step will be to estimate the MCID for the PANQOL total and subdomain scores from prospective analysis.

Since intervention does not offer an HRQOL advantage over observation, and because a significant number of small- to medium-sized VSs do not grow after diagnosis, tumors should be initially observed using serial imaging. We believe that it is hard to justify the added cost and potential risks, albeit small, of upfront SRS without evidence to support an HRQOL advantage over the natural history. If growth is documented on serial imaging, or additional tumor-related symptoms develop, active treatment with either SRS or microsurgery is indicated. For patients with tumor growth and minimal symptoms, SRS appears to offer the best HRQOL outcome, as long as long-term tumor control is achieved. Finally, microsurgery should be considered for patients with mass effect symptoms, growing cystic tumors, or strong preference.

In the current study, PANQOL scores were reduced in patients with VS compared with general population nontumor controls; however, the differences between individual management groups at 8 years following treatment were relatively small. These data support the notion that the diagnosis of VS imparts a significant reduction on HRQOL, while long-term differences between treatment strategies are generally less significant by comparison. We believe that further meaningful gains in HRQOL will require a redirection of efforts. While facial nerve function and tumor control are important, results from large-volume centers have largely plateaued within the last several decades. Moving forward, improvements in patient counseling regarding realistic long-term expectations of disease, as well as offering pre- and posttreatment psychosocial support, may provide the most impact on HRQOL improvement for patients with VS. For example, it is intriguing that there was no difference between groups in the PANQOL hearing domain despite patients undergoing SRS and microsurgery having significantly higher rates of severe hearing loss following treatment. This very well could be explained by the fact that we extensively counsel patients about the risks and consequences of unilateral deafness at initial consultation and during follow-up.

We wish to highlight some strengths of the current study and finally acknowledge several limitations. This is the first multicenter study evaluating HRQOL in VS and includes the largest number of patients to date. The combination of a high questionnaire response rate and international multicenter participation permitted enrollment of a large number of subjects, improving statistical inference, and enhancing external validity, and, therefore, the generalizability of our findings. Second, the majority of high-volume tertiary centers carry a strong preference for either SRS or microsurgery, and, consequently, most studies exclude one or more treatment arms. In the current study, both centers have significant experience managing VS using SRS and microsurgery with a relatively even distribution between treatment strategies. Additionally, a significant number of small- to medium-sized tumors are conservatively followed with serial MRI and audiograms at biannual or annual visits. As a result, we were able to compare HRQOL outcomes between all 3 management arms, in addition to a group of nontumor control subjects. Third, all patients were surveyed using 3 general HRQOL instruments and a recently validated disease-specific questionnaire to minimize the risk of monomethod bias present in earlier studies. To our knowledge, the PROMIS-10 has not been previously used in VS HRQOL assessment. Finally, long-term HRQOL analysis is critical to evaluating differences between microsurgery, SRS, and observation. The mean time interval between treatment and survey was 7.7 years—more than twice the follow-up of previous studies.

The primary limitations of the current study are nonrandomized treatment and the lack of longitudinal follow-up. Without patient randomization to observation, SRS, or microsurgery, it remains possible that treatment selection bias may result in baseline differences between groups, which may in turn influence long-term HRQOL outcome results. For example, patients who have resigned themselves to the diagnoses and are able to cope with disease-related symptoms are more likely to be observed or receive SRS, while those who are overly anxious about the diagnosis of a brain tumor or are most bothered by what are felt to be tumor-related symptoms, such as trigeminal neuropathy or recurrent vertigo, are more likely to receive microsurgery. Therefore, it is possible that the
slightly lower scores in the microsurgery arm are the result of pretreatment patient-related factors rather than treatment. Furthermore, without pretreatment HRQOL scores and longitudinal follow-up, time-dependent changes in HRQOL and causality between intervention and outcome cannot be confidently established. Finally, both centers in the current study are tertiary academic hospitals, and therefore it is possible that the VS population is biased by referral patterns. Once again, while considered the gold standard, randomized controlled trials comparing management strategies may never materialize given the relative rarity of disease, reluctance of patients to resign control over choice of management, strong biases of treating physicians, and the large number of patients that would be required to detect meaningful differences.22,34

Conclusions
Overall, the HRQOL differences following observation, microsurgery, and SRS are small. The diagnosis of VS, rather than treatment strategy, most significantly impacts patient quality of life. Therefore, further improvements in patient counseling regarding realistic long-term expectations of disease, as well as offering pre- and posttreatment psychosocial support, may provide the most impact on HRQOL improvement for patients with VS.

Since a significant number of tumors do not grow after the time of diagnosis, and because intervention does not appear to confer an HRQOL advantage over observation, small- and medium-sized VSs should be initially observed, while active treatment should be reserved for patients with unequivocal tumor growth or intractable symptoms that are amenable to treatment. As disease-specific HRQOL outcomes following SRS appear to be minimally better than microsurgery, we believe that SRS should be offered to those with growing tumors and minimal symptoms, while microsurgery should be considered for patients with symptoms associated with mass effect, enlarging cystic tumors, or in cases of patient preference. Moving forward, studies assessing HRQOL in VS patients should prioritize use of validated disease-specific measures, such as the PANQOL, given the significant limitations of generic instruments in distinguishing between treatment groups and tumor versus nontumor subjects.

Acknowledgment
We thank Christine M. Lohse, M.S., from the Department of Health Sciences Research for her valuable assistance with statistical analysis and manuscript review.

References
22. Myrseth E, Möller P, Pedersen PH, Lund-Johansen M: Vestibular schwannoma: surgery or gamma knife radiosur-

**Author Contributions**
Conception and design: Link, Carlson, Tveiten, Tombers, Lund-Johansen, Acquisition of data: Link, Carlson, Tveiten, Tombers, Castner, Finnkirk, Myrseth, Pedersen, Lund-Johansen. Analysis and interpretation of data: all authors. Drafting the article: all authors. Critically revising the article: all authors. Reviewed submitted version of manuscript: Link, Carlson, Tveiten, Driscoll, Neff, Pollock, Tombers, Castner, Finnkirk, Pedersen, Lund-Johansen.

**Correspondence**
Michael J. Link, Department of Neurologic Surgery, Mayo Clinic, 200 First St. SW, Rochester, MN 55905. email: link.michael@mayo.edu.