

Follow-up assessment of vestibular schwannomas: volume quantification versus two-dimensional measurements

Rick van de Langenberg · Bert Jan de Bondt ·
Patty J. Nelemans · Brigitta G. Baumert ·
Robert J. Stokroos

Received: 14 February 2009 / Accepted: 13 April 2009 / Published online: 6 May 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract

Introduction A conservative treatment strategy is often proposed as a primary treatment option in the management of vestibular schwannomas (VS). In this “wait and scan” policy, audiovestibular symptoms are monitored regularly, and VS growth is measured on consecutive magnetic resonance images (MRI). The aim of this study is validation of two-dimensional versus volume MRI assessment in the longitudinal follow-up of VS and to define tumor growth beyond measurement error.

Methods MRI scans of 68 consecutive patients with VS were analyzed retrospectively. Two-dimensional and volume measurements on contrast enhanced (CE) T1- and T2-weighted images were performed independently by two readers. Smallest detectable differences (SDD) were calculated, and intraclass correlation coefficients (ICCs) were determined for both assessment methods.

Results Two-dimensional and volume measurements both showed best reproducibility on CE T1-weighted images.

SDD for differences relative to baseline MRI [SDD (%) for two-dimensional measurements had a higher interobserver error compared to volume measurements (40% versus 19.7%), which decreases when tumor size increases. The ICC for two-dimensional measurements in three directions was 0.947, 0.974, and 0.978 and for volume measurements 0.999.

Conclusion Volume measurements are more accurate compared to two-dimensional measurements for the evaluation of VS growth. These measurements are assessed preferably on CE T1-weighted images. SDD (%) strongly depends on VS size. SDD between consecutive scans exceeds the common clinical applied criterion of 1 or 2 mm growth to define growth.

Keywords Vestibular schwannoma · Magnetic resonance imaging · Validity · Volume measurement

Introduction

The wide availability and technical improvement of magnetic resonance imaging (MRI) has led to an increase in detection of vestibular schwannoma (VS) at an early stage [1]. Accordingly, the incidence of diagnosed VS has increased [2, 3]. The natural development of VS remains uncertain as growth percentages between 30% and 90% have been reported, depending at least in part on the length of the observation period [2, 4]. So far, no clinical parameters have been identified to correlate with VS growth [5–7], and therefore, VS growth is objectified by performing consecutive MRI [2, 5]. If growth is found on MRI, an intervention may be chosen, such as surgical resection or radiation therapy [6, 8]. Most patients therefore enter the so-called wait and scan policy for a certain period, in which the audiovestibular

R. van de Langenberg (✉) · R. J. Stokroos
Department of Otorhinolaryngology—Head and Neck Surgery,
Maastricht University Medical Center,
P.O. Box 5800, 6202 AZ Maastricht, The Netherlands
e-mail: rickvandelangenberg@hotmail.com

B. J. de Bondt
Department of Radiology, Isala klinieken,
Zwolle, The Netherlands

P. J. Nelemans
Department of Epidemiology, Maastricht University,
Maastricht, The Netherlands

B. G. Baumert
Department of Radiation Oncology (MAASTRO),
GROW Research Institute, Maastricht University Medical Center,
Maastricht, The Netherlands

symptoms are monitored regularly and VS growth is measured on consecutive MRI [2].

Radiologists generally use two-dimensional measurements to assess VS growth, although volume measurements seem to provide more accurate growth assessment [1, 9, 10] since VS shows asymmetric growth in all directions. Little information is published about the diameter increase or volume increment between subsequent images that constitute to VS growth beyond measurement error [9–12]. Usually, when an increase in tumor diameter or volume is found, it is considered to be growing, but validation of this observation is lacking. Especially when invasive treatment decisions are based on these observations, it is of great importance to find the most suitable method to assess growth of VS on MRI and to provide a definition of growth beyond measurement error.

This study focuses on the accuracy and reproducibility of VS volume measurements compared to two-dimensional measurements to determine VS growth on MRI. The hypothesis is that measurement of tumor volume with specific area tracing software is a more accurate tool compared to two-dimensional measurements for determining tumor growth.

Materials and methods

All patients who received an MRI scan of the cerebellopontine angle (CPA) between January 2003 and March 2008 in our tertiary referral center were analyzed retrospectively. Patients were included in this study if a radiological diagnosis of a VS was made, resulting in 102 patients. Thirty patients were excluded who had been treated by surgery or radiotherapy, resulting in 72 patients. Four patients who had an intralabyrinthine schwannoma were also excluded. There were no patients with neurofibromatosis type 2. MRI images of 68 patients, 32 (47%) men and 36 (53%) women, age range of 36–84 years, median age of 63.5 years, were available; one scan was available in 21 patients, two scans were available in 22 patients, three scans in ten patients, four scans in eight

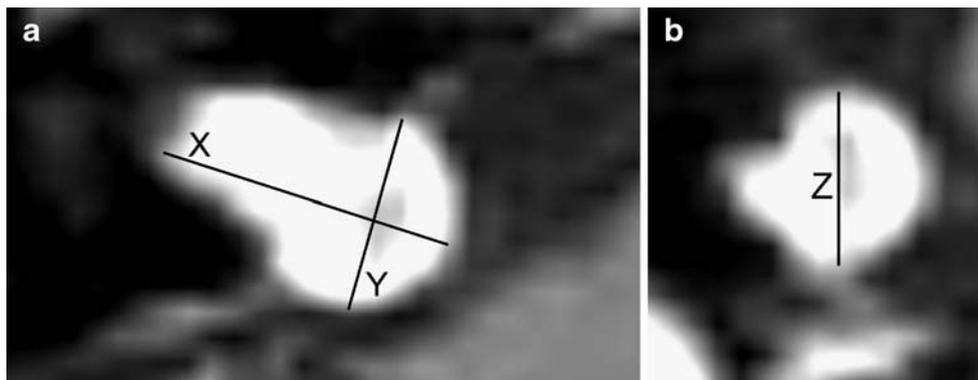
patients, five scans in four patients, and six scans in three patients, resulting in a total of 165 scans suitable for analysis. In patients with more than one scan, mean follow-up was 21.8 months (SD 15.7).

All examinations were performed at 1.5 T (Gyroscan, Powertrack 6000, Philips, Best, The Netherlands) using a head–neck coil (Philips, Best, The Netherlands). The MR protocol consisted of axial 2D SE T1-weighted images (TR/TE, 550/15 ms; slice thickness, 3 mm; inter slice gap, 0.3 mm; number of slices, 12; FOV, 180 mm (RFOV 80%); and matrix 256×256), axial 3D TSE T2-weighted images (TR/TE, 3,000/250 ms; slice thickness, 0.35 mm; number of overcontiguous slices, 30; FOV, 130 mm (RFOV 80%), and matrix 256×256) covering the skull base, and contrast-enhanced (gadolinium 0.2 ml/kg body weight) axial 3D ISO T1-weighted images (TR/TE, 8.9/4.6 ms; slice thickness, 1 mm; FOV, 256 (RFOV 80%); and matrix 256×256) covering the entire skull base and cranium. All patients underwent the same MRI protocol with similar parameters and planes of acquisition to ascertain an optimal correlation in serial scans.

Two readers, experienced in head and neck imaging, independently performed the measurements on contrast-enhanced T1-weighted images (CE T1-WI) and on the corresponding T2-weighted images (T2-WI). Both observers were blinded to each other's MR assessments and clinical information.

For two-dimensional assessment of VS, the maximum diameter was measured in three diameters: anteroposterior (AP), mediolateral (ML) [including the portion in the internal auditory canal (IAC)], and craniocaudal (CC) (Fig. 1a, b). To establish these dimensions, a digital submillimeter ruler was used. Volume assessment was done on a stereotactic radiotherapy treatment planning station, fitted with iPlan® RT image version 3—Advanced Contouring Workstation (BrainLAB Oncology Solutions, Feldkirchen, Germany). MR images were uploaded in this system, and area tracing software was used to outline the VS on each MR image that contained tumor tissue (Fig. 2a). If there was a sharp contrast with surrounding

Fig 1 a, b Contrast-enhanced T1-weighted image with a vestibular schwannoma in the cerebellopontine angle (CPA) on the right side. **a** Measurements in the axial plane: *X* is the maximum mediolateral, and *Y* the maximum anteroposterior dimension; **b** in the coronal plane, the *Z* demonstrates the maximum craniocaudal dimension



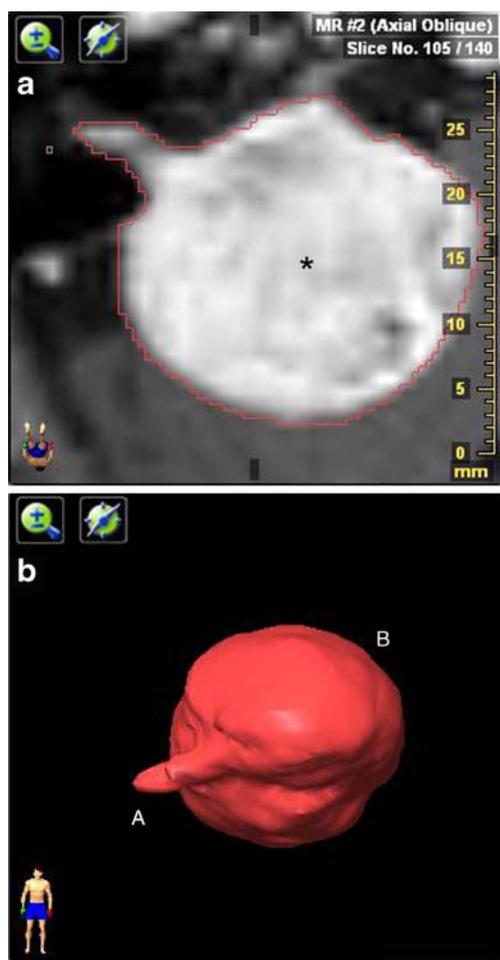


Fig. 2 a: Example of area tracing with volume software. Axial contrast-enhanced T1-weighted image shows a right-sided vestibular schwannoma (*asterisk*) with a large cerebellopontine angle component. The *red line* is the result of the autotracer which lines the vestibular schwannoma. **b** Three dimensional representation of a vestibular schwannoma (VS), integrating the surface of all slice intervals. The small intracanalicular (*A*) and large extracanalicular (*B*) portion of the VS can easily be identified

tissue, the auto brush function (surrounding the VS automatically) was used. Manual segmentation was necessary in cases in which differentiation with surrounding tissue was difficult because of the high sensitivity of this autotracer. Each segmentation result was checked visually. By tracing the VS surface on all slices, the software was able to calculate VS volume (Fig. 2b). Volumetric analysis was expressed in cubic centimeter.

To compare reproducibility of the measurements in different size categories, VS were classified into four stages, as defined by Hasegawa et al. [13]: stage A, intracanalicular VS; stage B, VS extending into the CPA; stage C, VS compressing the brain stem; and stage D, VS deviating the fourth ventricle.

Statistical analysis

Reproducibility measures consist of agreement and reliability parameters [14]. These parameters were evaluated for measurements by two different readers at one point in time. SPSS 15.0 statistical software (SPSS, Chicago, IL, USA) was used to perform the statistical calculations. For evaluation of interobserver reproducibility, we used baseline measurements of the MRI scans from the 68 patients for both readers, which can be considered as independent observations, whereas consecutive measurements within patients are correlated.

Agreement parameters

Agreement parameters measure the ability to achieve the same value in two measurements and give an indication of the size of measurement errors [14]. The agreement between two readers was evaluated using Bland and Altman plots [15]. The Bland and Altman plot is the most robust method to quantify agreement in clinical measurements. Here, the differences between measurements (on the Y-axis) are plotted against the mean of two measurements (on the X-axis). The visual representation of agreement illustrates the magnitude and range of the differences, bias or outliers, and the relation between the magnitude of the differences and the magnitude of the mean values [15]. This method also assesses 95% limits of agreement. These limits of agreement are used to define the smallest detectable difference (SDD) as $1.96 \times$ the standard deviation (SD) of the mean difference in measurements between the two readers. SDD represents the change that can be detected beyond measurement error [15]. We can define a tumor to have grown when the difference between two measurements falls outside this interval. In this way, it will be possible to discriminate between stable tumors and growing tumors, according to our measurements. For both two-dimensional and volume measurements, the SDD for absolute differences and for differences relative to baseline [SDD (%)] was calculated. Differences relative to baseline were calculated using the following formula: $(A-B)/[(A+B)/2] \times 100$, in which *A* is the result of reader A and *B* the result of reader B. The SDD and SDD (%) were presented for four VS stages: A, B, C, and D.

Use of the SDD (%) enables comparison of the different unities involved in the two different measurement techniques (millimeter and cubic centimeter). Because, in clinical practice, all three diameter measurements (CC, AP, and ML) are equally essential in assessing VS progression, we considered the diameter with lowest agreement as the limiting factor in diameter measurements. We compared the SDD (%) of this diameter with the SDD (%) of volume measurements.

Reliability parameters

Reliability parameters assess whether measurements can be used to distinguish patients from each other despite measurement error [14]. A parameter of reliability is the intraclass correlation coefficient (ICC). The ICCs are defined as the ratio of the variance among patients (patients variability) over the total variance (among patients, among readers, plus the error variance), which is expressed as a dimensionless number, being one (perfect reliability) in the most ideal case. ICC was calculated for interobserver diameter and volume VS measurement and was also presented for four VS stages: A, B, C, and D.

Results

Bland and Altman plots were constructed using data of the baseline MR images of the 165 scans from 68 patients (Figs. 3 and 4). The SD for each reader (A and B) and the SD of the mean difference between readers are presented in Table 1. The SDD and SDD (%) for absolute differences and differences relative to baseline MR images, respectively, are presented in Table 2. In Table 3, we present the ICC with 95% confidence intervals.

Contrast-enhanced T1- versus T2-weighted images

Two-dimensional and volume assessments of VS were performed on both CE T1-WI (CC dimension, T1CC; AP

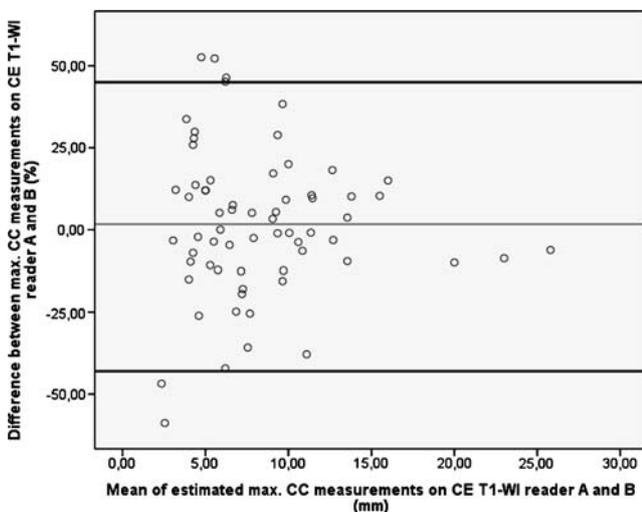


Fig. 3 Bland and Altman plot of baseline two-dimensional maximum craniocaudal (CC) dimension measurements on contrast-enhanced T1-weighted images (CE T1-WI). The values on the Y-axis represent the measurement differences between the two readers and their mean difference (*thin line*). The values on the X-axis represent the mean of both measurements. The *thick black lines* represent the 95% limits of agreement. Interobserver differences are larger in smaller vestibular schwannomas

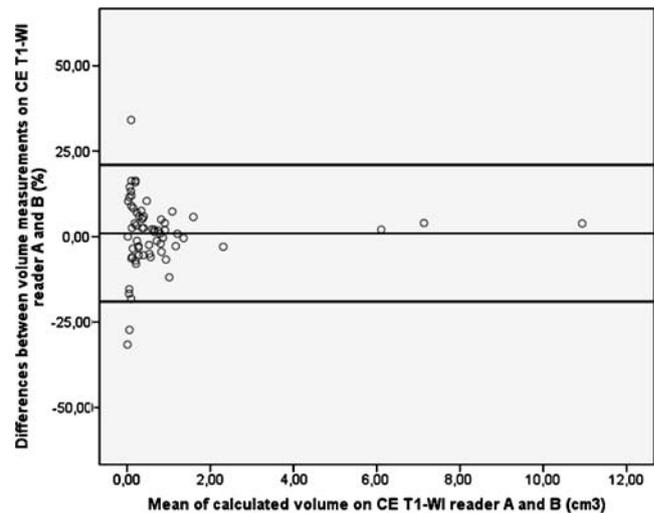


Fig. 4 Bland and Altman plot of baseline volume measurements on contrast-enhanced T1-weighted images (CE T1-WI). The values on the Y-axis represent the measurement differences between the two readers and their mean difference (*thin line*). The values on the X-axis represent the mean of both measurements. The *thick black lines* represent the 95% limits of agreement. Interobserver differences are larger in smaller vestibular schwannomas

dimension, T1AP; ML dimension, T1ML; volume, T1VOL) and T2-WI (CC dimension, T2CC; AP dimension, T2AP; ML dimension, T2ML; volume, T2VOL). The dimension with highest SDD (%) was taken as a limiting factor, when comparing both imaging modalities.

With two dimensional measurements, the SDD (%) for T1CC appeared to be equal to T2CC (40.3 and 40.1). However, the T2AP dimension showed higher SDD (%) compared to T1AP: 34.3 versus 28.3, respectively. Therefore, CE T1-WI showed highest agreement in two dimensional measurements. In addition, the ICC was consistently higher in T1CC, T1AP, and T1ML directions, compared to T2CC, T2AP, and T2ML directions (0.947, 0.974, and 0.978 versus 0.943, 0.961, and 0.948), reflecting higher reliability for CE T1-WI in two-dimensional measurements. For volume measurements, similar results were obtained: SDD (%) for T1VOL was 19.7 compared to 30.1 in T2VOL. ICC values for volume measurements for both sequences were 0.999.

Volume measurements versus two-dimensional measurements

The SDD (%) values for T1CC, T1AP, and T1ML were 40.3, 28.3, and 20.9, respectively (Table 2). All three dimensions are equally essential in estimating VS growth. Because the T1CC dimension is the limiting factor in these two-dimensional measurements, with its lowest agreement, it was used to compare the two-dimensional measurements with volume measurements. The SDD (%)

Table 1 Interobserver agreement parameters based on baseline T1-weighted, contrast-enhanced images (CE T1-WI) and T2-weighted images in 68 patients.

Baseline	Two-dimensional measurements on CE T1-WI (mm)			Volume measurements on CE T1-WI (cm ³)
	CC	AP	ML	VOL
Reader A, mean (SD)	8.40 (4.60)	8.70 (4.90)	12.70 (5.70)	0.85 (1.74)
A (n=29)	5.00	5.06	8.32	0.13
B (n=27)	9.05	9.71	14.28	0.63
C (n=6)	12.58	12.23	18.12	1.29
D (n=6)	17.50	18.33	21.62	4.71
Reader B, mean (SD)	8.20 (4.80)	8.60 (5.10)	13.20 (5.70)	0.81 (1.69)
A (n=29)	4.59	5.03	8.58	0.13
B (n=27)	9.27	9.86	15.06	0.63
C (n=6)	11.53	13.10	18.28	1.06
D (n=6)	18.33	18.93	22.28	4.57
Mean _{diff} (SD _{diff})	0.12 (1.53)	-0.18 (1.14)	-0.48 (1.08)	0.01 (0.055)

Baseline	Two-dimensional measurements on T2-WI (mm)			Volume measurements on T2-WI (cm ³)
	CC	AP	ML	VOL
Reader A, mean (SD) (n=68)	7.60 (4.5)	8.20 (5.00)	11.90 (5.50)	0.72 (1.49)
Reader B, mean (SD) (n=68)	7.20 (4.00)	8.40 (4.90)	12.10 (5.00)	0.72 (1.50)
Mean _{diff} (SD _{diff})	0.35 (1.41)	-0.18 (1.37)	-0.21 (1.69)	0.07 (0.069)

The rows of reader A and B of each imaging modality present the means and standard deviations of the measurements assessed by these readers. The means and standard deviations are also subdivided according to vestibular schwannoma stage for CE T1-WI. The mean difference row presents the mean differences between the measurements of reader A and B and the SD of these differences

A intracanalicular, B extracanalicular, C compressing brain stem, D deviating fourth ventricle, SD standard deviation, CC craniocaudal dimension, AP anteroposterior dimension, ML mediolateral dimension, VOL volume, Mean_{diff} mean of the observer difference, SD_{diff} standard deviation of the differences between observer A and B, CE T1-WI contrast enhanced T1-weighted images, T2-WI T2-weighted images

for T1CC was considerably higher than the SDD (%) for volume measurements (T1VOL): 40.3% versus 19.7%. The SDD (%) decreased with increasing tumor size from stage A to type D VS, in both two-dimensional and in volume measurements (Table 2). In all tumor stages, volume measurements were associated with smaller SDD (%) compared to CC dimensional measurements. The ICC

revealed that both two-dimensional and volume measurements showed high interobserver reliability. However, volume measurements revealed a higher reliability compared to the three diameter measurements (0.999 versus 0.947, 0.974 and 0.978). Reliability increased with tumor size, and volume measurements were more reliable in all VS stages (Table 3).

Table 2 SDD and SDD (%) for measurements performed on T1-weighted, contrast-enhanced (CE T1-WI) images and T2-weighted images.

	CC SDD (mm)	CC SDD (%)	AP SDD (mm)	AP SDD (%)	ML SDD (mm)	ML SDD (%)	VOL SDD (cm ³)	VOL SDD (%)
T1 A–D (n=68)	2.98	40.3	2.23	28.3	2.12	20.9	0.11	19.7
T2 A–D (n=68)	2.76	40.1	2.69	34.3	2.69	30.52	0.14	30.1
T1 A (n=29)	2.41	52.8	1.49	28.0	1.96	27.1	0.05	28.9
T1 B (n=27)	3.23	34.5	2.53	26.5	2.33	17.8	0.06	10.4
T1 C (n=6)	3.60	32.7	3.05	28.7	2.49	14.9	0.07	8.70
T1 D (n=6)	2.96	18.8	2.90	23.7	1.08	5.00	0.26	5.70

The SDD and SDD(%) for CE T1-WI are subdivided according to vestibular schwannoma (VS) stage: A intracanalicular VS, B VS extending into the cerebellopontine angle, C VS compressing brain stem, D VS deviating the fourth ventricle

SDD smallest detectable difference, SDD (%) smallest detectable difference relative to baseline measurements, CC craniocaudal dimension, AP anteroposterior dimension, ML mediolateral dimension, VOL volume

Table 3 Intraclass correlation coefficients (ICC) for two-dimensional and volume measurements on T1-weighted, contrast-enhanced images (CE T1-WI) and T2-weighted images with their 95% confidence intervals.

	CC	AP	ML	VOL
T1 A–D	0.947 (0.915–0.967)	0.974 (0.957–0.984)	0.978 (0.965–0.987)	0.999 (0.999–1.000)
T2 A–D	0.943 (0.909–0.965)	0.961 (0.938–0.976)	0.948 (0.917–0.968)	0.999 (0.998–0.999)
T1A (<i>n</i> =29)	0.564 (0.260–0.768)	0.862 (0.729–0.932)	0.947 (0.892–0.975)	0.987 (0.973–0.994)
T1B (<i>n</i> =27)	0.674 (0.393–0.841)	0.802 (0.604–0.907)	0.889 (0.767–0.949)	0.988 (0.973–0.994)
T1C (<i>n</i> =6)	0.840 (0.315–0.975)	0.917 (0.588–0.988)	0.968 (0.822–0.995)	0.999 (0.993–1.000)
T1D (<i>n</i> =6)	0.966 (0.810–0.995)	0.980 (0.883–0.997)	0.990 (0.942–0.999)	0.999 (0.992–1.000)

ICC values for CE T1-WI are subdivided per VS stage: *A* intracanalicular VS, *B* VS extending into the cerebellopontine angle, *C* VS compressing brain stem, *D* VS deviating the fourth ventricle

CC craniocaudal dimension, AP anteroposterior dimension, ML mediolateral dimension, VOL volume

Smallest detectable differences in two-dimensional and volume measurements

SDD for two dimensional measurements on CE T1-WI varied from 2.12 to 2.98 mm. For volume measurements on CE T1-WI, SDD was 0.11 cm³ (Table 2).

Discussion

VS are benign neoplasms, originating from the neurolemmal sheath of the eighth cranial nerve and are predominantly found in the CPA and in the internal auditory canal (IAC). The incidence of VS varies from 1 to 2/100,000 [2, 12, 16, 17], although postmortem histopathological examinations show a higher incidence of about 2.7% [18]. This discrepancy indicates that the vast majority of VS never become symptomatic, reflecting very slow or arrested growth.

Therefore, the wait and scan policy has gained popularity as an alternative or prelude to surgery and radiation therapy. This can be justified, as growth is known to be extremely variable with most VS remaining stable or showing minimal growth for many years [16, 19]. The goal of this regimen is to minimize therapeutic risks and complications and to preserve an optimal quality of life in selected patients. Because no single reliable clinical feature exists that predicts tumor growth [5–7], MRI is the mainstay in the conservative management of VS [2, 5]. It is essential to use a measuring method that provides reliable measurements with a high interobserver agreement, as change in size is—besides its clinical presentation—the most relevant parameter.

Various ways of describing VS tumor growth have been proposed. The conventional method of assessing VS is by performing two-dimensional measurements. Some authors use the largest AP and/or ML dimension, sometimes combined with the CC dimension [5, 8, 19–24]. Others use

the guidelines for measuring VS described by the American Academy of Otolaryngology, Head and Neck Surgery (AAO-HNS) [6, 7, 25]. The usefulness of two-dimensional measurements could be doubted. Firstly, two-dimensional measurements assesses VS growth in maximal three directions, while a VS shows asymmetric growth in all possible directions; therefore, a two-dimensional assessment does not approach “real” tumor growth. Secondly, in VS with large diameters, a small increase in diameter corresponds to a much larger increase in VS volume than a similar increase in diameter in a small VS [1]. Volume measurements can be performed in several ways; some authors consider VS to be ellipsoid and calculate the volume using a mathematical formula [11, 21, 26, 27]. However, this method has shown to produce a large overestimation of VS volume [1, 11]. Others have performed true volumetric measurements by using (semi)automatic software to calculate VS volume [1, 9, 11, 12, 21, 28, 29]. According to the few studies comparing two-dimensional versus non-formula-based volume measurements, the VS volume measurements are more accurate compared to two-dimensional measurements [1, 9, 10]. Other authors disagree with this [16, 21, 27], and also in clinical practice, most clinicians keep relying on two-dimensional measurements. However, the results of this study indicate that VS volume measurements, especially on the CE T1-WI, produce a better interobserver agreement and reliability compared to the two-dimensional measurements. This study therefore indicates that CE T1-WI volume measurements should replace two-dimensional measurements in evaluating VS growth. The difference in interobserver agreement and reliability between the two measurement methods is of clinical significance because invasive treatment decisions are based on these observations. Therefore, the measurement method with highest agreement on reliability is necessary in assessing VS growth.

An exception may be made concerning small (stage A) VS. In these small intracanalicular VS, the CC dimension does not play an important role, since the diameter of the

IAC is usually quite constant. When one uses the AAO-HNS guidelines, the CC direction is not taken into account at all. Then, the AP direction is the limiting factor of the remaining AP and ML directions (Table 2). In stage A VS, under these conditions, we found that measurement error [SDD(%)] of both two-dimensional and volume measurement techniques are comparable (AP 28.0, ML 27.1, and VOL 28.9). Therefore, both measurement techniques can be used to evaluate stage A VS. In larger VS (stages B and C), both AP and ML dimensions show higher measurement error [SDD (%)] compared to volume measurements. In stage D VS, the SDD (%) of the ML dimension equals the SDD (%) of volume measurements: 5 versus 5.7, respectively. This can be explained by the fact that the ML dimension is the longest distance in two-dimensional VS measurements, and measurement error will decrease when measurement distances increase. This occurs both in two-dimensional and in volume measurements and should be taken into account when evaluating growth in different VS stages (Table 2 and Figs. 3 and 4). Other authors confirm this phenomenon [1, 11, 30]. However, apart from this low SDD (%) of the ML dimension in stage D VS, the far worse SDD (%) of the AP dimension should also be taken into account in evaluating VS growth with two-dimensional measurements, thus pleading for the use of volume measurements in the assessment of stage B, C, and D VS.

Overall, contrast-enhanced images are necessary, both in volume and two-dimensional measurements, to maintain a high reproducibility, since it facilitates differentiation of VS from surrounding tissue.

The calculated agreement [SDD and SDD (%)] and reliability (ICC) values were compared with findings reported in the literature. Studies addressing the change in length and volume exceeding measurement error (SDD) are reported inconclusively. In two-dimensional VS measurements, many authors simply state a 1- or 2-mm increase in subsequent scans as evidence of growth without any statistical justification [2, 6–8, 20, 22, 31]. According to this present study, SDD varies between 2.12 mm for the ML dimension and 2.98 mm for the CC dimension on CE T1-WI (Table 2). This indicates that there is a measurement error in two-dimensional measurements, which is not recognized when a measured increase in size of only 1 or 2 mm between follow-up scans is considered as tumor growth, and treatment decisions should be used with caution when using these arbitrary criteria.

In volume measurements, an absolute increase above which one can consider a VS to have grown was not found in the literature. This study revealed SDD ranging from 0.05 to 0.26 cm³ (types A–D) on CE T1-WI (Table 2). Volume increase expressed as SDD (%) varied from 15 to 89% in the literature [9–12]. However, the numbers of patients in these studies were small [9, 11, 12], and

generally, not only baseline measurements of the VS were taken into account [9, 10, 12]. Therefore, the percentages reported in the literature could be questioned. In this study, SDD (%) in CE T1-WI volume measurements was 19.7%. ICC for two-dimensional measurements was not found in the literature. One study calculated ICC for volume measurements: Luppino et al. [9] calculated ICC for two different types of volume measurement. Their “contour method,” similar to our volume method, had an interobserver reliability of 0.96, which is comparable to the ICC of 0.99 in this study.

This study and technique also harbor some limitations. Firstly, the assessments of reproducibility parameters were based on interobserver differences and not on intraobserver differences. It was assumed that this approach better reflects clinical practice, where it is usual that different clinicians assess subsequent scans.

Secondly, BrainLAB volume software is not widely available in radiology departments

Thirdly, performing these volume measurements takes a little more time compared to the conventional two-dimensional measurements. In our experience, VS contouring and volume calculation typically took a few minutes. This could be a limiting factor in introducing this method for VS in daily clinical practice. In the literature, 10–25 min for manual segmentation has been described, although these calculations were performed with different software and on older systems [10, 11].

Fourthly, the used volume measurement method is semi-automatic. However, there is still a human component that is responsible for an interobserver difference. It is desirable to develop software able to perform even better automated volume measurements in order to further diminish reader-related measurement error.

Conclusion

CE T1-WI volume measurements show better interobserver agreement and reliability compared to two-dimensional measurements for the assessment of growth of VS. Small intracanalicular VS form an exception. When evaluating VS growth, one has to take VS baseline characteristics into account because SDD (%) strongly depends on VS size. The 1- or 2-mm difference commonly used to define growth of VS in consecutive scans in two-dimensional measurements lies within measurement error and should not direct clinical practice.

Conflict of interest statement We declare that we have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Shin YJ, Fraysse B, Cognard C, Gafsi I, Charlet JP, Berges C, Deguine O, Tremoulet M (2000) Effectiveness of conservative management of acoustic neuromas. *Am J Otol* 21:857–862
- Stangerup SE, Caye-Thomasen P, Tos M, Thomsen J (2006) The natural history of vestibular schwannoma. *Otol Neurotol* 27:547–552. doi:10.1097/00129492-200606000-00018
- Stangerup SE, Caye-Thomasen P, Tos M, Thomsen J (2008) Change in hearing during 'wait and scan' management of patients with vestibular schwannoma. *J Laryngol Otol* 122:673–681. doi:10.1017/S0022215107001077
- Charabi S, Tos M, Thomsen J, Charabi B, Mantoni M (2000) Vestibular schwannoma growth: the continuing controversy. *Laryngoscope* 110:1720–1725. doi:10.1097/00005537-200010000-00030
- Slattery WH 3rd, Fisher LM, Yoon G, Sorensen G, Lev M (2003) Magnetic resonance imaging scanner reliability for measuring changes in vestibular schwannoma size. *Otol Neurotol* 24:666–670. doi:10.1097/00129492-200307000-00022 discussion 670–661
- Walsh RM, Bath AP, Bance ML, Keller A, Tator CH, Rutka JA (2000) The role of conservative management of vestibular schwannomas. *Clin Otolaryngol Allied Sci* 25:28–39. doi:10.1046/j.1365-2273.2000.00317.x
- Raut VV, Walsh RM, Bath AP, Bance ML, Guha A, Tator CH, Rutka JA (2004) Conservative management of vestibular schwannomas—second review of a prospective longitudinal study. *Clin Otolaryngol Allied Sci* 29:505–514. doi:10.1111/j.1365-2273.2004.00852.x
- Smouha EE, Yoo M, Mohr K, Davis RP (2005) Conservative management of acoustic neuroma: a meta-analysis and proposed treatment algorithm. *Laryngoscope* 115:450–454. doi:10.1097/00005537-200503000-00011
- Luppino FS, Grooters E, de Bruine FT, Zwinderman AH, van der Mey AG (2006) Volumetric measurements in vestibular schwannoma, the influence of slice thickness and patient's repositioning. *Otol Neurotol* 27:962–968. doi:10.1097/01.mao.0000235371.39998.a7
- Vokurka EA, Herwadkar A, Thacker NA, Ramsden RT, Jackson A (2002) Using Bayesian tissue classification to improve the accuracy of vestibular schwannoma volume and growth measurement. *AJNR Am J Neuroradiol* 23:459–467
- Cross JJ, Baguley DM, Antoun NM, Moffat DA, Prevost AT (2006) Reproducibility of volume measurements of vestibular schwannomas—a preliminary study. *Clin Otolaryngol* 31:123–129. doi:10.1111/j.1749-4486.2006.01161.x
- Niemczyk K, Vaneecloo FM, Lemaitre L, Lejeune JP, Skarzynski H, Dubrulle F, Vincent C (1999) The growth of acoustic neuromas in volumetric radiologic assessment. *Am J Otol* 20:244–248
- Hasegawa T, Fujitani S, Katsumata S, Kida Y, Yoshimoto M, Koike J (2005) Stereotactic radiosurgery for vestibular schwannomas: analysis of 317 patients followed more than 5 years. *Neurosurgery* 57:257–265. doi:10.1227/01.NEU.0000166542.00512.84 discussion 257–265
- de Vet HC, Terwee CB, Knol DL, Bouter LM (2006) When to use agreement versus reliability measures. *J Clin Epidemiol* 59:1033–1039. doi:10.1016/j.jclinepi.2005.10.015
- Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1:307–310
- Rosenberg SI (2000) Natural history of acoustic neuromas. *Laryngoscope* 110:497–508. doi:10.1097/00005537-200004000-00002
- Tos M, Charabi S, Thomsen J (1998) Clinical experience with vestibular schwannomas: epidemiology, symptomatology, diagnosis, and surgical results. *Eur Arch Otorhinolaryngol* 255:1–6. doi:10.1007/s004050050012
- Yoshimoto Y (2005) Systematic review of the natural history of vestibular schwannoma. *J Neurosurg* 103:59–63
- Selesnick SH, Johnson G (1998) Radiologic surveillance of acoustic neuromas. *Am J Otol* 19:846–849
- Doherty JK, Friedman RA (2006) Controversies in building a management algorithm for vestibular schwannomas. *Curr Opin Otolaryngol Head Neck Surg* 14:305–313. doi:10.1097/01.moo.0000244186.72645.d4
- Fiirgaard B, Pedersen CB, Lundorf E (1997) The size of acoustic neuromas: CT and MRI. *Neuroradiology* 39:599–601. doi:10.1007/s002340050475
- Fucci MJ, Buchman CA, Brackmann DE, Berliner KI (1999) Acoustic tumor growth: implications for treatment choices. *Am J Otol* 20:495–499
- Tanaka Y, Hongo K, Tada T, Kobayashi S (2003) What is the best method for reporting tumor diameter in vestibular schwannoma? *Neurosurgery* 53:634–637. doi:10.1227/01.NEU.0000080062.61335.A5 discussion 637–638
- Saito AI, Morris CG, Ito K, Watanabe F, Karasawa K, Mendenhall WM, Naoi Y (2007) Comparing size evaluation methods for acoustic neuroma after stereotactic radiosurgery. *Radiat Med* 25:339–345. doi:10.1007/s11604-007-0150-6
- American Academy of Otolaryngology—Head and Neck Surgery Foundation (1995) Committee on Hearing and Equilibrium guidelines for the evaluation of hearing preservation in acoustic neuroma (vestibular schwannoma). *Otolaryngol Head Neck Surg* 113:179–180. doi:10.1016/S0194-5998(95)70101-X
- Caye-Thomasen P, Hansen S, Dethloff T, Stangerup SE, Thomsen J (2006) Sublocalization and volumetric growth pattern of intracanalicular vestibular schwannomas. *Laryngoscope* 116:1131–1135. doi:10.1097/01.MLG.0000217528.37106.2D
- Charabi S, Thomsen J, Mantoni M, Charabi B, Jorgensen B, Borgesen SE, Gyldensted C, Tos M (1995) Acoustic neuroma (vestibular schwannoma): growth and surgical and nonsurgical consequences of the wait-and-see policy. *Otolaryngol Head Neck Surg* 113:5–14. doi:10.1016/S0194-5998(95)70138-9
- Laasonen EM, Troupp H (1986) Volume growth rate of acoustic neurinomas. *Neuroradiology* 28:203–207. doi:10.1007/BF00548193
- Meijer OW, Weijmans EJ, Knol DL, Slotman BJ, Barkhof F, Vandertop WP, Castelijns JA (2008) Tumor-volume changes after radiosurgery for vestibular schwannoma: implications for follow-up MR imaging protocol. *AJNR Am J Neuroradiol* 29:906–910. doi:10.3174/ajnr.A0969
- Sarazin L, Jolivet O, Doyon D (1993) Computerized evaluation of tumor volume with MRI. Applications to the surveillance of acoustic neurinoma. *J Radiol* 74:455–460
- Bozorg Grayeli A, Kalamirides M, Ferrary E, Bouccara D, El Gharem H, Rey A, Sterkers O (2005) Conservative management versus surgery for small vestibular schwannomas. *Acta Otolaryngol* 125:1063–1068. doi:10.1080/00016480510038013