Methodological issues in longitudinal studies: vestibular schwannoma growth rates in neurofibromatosis 2

Michael E. Baser,1 Victor-F. Mautner,2 Dilys M. Parry,3 D. Gareth R. Evans4

1Los Angeles, CA, U.S.A., 2Department of Maxillofacial Surgery, University Hospital Eppendorf, Hamburg, Germany, 3Genetic Epidemiology Branch, National Cancer Institute, Bethesda, MD, U.S.A., 4Academic Department of Medical Genetics, St. Mary's Hospital, Manchester, U.K.

Correspondence and reprints to:
Michael E. Baser, Ph.D.
10622 Kinnard Ave., #203
Los Angeles, CA 90024
e-mail: michael.baser@verizon.net
tel: 310-474-1608; fax: 310-472-3614
Abstract

Introduction: Four longitudinal studies of vestibular schwannoma (VS) growth rates in neurofibromatosis 2 (NF2) have yielded very different results on the relationship of VS growth rates to age. The study designs differed with respect to patient eligibility criteria, indices of VS growth rates, VS volumetric methods, and sample sizes. The purposes of this study were to determine the most likely reason for the different results and to determine the actual relationship of VS growth rates to age.

Methods: Each of the differences in study design was tested as a possible reason for the different results using data from two of the longitudinal studies and data from the population-based United Kingdom NF2 Registry.

Results: The patient eligibility criterion in one study caused selection bias for slower-growing VS. The proper interpretation of the results from the four studies is that VS growth rates in NF2 are highly variable but tend to decrease with increasing age.

Discussion: Clinical trials for VS in NF2 should focus on younger patients because VS growth rates tend to decrease with increasing age, and because faster-growing VS are more likely to be excised with increasing age than slower-growing VS.

Key words: neurofibromatosis 2; NF2; vestibular schwannoma; longitudinal; natural history
NF2 is a rare autosomal dominant disease that is characterized by vestibular schwannomas (VS), meningiomas, non-VS schwannomas, and ocular abnormalities. Bilateral VS, which occur in almost all adults with NF2, are considered to be the hallmark of the disease. The “gold standard” for imaging VSs is gadolinium-enhanced magnetic resonance imaging (GE-MRI).

Four longitudinal studies of VS growth rates in NF2 using GE-MRI have yielded very different results on the fundamental question of the relationship of VS growth rates to age. In the studies of Baser et al. and Mautner et al., VS growth rates decreased with increasing age. Data from these two studies can be combined because each study had the same patient eligibility criteria and volumetric methods. In combined data from the two studies, age at baseline VS measurement accounted for 23% of the variance in VS growth rates, and 79 (85%) of the 93 VS changed by more than five mm$^3$ per year. In the study of Slattery et al., VS growth rates did not significantly change with increasing age. VS growth rates were not reported in terms of volumes, but the average change in greatest VS diameter in 21 patients with the longest follow-up was only 1.9 mm per year; only four (13%) of the 30 VS in these patients changed by at least five mm in one or more linear dimensions during the study period. In the study of Abaza et al., VS growth rates increased with increasing age.

An understanding of the natural history of NF2 is necessary to properly design and interpret clinical trials. What can be learned from the conflicting results of these longitudinal studies? To answer this question, the reasons for the different results must be identified. The studies had different patient eligibility criteria, indices of VS growth rates, VS volumetric methods, and sample sizes (Table 1). However, not all of these differences are equally likely explanations for the conflicting results.
Table 1. Description of four longitudinal studies of VS growth rates in NF2

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design</th>
<th>Eligibility criterion for VS measurements</th>
<th>No. of patients</th>
<th>Length of follow-up</th>
<th>Volumetric method</th>
<th>Index of V*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abaza et al. (1996)</td>
<td>Review of previously collected MRIs</td>
<td>All VS measurements, including post-operative measurements on recurrent or residual VS</td>
<td>22</td>
<td>3.7 years (mean)</td>
<td>Box model</td>
<td>Volume c</td>
</tr>
<tr>
<td>Baser et al. (2002)</td>
<td>Review of previously collected MRIs</td>
<td>All pre-operative VS measurements, even if both VS were resected prior to enrollment</td>
<td>17</td>
<td>4.0 years (median)</td>
<td>Box model</td>
<td>Tumor dc</td>
</tr>
<tr>
<td>Mautner et al. (2002)</td>
<td>Review of previously collected MRIs</td>
<td>All pre-operative VS measurements, even if both VS were resected prior to enrollment</td>
<td>38</td>
<td>3.9 years (median)</td>
<td>Box model</td>
<td>Tumor dc</td>
</tr>
<tr>
<td>Slattery et al. (2004)</td>
<td>Prospective</td>
<td>At least one currently unoperated VS</td>
<td>56</td>
<td>2.4 years (mean) in 21 patients with long-term follow-up</td>
<td>Linear dimensions (AP, ML, GP)</td>
<td>Change in</td>
</tr>
</tbody>
</table>

VS = vestibular schwannoma, NF2 = neurofibromatosis 2, TDT = tumor doubling time, AP = anterior-posterior, ML = medial-lateral, GD = greatest diameter.

1 Each of these studies was based on patients from the U.S. National Institutes of Health longitudinal study of NF2.

2 Data from these two studies were re-analyzed in the present study using cm³ per year as the index of VS growth rate.
Previously, we re-evaluated data from the study of Abaza et al. and noted that the study included post-operative measurements on recurrent or residual VS, which could cause a bias toward apparently faster-growing VS in older people. In the present study, we evaluated possible reasons for the discrepant results of the other three longitudinal studies by empirically testing the effects of the differences in study design noted above. We re-analyzed data from two of the three longitudinal studies and used data from the population-based United Kingdom NF2 registry.

**Patient eligibility criteria.** In the studies of Baser et al. and Mautner et al. (which each found that VS growth rates decreased with increasing age), patients who had both VS excised prior to enrollment were still eligible for study if pre-operative VS volumetric measurements were available. But in the study of Slattery et al. (which found that VS growth rates did not change with increasing age), patients were eligible for study only if they currently had at least one unoperated VS. This eligibility criterion causes selection bias for slower-growing VS because several conditions are fulfilled.

First, excised VS grow significantly faster prior to excision than non-excised VS. Using linear regression analysis of the combined data from Baser et al. and Mautner et al., there was a significant difference between the mean growth rates of excised VS and non-excised VS when the index of VS growth rates was either cm$^3$ per year or tumor doubling time (TDT) (Table 2).
Table 2. Results of linear regression analyses with VS growth rates as the outcome

<table>
<thead>
<tr>
<th>Index of VS growth rate and covariates</th>
<th>Regression coefficient</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cm³ per year (log₁₀)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excised VS compared to non-excised VS</td>
<td>0.85</td>
<td>0.11 to 1.58</td>
</tr>
<tr>
<td>Age at baseline VS measurement (per year)</td>
<td>-0.04</td>
<td>-0.07 to –0.20</td>
</tr>
<tr>
<td>TDT (log₁₀)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excised VS compared to non-excised VS</td>
<td>-0.38</td>
<td>-0.64 to –0.12</td>
</tr>
<tr>
<td>Age at baseline VS measurement (per year)</td>
<td>0.01</td>
<td>0.00 to 0.02</td>
</tr>
</tbody>
</table>

CI = confidence interval of regression coefficient, VS = vestibular schwannoma, TDT = tumor doubling time.
There are two separate linear regression analyses with outcomes of cm³ per year and TDT. TDT is calculated as twice the baseline VS volume divided by the average annual VS growth rate (in cm³ per year). The analysis is based on combined data of 55 patients from references 4 and 5.
Second, the proportion of NF2 patients with both VS excised increases with increasing age. Using Kaplan-Meier analysis of data from 479 patients in the United Kingdom NF2 registry, the cumulative proportion of NF2 patients with excised bilateral VS (i.e., who would be ineligible for the study of Slattery et al.6) was about 1% by age 20, 3% by age 25, and thereafter increased more rapidly to 37% by age 50. Finally, the study of Slattery et al.6 has only a small proportion of young NF2 patients, the group that is least likely to be affected by selection bias. Only about 15% of NF2 patients in the study were enrolled before age 20 (Laurel Fisher, personal communication). In sum, the eligibility criterion in the study of Slattery et al.6 causes selection bias for slower-growing VS because excised VS grow faster prior to excision than non-excised VS, faster-growing VS are more likely to be excised with increasing age than slower-growing VS, and the study has only a small proportion of young NF2 patients.

Another type of selection bias for slower-growing VS would occur if people with severe NF2, who have an early age at onset of symptoms and a high risk of mortality,8 also have high VS growth rates. The increased risk of mortality in NF2 patients with severe disease would cause selection for older NF2 patients with slow-growing VS. There are too few deceased NF2 patients with pre-mortem longitudinal VS growth measurements to directly address this question. However, selection bias due to an increased risk of mortality in NF2 patients with severe disease probably would have a similar effect on all longitudinal studies and therefore would not explain the different results of the studies under consideration.

Indices of VS growth rates. Many indices can be used to estimate VS growth rates. TDT is a widely-used index. In the studies of Baser et al.4 and Mautner et al.5, TDT was calculated as twice the baseline VS volume divided by the VS growth rate, where growth rate (in cm³ per year) was calculated as a linear fit through sequential VS volumes. Baseline VS volume is factored into TDT because the growth rate of VS may depend, in part, on their size.

Slattery et al.6 suggested that TDT might accentuate change over time because, in several studies, there was not a significant relationship between baseline VS volumes and subsequent VS growth rates.6,10,11 We evaluated this relationship in the combined data of Baser et al.4 and Mautner et al.5 using logarithmic transformations because baseline VS volumes and VS growth rates each ranged from low to high values. There was a strong association between baseline VS volumes and subsequent VS growth rates (Figure 1). It is possible that previous studies did not find this association because the data were not transformed. Also, two of the cited studies were of unilateral sporadic VS, not NF2-associated VS.10,11 Compared to unilateral sporadic VS, NF2-associated VS are more lobular, less vascular, and have higher proliferation indices.12,13

In some studies, VS growth rates have been estimated using a model that assumes logarithmic growth,14,15 and it has been suggested that linear models of VS growth are less appropriate with increasing length of observation because the underlying growth pattern is sigmoidal.14 However, VS growth patterns are highly variable and the probability of a sigmoidal VS growth pattern in NF2 patients is not significantly associated with the length of observation.5

VS growth curves have many patterns, but a single type of pattern does not predominate, so a linear fit will not systematically overestimate VS growth rates in the aggregate. Some common patterns of VS growth are increasing but asymptotic growth, exponentially increasing
growth, linearly increasing growth, and no growth. A linear fit overestimates growth in VS with increasing but asymptotic growth, underestimates growth in VS with exponentially increasing growth, and accurately estimates growth in VS with linearly increasing growth or no growth.

Because VS growth patterns are highly variable, it is most practical to use a curve fitting method that can be applied to all growth curves, as is the case for a linear fit. In addition to a linear fit, other indices of VS growth that do not depend on any particular pattern of growth are:
(a) The proportion of VS that increase by a specified change in volume within a set time.
(b) Kaplan-Meier analysis of VS volumes that increase by a specified amount, in which "failure" is defined as the VS volume increasing by the specified amount.
(c) The proportion of VS that increase in size more than is attributable to measurement error within a set time (or over time, as in a survival analysis). This requires statistical definition of "no growth", which can be determined by repeated measures of the same VS over a short period of time in a few patients.

VS volumetric methods. Precise MRI-determined volumetrics are more accurate than box models or single-dimension linear measurements, but there are several sources of variability in VS growth rates and volumetric accuracy is not necessarily the most important one. Overall variability in VS growth rates is the sum of biological variability (variability over time in the same person and variability between people) and measurement error (e.g., from approximations to VS volumes using box models). In most cases, the magnitude of the biological variability of VS growth rates is much higher than the magnitude of measurement error.

To estimate the effect of volumetric precision on VS growth rates, we compared VS growth rates that were obtained when two different types of box models were used (we did not have information on MRI-determined volumes). Box models estimate tumor volumes by setting a three-dimensional box around the tumor. The three dimensions are the maximum diameters in anterior-posterior, medial-lateral, and superior-inferior dimensions. In the one-component model, a single box is fit to the entire VS. In the two-component model, separate boxes are fit to extracanalicular and intracanalicular parts of the VS. The two-component model is more accurate than the one-component model when the VS has extracanalicular and intracanalicular parts. In Figure 2, data from the study of Baser et al. is used to compare the estimated VS growth rates of 26 VS with both extracanalicular and intracanalicular parts using one- and two-component models. There is a strong association, suggesting that less precise volumetrics do not have a major overall effect on estimated VS growth rates.

Sample size. Slattery et al. suggested that the decrease in VS growth rates with increasing age in the studies of Baser et al. and Mautner et al. could be due to small sample size. But data from the latter two studies can be combined, and as expected, the same general relationship obtains in the combined group of 55 patients as in each of the individual studies. Using linear regression analysis, VS growth rates decreased with increasing age at baseline VS measurement, whether the index of VS growth rate was cm³ per year (r² = 0.23, P < .001) or TDT (r² = 0.12, P = .010). It is more important to note that VS growth patterns are highly variable and that none of the studies under consideration has a sufficient number of patients to adequately characterize this variability.
This study and our previous study\textsuperscript{4} illustrate that erroneous results in longitudinal studies can be caused by flawed study design or by flawed data analysis. In the study of Slattery et al.\textsuperscript{6}, the patient eligibility criterion causes selection bias for slower-growing VS. The proper interpretation of the results from the four longitudinal studies\textsuperscript{4,5,6,7} is that VS growth rates in NF2 are highly variable but tend to decrease with increasing age. Clinical trials for VS in NF2 should focus on younger patients because VS growth rates tend to decrease with increasing age, and because faster-growing VS are more likely to be excised with increasing age than slower-growing VS.

Acknowledgments

We thank J. M. Friedman and Harry Joe for helpful discussions.

Figure legends

Figure 1. Association of baseline VS volume with subsequent VS growth rate. The analysis is based on 101 VS that changed in volume during the study period (combined data from references 4 and 5). The slope of the linear regression analysis was 0.87 cm\textsuperscript{3}/year of VS growth rate per unit (in cm\textsuperscript{3}) of baseline VS volume (95\% CI, 0.69 to 1.02). A separate analysis was done on 122 VSs that included 21 VS that did not change in volume during the study period (graph not shown). The slope of the linear regression analysis was 0.80 cm\textsuperscript{3}/year of VS growth rate per unit (in cm\textsuperscript{3}) of baseline VS volume (95\% CI, 0.54 to 1.06).

Figure 2. Comparison of VS growth rates (in cm\textsuperscript{3} per year) using one- and two-component box models of VS volumes (Wilcoxon signed rank test, \(P = .008\)). The analysis is based on 26 VS with extracanalicular and intracanalicular parts (data from reference 4). Two outliers are not pictured for graphical clarity.

Licence for publication

The corresponding author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd to permit this article (if accepted) to be published in the Journal of Medical Genetics and any other BMJPGL products and sublicences such use and exploit all subsidiary right, as set out in our licence (http://JMG.bmjournals.com/misc/ifora/licenseform.shtml).

References


VS growth rate, two-component model (cc per year)

VS growth rate, one-component model (cc per year)