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

M E Baser, V-F Mautner, D M Parry, D G R Evans

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 studies had different patient eligibility criteria, indices of VS growth rates, VS volumetric methods, and sample sizes. We reanalysed data from two of the longitudinal studies and data from the population based United Kingdom NF2 Registry to determine the most likely reason for the different results and the actual relationship of VS growth rates to age. We found that the 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. The studies had different patient eligibility criteria, indices of VS growth rates, VS volumetric methods, and sample sizes. However, not all of these differences are equally likely explanations for the conflicting results.

Previously, we re-evaluated data from the study of Abaza et al, and noted that the study included postoperative 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 reanalysed data from two of the three longitudinal studies and used data from the population based United Kingdom NF2 Registry.

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 design and interpret clinical trials properly. 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. However, not all of these differences are equally likely explanations for the conflicting results.

N2 is a rare autosomal dominant disease characterised 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 19% of the variance in VS growth rates, and 75 (82%) of the 93 VS grew at least 5 mm/year. In the study of Slattery et al, VS growth rates did not change significantly with increasing age. VS growth rates were not reported in terms of volumes, but in patients with longer follow-up, only four (13%) of the 30 VS in these patients grew at least 5 mm in one or more linear dimensions during the study period.

In the study of Slattery et al, VS growth rates increased with increasing age.

Previously, we re-evaluated data from the study of Abaza et al, and noted that the study included postoperative 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 reanalysed data from two of the three longitudinal studies and used data from the population based United Kingdom NF2 Registry.

SOURCES OF ERROR
Patient eligibility criteria

In the studies of Baser et al and Mautner et al (each of which found that VS growth rates decreased with increasing age), patients who had undergone excision of both VS prior to enrolment were still eligible for study if pre-operative VS volumetric measurements were available. In contrast, Slattery et al (who found that VS growth rates did not change with increasing age) included patients in the study only if they had at least one unoperated VS at the time. This eligibility criterion causes selection bias for slower growing VS because several conditions are fulfilled. Firstly, the 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 and non-excised VS when the index of VS growth rates was either cm/year or tumour doubling time (TDT) (table 2). Secondly, 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.

Abbreviations: GE-MRI, gadolinium enhanced magnetic resonance imaging; NF, neurofibromatosis; TDT, tumour doubling time; VS, vestibular schwannoma
NF2 Registry, the cumulative proportion of NF2 patients with excised bilateral VS (that is, those who would be ineligible for the study of Slattery et al.6) was about 1% by 20 years of age, 3% by 25 years, and thereafter increased more rapidly, reaching 37% by 50 years. Finally, the study of Slattery et al.6 had only a small proportion of young NF2 patients, the group least likely to be affected by selection bias. Only about 13% of NF2 patients in the study were <20 years old (L. Fisher; personal communication). In sum, the eligibility criterion in the study of 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.4,6,10–11 We evaluated this relationship in the combined data of Baser et al.5 and Mautner et al.4 Logarithmic transformations were used 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 (fig 1). It is possible that previous studies did not find this association because the data were not transformed. In addition, two of the cited studies were of unilateral sporadic VS, not NF2 associated VS.10–11 Compared with unilateral sporadic VS, NF2 associated VS are more lobular, less vascular, and have higher proliferation indices.12–13

In some studies, VS growth rates are 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.16 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.17 VS growth curves have many patterns and a single 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

### Table 1: Description of four longitudinal studies of VS growth rates in NF2

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Study design</td>
<td>Review of previously collected MRIs</td>
<td>Review of previously collected MRIs</td>
<td>Review of previously collected MRIs</td>
<td>Prospective</td>
</tr>
<tr>
<td>Eligibility criterion for VS measurements</td>
<td>All VS measurements, including postoperative measurements on recurrent or residual VS</td>
<td>All pre-operative VS measurements, even if both VS were resected prior to enrolment</td>
<td>All pre-operative VS measurements, even if both VS were resected prior to enrolment</td>
<td>At least one currently unoperated VS</td>
</tr>
<tr>
<td>No. of patients</td>
<td>22</td>
<td>17</td>
<td>38</td>
<td>38</td>
</tr>
<tr>
<td>Length of follow up</td>
<td>3.7 years (mean)</td>
<td>4.0 years (median)</td>
<td>3.9 years (median)</td>
<td>2.4 years (mean) in 21 patients with long term follow up</td>
</tr>
<tr>
<td>Volumetric method</td>
<td>Box model</td>
<td>Box model</td>
<td>Box model</td>
<td>Linear dimensions (AP, ML, GP)</td>
</tr>
<tr>
<td>Index of VS growth rate</td>
<td>Volume change (in cm³)</td>
<td>TDT</td>
<td>TDT</td>
<td>Change in linear dimension</td>
</tr>
<tr>
<td>Relationship of VS growth rate to increasing age</td>
<td>Increasing</td>
<td>TDT</td>
<td>Decreasing</td>
<td>No significant association</td>
</tr>
</tbody>
</table>

VS, vestibular schwannoma; NF2, neurofibromatosis 2; TDT, tumor doubling time; AP, anterior-posterior; ML, medial-lateral; GD, greatest diameter. *Each of these studies was based on patients from the U.S. National Institutes of Health longitudinal study of NF2. †Data from these two studies were re-analysed in the present study using cm³/year as the index of VS growth rate.

### 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/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/year and TDT. TDT is calculated as twice the baseline VS volume divided by the average annual VS growth rate (in cm³/year). The analysis is based on combined data of 55 patients from Baser et al. and Mautner et al.
with increasing age in the studies of Baser et al.4 and Mautner et al.5 Suggested that the decrease in VS growth rates with increasing age in the studies of Baser et al.4 and Mautner et al.5 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 volume increasing by the specified amount; and (c) the proportion of VS that increase in volume more than is attributable to measurement error within a set time (or over time, as in a survival analysis). The latter method 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.

Volumetric methods for vestibular schwannoma

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 (for example, 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 obtained when two different types of box models were used (we did not have information on MRI determined volumes). Box models estimate tumour volumes by setting a three dimensional box around the tumour. 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 fitted to the entire VS. In the two component model, separate boxes are fitted to the 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.6 In fig 2, data from the study of Baser et al.4 was used to compare the estimated VS growth rates of 26 VS with both extracanalicular and intracanalicular parts, using one and two component models. There was 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.2 suggested that the decrease in VS growth rates with increasing age in the studies of Baser et al.4 and Mautner et al.5 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/year ($r^2 = 0.23$, $p < 0.001$) or TDT ($r^2 = 0.12$, $p = 0.010$). 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 characterise this variability adequately.

CONCLUSION

This study and our previous study1 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.,4 the patient eligibility criterion causes selection bias for slower growing VS. The proper interpretation of the results from the four longitudinal studies4–7 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 are slower growing VS.

ACKNOWLEDGEMENTS

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

Authors’ affiliations

M E Baser, Los Angeles, CA, USA
V F Mautner, Department of Maxillofacial Surgery, University Hospital Eppendorf, Hamburg, Germany
D M Parry, Genetic Epidemiology Branch, National Cancer Institute, Bethesda, MD, USA
D G R Evans, Academic Department of Medical Genetics, St Mary’s Hospital, Manchester, UK

REFERENCES


