Suggested response criteria for phase II antitumor drug studies for neurofibromatosis type 2 related vestibular schwannoma

Scott R. Plotkin,
Department of Neurology and Cancer Center, Massachusetts General Hospital, Boston, MA, USA

Chris Halpin,
Audiology Department, Massachusetts Eye and Ear Infirmary, Boston, MA, USA

Jaishri O. Blakeley,
Department of Neurology, Johns Hopkins Medical Institute, Baltimore, MD, USA

William H. Slattery III,
Department of Otolaryngology, House Ear Institute, Los Angeles, CA, USA

D. Bradley Welling,
Department of Otolaryngology, Ohio State University, Columbus, OH, USA

Susan M. Chang,
Department of Neurological Surgery, University of California, San Francisco, CA, USA

Jay S. Loeffler,
Department of Radiation Oncology, Massachusetts General Hospital, Boston, MA, USA

Gordon J. Harris,
Department of Radiology, Massachusetts General Hospital, Boston, MA, USA

A. Gregory Sorensen,
Department of Radiology, Massachusetts General Hospital, Boston, MA, USA

Michael J. McKenna, and
Department of Otolaryngology, Massachusetts Eye and Ear Infirmary, Boston, MA, USA

Fred G. Barker II
Neurosurgical Service, Pappas Center for Neuro-Oncology, Yawkey 9E, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA

Fred G. Barker: barker@helix.mgh.harvard.edu

Abstract

Neurofibromatosis type 2 (NF2) is a tumor suppressor gene syndrome characterized by multiple schwannomas, especially vestibular schwannomas (VS), and meningiomas. Anticancer drug trials are now being explored, but there are no standardized endpoints in NF2. We review the challenges of NF2 clinical trials and suggest possible response criteria for use in initial phase II studies. We

Correspondence to: Fred G. Barker, II, barker@helix.mgh.harvard.edu.

suggest two main response criteria in such trials. *Objective radiographic response* is defined as a durable 20% or greater reduction in VS volume based on postcontrast T1-weighted MRI images collected with 3 mm or finer cuts through the internal auditory canal. *Hearing response* is defined as a statistically significant improvement in word recognition scores using 50-word recorded lists in audiology. A possible composite endpoint incorporating both radiographic response and hearing response is outlined. We emphasize pitfalls in response assessment and suggest guidelines to minimize misinterpretations of response. We also identify research goals in NF2 to facilitate future trial conduct, such as identifying the expectations for time to tumor progression and time to measurable hearing loss in untreated NF2-related VS, and the relation of both endpoints to patient prognostic factors (such as age, baseline tumor volume, and measures of disease severity). These data would facilitate future use of endpoints based on stability of tumor size and hearing, which might be more appropriate for testing certain drugs. We encourage adoption of standardized endpoints early in the development of phase II trials for this population to facilitate comparison of results across trials of different agents.

Keywords

Vestibular schwannoma; Acoustic neuroma; Neurofibromatosis type 2; Trial design; Phase II trial

Background

Vestibular schwannomas and neurofibromatosis type 2

Vestibular schwannomas (VS; also known as acoustic neuromas) are among the most common benign central nervous system tumors, with an annual incidence in the United States of approximately 1.5 per 100,000 person years and comprising about 9% of all brain tumors [1]. No medical treatment for VS currently exists. Patients are managed either through observation with serial imaging studies or, when active treatment is required, via either surgical resection or radiation [2]. Vestibular schwannomas commonly cause unilateral hearing loss, trigeminal deficit, and (usually after treatment) facial weakness. However, in most patients with sporadic, unilateral VS good functional status and quality of life are accomplished with observation, radiation treatment, surgery, or some combination of these; many are cured by treatment.

In neurofibromatosis type 2 (NF2), a dominantly inherited syndrome caused by alterations in the *NF2* tumor suppressor gene, VS occur bilaterally, commonly accompanied by other cranial and spinal nerve schwannomas and meningiomas. NF2 affects about one person per 25,000 live births, and accounts for 2–7% of patients with VS [3–5]. In contrast to sporadic VS, in NF2 patients, the inexorable growth of VS and other tumors and cumulative morbidity from treatment typically cause multiple cranial nerve deficits, often followed by death in young adulthood [6]. In one registry study, 60% of patients died before age 44 [7]. Nearly all NF2 patients die from sequelae of tumor growth or from treatment mortality [8]. Unlike the unilateral cranial nerve defects encountered in sporadic VS patients, the cranial nerve deficits in NF2 patients are often bilateral, including complete hearing loss, facial diplegia causing dysarthria and difficulty eating, and inability to swallow or handle secretions because of bilateral lower cranial nerve dysfunction. These can have profound

effects on quality of life and may contribute to mortality. In addition, the morbidity of treatment from either surgery [9] or radiation [10–12] is greater than for sporadic VS, probably because nerve fibers are much more likely to run through NF2 associated schwannomas rather than on the tumor surface, as in sporadic VS [13–16].

Treatment choices in NF2 often weigh the possible consequences of progressive brainstem compression and gradual cranial nerve dysfunction from multiple enlarging tumors against the potential for devastating combinations of cranial nerve deficits from surgical or radiation treatment (Fig. 1). A particularly common clinical scenario in NF2 is progressive loss of hearing from VS growth in an only hearing ear, after the other ear has already suffered complete hearing loss due to tumor growth or treatment morbidity. For patients such as these, in contrast to patients with sporadic, unilateral VS, both surgery and radiation treatment may pose unacceptable risks, and novel medical treatments are urgently needed.

Need for standardized response criteria in NF2

Although radiation therapy has emerged as an effective treatment for VS, the lack of standardized response criteria complicates attempts to compare results in separate clinical series, e.g., with different fractionation schemes for fractionated stereotactic radiotherapy [17–19]. Currently there are no established criteria specifically designed for assessing responses to antitumor drugs in benign nervous system tumors such as schwannomas and meningiomas. More recently, the introduction of targeted therapies developed for treatment of common cancers, combined with recent work on the tumor biology of schwannomas, has raised the possibility of designing rational drug treatments for NF2 patients with progressive intracranial tumors [20–26]. For example, the treatment of small series of NF2 patients using erlotinib [27, 28] and bevacizumab [29] has recently been reported, with apparent benefit in some cases—including stabilization and shrinkage of some VS and/or substantial reversal of chronic hearing loss. This initial experience suggests that a range of responses (and nonresponses) are possible after drug treatment for benign nervous system tumors.

In this communication, we suggest response criteria for measuring the success of antitumor drugs used to treat NF2-related VS. Specifically, we highlight the suggested roles of volumetric tumor measurement and word recognition hearing scores in assessing both direct antitumor effects and an important component of health-related quality of life in NF2 patients, especially those who are losing the hearing in an only hearing ear. Although other potential endpoints in such trials are of great interest, and may have a place as secondary endpoints in trials, we argue here, based on our initial experience with small numbers of patients as well as published studies on NF2 and VS, that a standardized assessment of these two endpoints should be included routinely in trials of antitumor drugs in NF2. We do not here systematically address certain other important aspects of phase II trial design, such as entry criteria (including definitions of NF2 and of VS in the absence of histological diagnosis), choice of trial design (e.g., singlearm, randomized with active or placebo comparator arm [30, 31], crossover [32], or randomized discontinuation [30, 33]), measuring responses in non-VS lesions such as meningiomas or other schwannomas, toxicities that should be monitored with special significance in NF2 such as ototoxicity [34] and corneal keratopathy (to which patients with trigeminal lesions may be especially susceptible) [35],

or the "go/no go" choice of what level of response justifies proceeding to phase III trials for a specific agent [36].

Justification of endpoint selection: tumor size

Typical endpoints used in phase II cancer trials

Phase II trials are trials designed to detect the first evidence of an anticancer drug's activity against tumors in patients. (Phase I trials test the safety of drugs that have not previously seen human use, usually in a small single-arm design, and phase III trials aim to prove drug efficacy, usually in a large randomized trial.) A typical phase II design using tumor response as an endpoint enrolls 40 patients who are treated with a novel drug; if the drug produces a prespecified number of "responses", the drug may proceed to definitive phase III testing.

Historically, most phase II trials of anticancer drugs have relied on objective tumor response rates (i.e. tumor shrinkage, by imaging or by direct measurement, in patients whose tumors are shown to be enlarging) in justifying the selection of active agents for further testing in definitive phase III trials. The assumptions behind this choice of endpoint are that most anticancer drugs work through a cytotoxic mechanism, that more cytotoxicity is reflected by greater macroscopic tumor shrinkage, and that objective tumor shrinkage is an adequate surrogate for other endpoints—such as longer overall or progression-free survival—that usually serve as primary endpoints in definitive phase III trials [37–40].

Recent developments in phase II drug testing in some other tumors have emphasized alternative imaging endpoints instead of or in addition to objective reduction of tumor size measurements, for several reasons. Many targeted agents are assumed to have cytostatic, rather than cytotoxic effects, and lack of tumor shrinkage may not reflect lack of drug activity. Conversely, some tumor enlargement by imaging may represent "pseudoprogression"— an increase in enhancement due to treatment that does not reflect actual tumor growth [41, 42]—and spontaneous regression of these changes can falsely mimic active drug response. Under either of these circumstances, freedom from disease progression over time may be a more appropriate marker of drug activity than objective response. The increasingly common application of progression- free survival at 6 months of treatment as the primary endpoint in recurrent malignant glioma drug trials reflects these concerns [43–45]. Even given the difficulties in accurately defining tumor progression cited above, progression-free survival deserves consideration as a primary endpoint in NF2-related tumors, particularly since these tumors are histologically benign, typically slowgrowing, and metastasis is exceptionally rare.

Additionally, imaging changes other than changes in size are sometimes more sensitive early indicators of anticancer drug activity. Examples are reduction in tumor density on contrast-enhanced computed tomographic (CT) scans [46, 47], reductions in tumor metabolism detected using positron emission tomography [46–48], and changes in diffusion or perfusion within tumors detected using MRI [49–51]. To date, there is little clinical experience with these imaging modalities in either sporadic or NF2-related VS.

Special considerations for NF2 trial endpoints

For growing VS in NF2 patients, clinical experience suggests that a durable stabilization or reduction in tumor size from treatment would be a direct clinical benefit. NF2 patients suffer progressive declines in hearing over time, and these declines are associated with gradual VS growth [52]. In one retrospective study, 27% of 108 ears suffered decreased hearing within 2 years of diagnosis of NF2 [53]. Although hearing can decline in ears during an interval of time without measurable VS growth, sometimes suddenly, hearing decline in VS is significantly more common in actively growing tumors [54–57]. In addition, the morbidity of surgery or radiation treatment of VS is highly dependent on tumor size [9, 10, 58]. Other symptoms, such as trigeminal deficits, are directly correlated with tumor size, and can resolve if mass effect on the nerve is reduced by surgery.

Response rate versus time to progression in NF2

We suggest reduction in tumor size as an endpoint for initial single-arm phase II drug trials in NF2-related VS, rather than time before tumor progression, because of our present imperfect understanding of NF2 natural history. Several studies have addressed the natural history of NF2, including the course of untreated VS [52, 53, 59–63]. Most of these studies reported growth rates of VS over various periods of time in individuals or in groups [52, 59, 60, 62, 64]; one study reported the fraction of untreated tumors that grew or were stable over an unspecified period [63]. None of these studies reported time until progression in individual untreated NF2 VS. One study reported time to tumor progression in 11 untreated VS in NF2 patients; median time to tumor growth was 6 months [61]. These tumors were under observation because they were contralateral to growing VS that were treated with radiosurgery, but whether the observed tumors were growing at the time observation started was not reported. In a similar report [65], 15 of 29 untreated VS contralateral to a VS undergoing radiosurgery required treatment at an average of 27 months after starting observation. In addition, in sporadic VS followed conservatively, initial growth rate is a good predictor of subsequent continued growth [57, 66], although sporadic VS appear to have a substantially lower growth rate than that reported for NF2 VS [67–69]. Although tumor growth is correlated with loss of hearing, reduction of tumor size after treatment with surgical resection or radiosurgery is only rarely accompanied by an improvement in hearing, even in sporadic VS [10, 70–76]. Whether this also holds true for drug treatment remains to be seen.

These data suggest that time to tumor progression in initially growing NF2-related VS is an endpoint that, without effective treatment, would be reached in many patients in a relatively short period of time. However, a much better understanding of the natural history of untreated NF2 VS, including potentially important prognostic factors for growth such as age, the presence of meningiomas, other measures of disease severity such as number of prior surgical procedures or overall tumor burden, measured baseline tumor volume and growth rates, family history and familial pattern of disease, and specific NF2-defining genetic defects [59, 60, 62, 64] will be necessary before this endpoint can be used in single-arm phase II trials that use a historical control group. Progression-free survival could potentially be used at present in randomized phase II trials of NF2-related VS with an untreated control cohort, or placebo controlled crossover trials. Randomized phase II trials

that compare active treatment to a placebo control arm are considered to be most useful when there is little prior information on the expected efficacy rate of a novel treatment, or when the endpoint is thought to be dependent on patient selection factors (such as time to tumor progression) [77]—both criteria that are applicable to NF2-VS studies. However, the present understanding of progression-free survival in NF2 related VS is sufficiently vague that it would be difficult to determine accrual goals and planned follow-up periods to properly power such trials, which also risk poor patient acceptance because of the use of a placebo arm. Better knowledge of the possible influence of patient-related prognostic factors noted above would probably also be necessary in properly stratifying the randomization and analysis of such trials. Crossover trials in which each patient alternates between active and placebo treatment periods would also be difficult to design in the absence of better natural history data.

Recommended use of radiographic response rate

We propose that *objective radiographic response rate* is a better choice for initial drug trials in NF2 VS. In contrast to time to progression, the expectation for rates of spontaneous tumor shrinkage in untreated VS is better understood. Spontaneous involution of untreated sporadic VS is well known; it occurs in about 8% of tumors in most reports [68, 78]. Tumors can reduce by 3–15 mm in diameter [79], usually during a long period of observation. Regression has occasionally been reported in VS that were initially growing [80, 81], although most reports do not specify initial growth. In the NF2 Natural History Consortium patients, 3 of 84 tumors (3.6%) regressed by 2 mm or more during 9 months to 2 years of followup [60]. This study was not limited to patients with initially growing tumors, in which spontaneous regression would probably be expected to be even less frequent. This low rate of spontaneous regression in NF2 related VS suggests that most objective responses in a drug trial for these tumors would likely represent true activity of the agent being tested.

Methods of measuring tumor size

Tumor response has historically been measured in relation to linear dimension, as in the 2000 RECIST criteria [82]; bidimensional area, as in the 1979 WHO criteria [83], and the 1990 "Macdonald" criteria for malignant glioma, which also incorporated steroid dose and neurological condition [84]; and volumetric measures based on crosssectional imaging such as CT or MRI [51, 85]. For sporadic VS, measures based on greatest linear dimension are typically used in clinical practice to assess growth over time [54, 57] and operative risk of complications [58]. In following VS for growth in a "wait and scan" strategy, 2 mm of growth is commonly chosen as the threshold before recommending treatment [86, 87], or less often 1 mm [88]. Either the greatest tumor diameter measured parallel to the petrous face or simply the greatest tumor dimension (potentially including the intracanalicular portion), both on axial images, are common assessment methods [60, 89]. An average tumor diameter calculated as the square root of the product of the greatest dimension parallel to the petrous ridge and the greatest perpendicular diameter on axial images has been recommended as a standard measure of VS size [90]. While such measures may adequately summarize tumor size in sporadic VS, which typically have a regular, ovoid shape in their extracanalicular portion, NF2-related VS are often highly irregular in form (particularly after prior treatment) (Fig. 2), and such measures are likely to be inaccurate surrogates for true

NF2 VS volumes. A comparative study showed that volumetric measures were much more sensitive to tumor growth (and, presumably, to tumor shrinkage) than linear measures [89], and volume estimates based on bidimensional area have been shown to be a poor surrogate for actual tumor volumes [91].

For volumetric measure of a VS, a simple method is the "box model" in which the greatest anteroposterior, mediolateral, and superoinferior measures are multiplied [52, 59], or a two-box modification that measures intracanalicular and extracanalicular volumes separately [62]. Two limitations of the box method are the considerable interobserver variation in measuring the three tumor dimensions used in this method [87, 92] and the irregular shape of most NF2 VS, in contrast to the regular shape of sporadic VS. A method that uses summed areas from individual crosssectional imaging slices (typically axial slices) has been shown to yield accurate volumetric measurements, within 10% of true volumes when five or more axial slices through the target volume are used [93]. A study of volumetric measurement of NF2 VS using these methods, using 3 mm axial slices, showed an intrarater coefficient of variation of measurements of 5% or less for tumors larger than 1 cm3 [89]. 1 mm isotropic voxels are now routinely feasible on many MRI systems, and the smaller the voxel size, the smaller the VS size and change that can be accurately detected. We recommend volumetric measurement of VS volume as the means of assessing objective tumor response in NF2 VS trials.

Volumetric measurement of VS: special considerations in NF2

An additional issue in assessing tumor volume for trials is the number of lesions that can contribute to volume measurement. In cancer trials, the widely used RECIST criteria specify that up to five lesions per organ and up to 10 in total can contribute to tumor size measurement [82]. NF2 differs from malignant primary tumors in that, although multiple lesions are usually present, they do not represent metastases. In addition, because of their critical sites of origin, most individual NF2 lesions have the potential to cause distinct neurological symptoms through growth. We recommend that individual lesions be measured separately. Most trials will probably specify a single index or "target" lesion that qualifies a patient for trial entry and is tracked as the principal endpoint. For example, a growing VS on the side of an only hearing ear might be the target lesion, and the sizes of other schwannomas and meningiomas would then be followed as secondary endpoints.

Threshold for defining objective response

No universally agreed definition of the degree of tumor volume reduction necessary to declare response in phase II trials has emerged, although a 50% change in volume has been proposed for malignant glioma trials [51, 94]. For benign tumors such as VS, as well as for agents assumed to have a cytostatic rather than cytotoxic mechanism of action, a smaller change in volume may be more sensitive in detecting drug activity. Volumetric changes of 15% [95] or 20% [96] have been used to define response and progression in early phase trials in neurofibromatosis type 1 associated plexiform neurofibromas. We propose using a 20% change in comparison with baseline as the threshold to define response or progression in NF2 VS. This corresponds roughly to a 2 mm change in diameter of a 22 mm diameter

sphere, which would be the typical threshold for reliable clinical measurement of a minimal change in a VS of average size.

Response should be measured at standardized timepoints to facilitate comparison across trials. We suggest the inclusion of (at least) 6-, 12-, and (when available) 24-month assessments in trial reports, although investigators often would specify more frequent scanning in trial protocols. Responses must be durable, i.e., confirmation of radiographic response should be performed at least 2 months after the criteria for response are first met.

Justification of endpoint selection: hearing response

Hearing loss and quality of life in NF2 patients

Bilateral VS are the hallmark of NF2, and most NF2 patients eventually suffer complete bilateral hearing loss either from tumor growth or from the morbidity of surgical or radiation treatment. Clinical experience indicates that many, if not most NF2 patients lose all useful hearing before death. Because many patients have their VS completely excised after hearing is lost through tumor growth, prolonged survival with complete hearing loss as the dominant neurological deficit is possible if the burden of other NF2- related tumors (such as meningiomas and spinal tumors) is low; such patients are frequently encountered in clinical practice. Although data on the frequency of hearing loss in an NF2 population are not available, Baser et al. [64] reported that 37% of patients in the UK NF2 registry had undergone bilateral VS removal by age 50; most such patients would have no useful hearing, and many more who had not had bilateral surgery would have lost all hearing as well through tumor growth. Because most NF2 patients lose hearing after acquisition of language (i.e., they are postlingually deafened), the effect of hearing loss on their quality of life is typically severe, affecting psychological and social function, and employment [97]. Although no quality of life studies specific to NF2 patients have been performed, the adverse effect of severe and even moderate hearing loss on healthrelated quality of life (HRQoL) has great face validity and has been well documented in children and adults who lose hearing from other causes [98–102]. Conversely, treatments that improve hearing function in patients with moderate to severe hearing loss, such as hearing aids, cochlear implants, and ear surgery, cause improvements in HRQoL that are easily measured [103–110].

Non-hearing neurological consequences of VS

Other neurological consequences of VS deserve consideration as possible response measures in NF2 clinical trials. Balance difficulties and tinnitus are common in NF2 patients, especially those with bilateral surgery, but both are too difficult to quantify to use as a primary endpoint in trials [111–115]. Facial nerve palsy, unilateral or bilateral, is frequent in NF2 patients, can be graded using standard scales [116], and has a significant impact on HRQoL [117], but it more often represents an effect of surgical treatment or of coincident facial nerve schwannomas than of growing VS. Similarly, trigeminal nerve deficits often reflect consequences of treatment or of trigeminal schwannomas, and are difficult to quantify. Although we know of no comparable NF2-specific research, quality of life studies in patients with sporadic, unilateral VS consistently identify hearing loss as the deficit with

most impact on their subjective well being, particularly after surgery [115, 118, 119]. Although non-VS tumors such as schwannomas of other cranial nerves, meningiomas, and spinal tumors can certainly affect HRQoL adversely in some NF2 patients, not all patients have such tumors [7, 59, 120] and their neurological consequences vary so widely that uniform quantification is not possible.

Objective measurement of hearing function

Although hearing seems to be the obvious choice for a trial endpoint with direct relevance to patients' quality of life, hearing can be measured in many ways in modern audiometric practice, and adoption of a uniform hearing measure for the purposes of NF2 trials would be an advantage. Hearing measures include those included in most routine audiograms, such as simple detection of sound (measured as pure tone thresholds at individual stimulus frequencies or their average of several frequencies [pure tone average, PTA]) and speech discrimination (measured as word recognition score based on standardized word lists and presentation conditions). One composite hearing measure, commonly known as the Gardner–Robertson classification [90, 121] has been recommended for reporting hearing preservation after VS surgery and is commonly used to grade hearing after radiosurgical VS treatment as well. This scale combines PTA and word recognition scores into a four-level scale. Other hearing tests are frequently abnormal in VS patients, such as electrocochleography or auditory brainstem response (ABR) testing, but these methods commonly detect changes in auditory function that are not subjectively perceived by patients —that is, these abnormalities may be statistically, but not clinically significant. Other tests appraise the subject's assessment of their own hearing function in more subjective terms, such as the Glasgow Benefit Inventory [122], the Hearing Participation Scale [123], the Hearing Handicap Inventory for Adults [103], and the Hearing Satisfaction Scale [124], among others. These scales are commonly used in assessing the societal benefit of interventions that improve hearing (i.e., cost-utility studies, especially across different interventions), but are not typically used as primary outcome measures in such trials.

Word discrimination and hearing in NF2

Of the available means of testing hearing, we suggest that speech discrimination (word recognition score) be used as the primary hearing endpoint in phase II NF2 VS trials, for several reasons. First, word recognition scores directly measure the aspect of severe hearing loss that affects patients most acutely in daily life, namely communication [125]. Second, word recognition is affected disproportionately to pure tone thresholds in VS patients. VS cause hearing loss primarily through dysfunction of the auditory nerve, presumably because of direct compression from the tumor. Recent histopathologic data suggest that secondary degeneration of the cochlea may also be present, as manifest by loss of cochlear hair cells and spiral ganglion neurons and other histopathologic changes [126]. Together, these cochlear and neural alterations lead to the characteristic hearing loss pattern in VS patients, in which word recognition is affected to a greater extent than detection of pure tones [127]. In many cases of hearing loss from non- VS etiologies, such as conductive hearing loss, hearing aids can improve hearing loss with elevated thresholds by amplifying incoming sound presented to the cochlea. However, hearing aids cannot reliably address hearing lossrelated impaired word recognition, such as VS-related hearing loss, because poor quality

word information delivered at loud levels provides limited improvement. Indeed, while hearing with poor word recognition but relatively preserved pure tone thresholds may still have value in detecting environmental sounds (such as a car horn), its main value in communication may be only in detecting speech onset and orienting toward the speaker to assist lip reading. An improvement in word recognition from treatment would therefore be of special value in NF2 patients, who often have only one hearing ear and whose hearing deficits usually impose limited benefit from amplification. As a third reason for basing hearing response assessment on word recognition scores, we cite our own small clinical experience with treating NF2 patients with targeted agents, in which we observed some cases of striking improvement in word recognition with only modest PTA improvement; these patients considered their gains in hearing to be of real benefit [27–29] (S.R. Plotkin, unpublished data).

Gardner–Robertson scale as an NF2 trial endpoint

The Gardner–Robertson scale, a composite hearing endpoint that is in wide clinical use, has previously been recommended for use in reporting hearing preservation results after VS surgery [90, 121]. This scale uses both PTA and word recognition scores and classifies hearing into four levels. This scale was designed to identify "useful" hearing in ears where the ear being tested was the worse-hearing ear and the contralateral ear was normal, or nearly so. It therefore uses only one category to score all hearing ears with word recognition less than 50%. We feel that hearing differences that fall within this broad range have clinical significance when the ear being tested is the only hearing ear, as in many NF2 patients, and do not recommend this scale for use as a primary endpoint in NF2 trials because the score lacks sensitivity in this application. Similarly, the various criteria for success of steroid treatment of idiopathic sudden sensorineural hearing loss often assume good hearing function in the contralateral ear in defining response (that is, the affected ear must improve to a certain degree toward "normal" as defined by the unaffected ear) and are typically insensitive to minor improvements in the affected ear [128].

Time to hearing failure in NF2

As for tumor size, the natural history of hearing changes for PTA and word recognition in NF2 are only partly understood; few NF2 natural history reports have included data on hearing. No study reports the time over which PTA or word recognition scores can be expected to be stable in NF2 ears with growing VS, values which would be required to design single-arm phase II trials with a time to progression hearing endpoint. Abaza et al. [52] reported that baseline PTAs were worse in NF2 patients (ears) with larger VS and that older patients had worse PTAs. PTAs were more likely to decline over time intervals with active tumor growth. Similarly, worse speech recognition scores were found in ears with larger VS tumors. In the NF2 Natural History Consortium patients, PTA values showed a slow decline, with differences becoming significant only after 3–5 years of observation [53]. Only 4 of 106 ears (4%) declined from Gardner-Robinson hearing class A or B to class C or D over 7 months to 2 years of followup, and only 3 of 36 ears (8%) showed similar declines over 3–5 years of observation. This cohort was not limited to patients with growing VS, whose hearing might be expected to be at higher risk. The authors recommended against

using decline in hearing as an endpoint in NF2 trials because the endpoint was too infrequent (using their definition) and unrealistically large trials would be required.

Proposed hearing response criteria

We propose that *improvement in hearing as measured by word recognition scores* be considered as an important endpoint in initial NF2 VS phase II trials. Information from natural history studies does not exclude possible transient improvements in hearing in some patients, although the overall pattern of progressive decline over time is very clear. Clinical experience would indicate that such improvement in untreated patients is uncommon, small in magnitude, and transient in duration. In addition, neither surgery nor radiation treatment of VS is followed by improvement in hearing except in unusual cases [70–76]. For initial phase II trials in NF2 related VS, we suggest using the smallest change in word recognition score that can be detected with statistical significance as the criterion for a hearing response (see below). If an additional category of "minor hearing response" is desired, an appropriate change in PTA can be used (see below). Responses must be durable, i.e., confirmation of hearing response should be performed at least 2 months after the criteria for response are first met.

As an important research goal, we propose including an additional test such as the Glasgow Benefit Inventory to confirm the clinical significance of hearing improvement in only hearing ears when responses measured in this way are achieved.

In the next sections, we offer more detailed suggestions for assessing tumor volume and hearing function in phase II drug trials in NF2-related VS.

Measuring VS tumor response by volume

Definitions

At baseline, VS will be categorized as follows: measurable (lesions that are 0.5 ml or greater in volume when imaged with contrast enhanced MRI as detailed below) or nonmeasurable (all other lesions). VS previously treated with radiation are considered nonmeasurable, because both early increases and later gradual decreases in tumor volume after radiation might confound response assessment [129–135]. (An exception may be made for index VS that are unequivocally enlarging >5 years after prior radiation treatment.) VS that have previously been partially resected are measurable if investigators are certain that enhancement represents solid schwannoma tissue rather than linear postoperative enhancement that represents scar or inflammatory tissue, as is sometimes seen in the IAC after a complete resection [136–139]. Measurements should be recorded in metric notation. Baseline evaluations should be performed as closely as possible to the beginning of treatment and not more than 6 weeks prior to beginning treatment.

Patients with NF2 may have special circumstances that hinder measurement of a VS. Collision tumors (i.e., tumor masses formed by multiple tumors that abut each other; Fig. 3) in which the VS cannot be clearly distinguished from the neighboring tumor(s) are considered nonmeasurable. VS obscured by artifact related to placement of an auditory

brainstem implant (Fig. 4) or other devices such as some cerebrospinal fluid shunts are also considered nonmeasurable.

Specification of methods of measurement

The same techniques should be used to characterize VS at baseline and during follow-up. To declare a tumor response, patients should be on stable or decreasing corticosteroid doses compared to baseline.

Radiographic response is determined by quantitative volumetric tumor analysis of contrast-enhanced MRI scans [89]. MRI scans should include pre- and post-contrast images of the brain including thin spin-echo or gradientecho T1-weighted axial images through the internal auditory canal (slice 1.5 mm or thinner, skip 0 mm). Fatsuppression techniques should be used for post-contrast studies if there is a prior history of surgery in the internal auditory canal to avoid confusing fat grafts in the internal auditory canal for enhancing tumor [140]. Similarly, true fluid cysts within a VS or arachnoid cysts abutting a VS should be excluded from volume measurement. Tumor volumetry will typically be performed centrally in multicenter trials, and should be performed using post-contrast images through the internal auditory canal with the Vitrea2 (Vital Images, Minnetonka, MN) semi-automated segmentation software (or comparable software), edited manually as needed by the reviewer. The coefficient of variation for VS greater than 1 cc using this technique is generally below 5% [89], so we suggest that 5% would be a reasonable lower bound for definition of minor response. Although drugs tested to date do not affect VS enhancement, T2 images should also be examined to confirm that apparent volume reduction represents a true change in tumor size (as apparent from T2 visualization of tumor borders) rather than a change in intrinsic enhancement of VS tissue.

We anticipate that most NF2 phase II trials will define an "index" lesion to be used for response assessment, such as the VS on the side of an only hearing ear. (The definition of an index lesion may vary for trials that focus on other NF2 histologies such as meningiomas or ependymomas, or for patients with bilateral VS and no hearing in either ear.) In patients with bilateral VS the index VS should be selected before trial enrollment. Some NF2 patients have seemingly innumerable meningiomas and schwannomas; we propose that up to five lesions other than the index VS that can each be accurately measured by MRI scan should be identified as non-target lesions. Non-target lesions for NF2 patients should be defined by the study radiologist at baseline as contralateral VS, non-VS, and meningiomas. The volume of the contralateral VS and individual volumes for all non-VS and meningiomas (or, alternatively, total non-target schwannoma and total meningioma volumes) should be recorded at baseline. The change in volume of non-target lesions will not be used for response criteria but can be used and reported as secondary endpoints, if desired.

Radiographic response

In the proposed criteria, an objective radiographic response is defined as 20% reduction in tumor volume, comparing the nadir of tumor volumes measured during treatment to the baseline volume. If desired, a minor radiographic response may be defined as a reduction in tumor volume between 5% and 19%. Radiographic progression is defined as an increase in

tumor volume greater than 20% compared to the baseline evaluation. In general, only measurable tumors can be assessed for response. Occasionally, a VS that was nonmeasurable at baseline may cause progression to be declared, i.e., a tumor that was smaller than 0.5 ml at baseline that enlarges to greater than 0.5 ml during treatment.

Measuring VS tumor response by hearing assessment

Definitions and methods of measurement

Hearing in phase II studies will be measured by standard audiologic techniques including determination of word recognition scores and pure-tone averages ipsilateral to the target vestibular schwannoma.

Word recognition is measured using 50-item CID-W22 (Central Institute for the Deaf; Ira Hirsh recording) monosyllable wordlists and ranges from 0% to 100% [141]. Other standardized monosyllable recorded lists may be used including NU#6 [142], CNC [143] and others as long as similar performance-to-intensity functions can be shown for the recording used. Because word lists can differ in difficulty, the word list used for an individual in a trial should be noted and kept consistent for all testing throughout the protocol treatment. The principle of using equivalent performance functions may also be applied to word lists in other languages, such as bisyllabic Spanish words [144]. When the tester is an English speaker, we recommend using a trial entry criterion of "testable in English" (rather than "English as a first language") which promotes inclusion without reducing validity [145]. In all cases, live voice presentation and truncated (e.g., 25-word) lists are not acceptable.

Pure-tone thresholds are most often measured at the frequencies of 250, 500, 1000, 2000, 3000, 4000, 6000 and 8000 Hz. The pure-tone average is calculated as the average of thresholds at 500, 1000, 2000 and 3000 Hz as recommended by the American Academy of Otolaryngology (AAO) [146]. In cases of VS, the hearing loss is sensorineural. However, an unrelated, additional conductive loss may also be present which would further elevate the air conduction thresholds. While air conduction thresholds (with masking as necessary) are most often the best data, in cases with additional conductive loss, the masked bone conduction results are a better measure of the sensorineural hearing loss. It is recommended that masked bone conduction results be used when air-bone gaps larger than 10 dB are found at more than one frequency.

When forming a PTA, it is sometimes necessary to include data where the upper limits of the audiometer is reached and the patient has not responded. This should not be treated as missing data. Our recommended approach is to extrapolate conservatively by adding one audiometric step (5 dB) to the upper limit value, thereby capturing the large magnitude of the loss at that frequency [145].

Hearing response

Because reduced word recognition directly limits patients' quality of life, the primary audiologic outcome will be word recognition, with response defined as the smallest improvement in word recognition score over baseline that meets criteria for statistical

significance at the P = 0.05 level. Table 1 shows values for detecting a statistically significant change in word recognition at the P =0.05 level for different baseline scores [147]. The table can be used to compare any two scores from the same patient, such as before and after treatment [148].

When hearing variables are used as response criteria, it is important to recognize the presence of "ceiling" and "floor" effects. Normal audiologic values (100% correct for word recognition and 0 dBHL for thresholds and PTA) represent a limit for improvement at one extreme of the range (i.e., "ceiling" effect). At the other end of the spectrum, non-detectable audiologic values (0% correct for word recognition and equipment maximum [e.g., 110 dBHL] for thresholds and PTA) represent a limit for decline at the other extreme of the range (i.e., the "floor" effect). As a result of the ceiling effect, patients with baseline word recognition scores above 94% cannot achieve a statistically significant improvement in hearing at the P = 0.05 level with treatment and hence are nonevaluable for hearing response using these criteria. (As a practical matter, one or two-word lapses in attention will prevent patients with baseline word recognition scores ranging between 86% and 94% from achieving 100% on subsequent tests even if their hearing improved to normal. Investigators may wish to consider such patients nonevaluable for hearing responses in designing protocols for which hearing response is a primary endpoint.) In practice, the magnitude of potential improvement on audiologic measures is related to the baseline hearing loss. For example, a patient with word recognition scores of 60% can potentially recover "twice as much" as a patient with word recognition scores of 80%. This concept applies equally to PTA: patients with worse (larger) starting threshold values have more freedom to improve than those who with better threshold values.

A hearing response is defined as an improvement in the word recognition scores above the 95% critical difference threshold (Table 1), comparing the best word recognition score achieved during treatment to the baseline word recognition score. To declare a hearing response, patients should be on stable or decreasing corticosteroid doses compared to baseline. Hearing should be stable at baseline or slowly decreasing, i.e., patients with sudden hearing loss immediately before beginning treatment should not be considered eligible for a hearing response. Such patients often regain hearing spontaneously or with corticosteroid treatment, at least in sporadic VS [149–151].

If desired, investigators may also define a minor hearing response as an improvement in the pure-tone average of 10 dB in the setting of stable word recognition scores (i.e., within the 95% critical difference threshold defined in Table 1), taking as reference the baseline pure-tone average.

Progression in hearing deficit is defined as a decrease in the word recognition scores below the 95% critical difference threshold (Table 1), taking as reference the baseline word recognition score. Stable hearing is defined as all other responses.

Discussion

In this communication, we suggest two important means of assessing tumor response of NF2-related VS to drug treatment: radiographic tumor response (defined as a 20% reduction in tumor volume) and hearing response (defined as a statistically significant improvement in word recognition score). Prior experience with chemotherapy for VS has been very limited, and no standard response criteria for such trials have been defined. Riccardi reported improvement in depression, weight gain and improved muscle mass, gait improvement, absence of tumor-associated pain, and "slowing of neurofibromas growth" in one NF patient treated using ketotifen [152], and Jahrsdoerfer and Benjamin [153] reported cessation of tumor growth and stabilization or improvement in hearing in two patients treated with cyclophosphamide, doxorubicin and dacarbazine. Neither report used predefined criteria to classify response.

Recent advances in the understanding of schwannoma tumor biology [20–26], combined with development of active anticancer drugs designed to target specific molecular pathways in neoplastic cells, suggest that trials of drug treatment for NF2-related VS may now be warranted. We suggest that the criteria defined here deserve consideration for inclusion as primary endpoints in such trials, at least initially. When more detailed information becomes available on the time to tumor progression and time to hearing loss in untreated NF2 VS, these endpoints are likely to replace tumor shrinkage and hearing improvement in trials. However, because many NF2 patients represent new somatic mutations in the NF2 gene and do not belong to known NF2 kindreds [4], some patients will continue to present as new diagnoses with poor hearing in both ears. Such patients would benefit from drugs that cause hearing improvement, rather than stability. In addition, a recent review of phase II trial design for targeted anticancer agents (many of which would be expected to be cytostatic) showed that most agents reaching regulatory approval generated objective responses in phase II testing; this was correlated with endpoints reflecting tumor stability [154].

Special considerations for primary endpoints in NF2 trials, and composite endpoints

Some special considerations in NF2 phase II trials may prevent either radiographic response or hearing response from serving all of the typical functions of a phase II primary endpoint. For example, patients in a phase II trial whose tumors progress according to the primary endpoint (e.g., radiographic criteria) do not normally continue to receive the study drug and are no longer treated on protocol. However, in an NF2 trial some patients may achieve a hearing response to treatment while developing progressive disease by MRI scan. Although treatment is deemed to have failed in these patients by traditional response criteria, we feel these patients should be allowed to continue treatment with the study drug on a compassionate- use basis at the discretion of the study chair in an NF2 trial. We believe this can be justified based on the current standard of care for NF2-related VS, in which growing tumors are often monitored without active intervention to preserve hearing. Conversely, some patients might enter an NF2 VS trial because of a growing VS in an only hearing ear with word recognition scores above 94%. Such patients would be ineligible for hearing responses because of the "ceiling effect" described above, although drug activity could still be detected through changes in tumor size.

One approach to the problem of two endpoints, either one of which might be more important in a given clinical situation, is to combine the two into a composite endpoint. Table 2 shows one possible means of integrating objective radiographic response and hearing response in NF2 VS. As shown, the combination of major and minor radiographic and hearing responses can be used to construct a composite endpoint with four response categories: major response, minor response, stable disease, and progressive disease. Whether a composite endpoint of this nature had value would probably become clear as clinical experience with such trials accrues. Initially, we suggest using objective (radiographic) tumor response as the major endpoint for designing phase II trials, because its measurement in relation to known NF2 VS natural history seems most secure. All patients enrolled in an NF2 VS trial would be expected to be eligible for volumetric response, while our preliminary experience suggests only half of patients might be eligible for a hearing response or composite endpoint response. A volumetric primary endpoint would increase efficiency in this rare disease. Hearing response rates, though, also deserve significant weight in judging a drug's clinical benefit in NF2 related VS and whether to progress to phase III testing—as do response durability, treatment toxicity, and other factors.

Future research and novel endpoints

This communication also identifies several important research goals in NF2 natural history in order to facilitate future development of new endpoints for NF2 trials. These goals include identifying the time to tumor progression and time to measurable hearing loss in untreated NF2-related VS, and the relation of both endpoints to patient prognostic factors (such as age, baseline tumor volume, and measures of disease severity). In addition, researchers should seek better definition of the smallest improvement in hearing (specifically word recognition) that has clinical significance for NF2 patients with only hearing ears (i.e., the minimal clinically significant difference). If the test for word recognition using standard audiometric technique that we describe proves to be insensitive to small but meaningful hearing improvements in the context of an only hearing ear, more sensitive hearing endpoint measures may need to be developed. Finally, the possibility of using a composite endpoint in NF2 trials, such as the one we describe here, should be considered for future use in selecting drugs for phase III trials in this rare disease.

References

1. Central Brain Tumor Registry of the United States. Statistical Report: primary brain tumors in the United States 2000–2004. Hinsdale, IL: CBTRUS; 2008.

2. Regis, J.; Roche, P-H., editors. Modern management of acoustic neuroma. Basel: Karger; 2008.

3. Evans DG, Moran A, King A, Saeed S, Gurusinghe N, Ramsden R. Incidence of vestibular schwannoma and neurofibromatosis2 in the North West of England over a 10-year period: higher incidence than previously thought. Otol Neurotol. 2005; 26:93–97. [PubMed: 15699726]

4. Evans DG, Huson SM, Donnai D, Neary W, Blair V, Teare D, Newton V, Strachan T, Ramsden R, Harris R. A genetic study of type 2 neurofibromatosis in the United Kingdom. I. Prevalence, mutation rate, fitness, and confirmation of maternal transmission effect on severity. J Med Genet. 1992; 29:841–846. [PubMed: 1479598]

5. Barker FG II, Carter BS, Ojemann RG, Jyung RW, Poe DS, McKenna MJ. Surgical excision of acoustic neuroma: patient outcome and provider caseload. Laryngoscope. 2003; 113:1332–1343. [PubMed: 12897555]

6. Evans DG, Huson SM, Donnai D, Neary W, Blair V, Newton V, Harris R. A clinical study of type 2 neurofibromatosis. Q J Med. 1992; 84:603–618. [PubMed: 1484939]

7. Otsuka G, Saito K, Nagatani T, Yoshida J. Age at symptom onset and long-term survival in patients with neurofibromatosis Type 2. J Neurosurg. 2003; 99:480–483. [PubMed: 12959433]

8. Baser ME, Friedman JM, Aeschliman D, Joe H, Wallace AJ, Ramsden RT, Evans DG. Predictors of the risk of mortality in neurofibromatosis 2. Am J Hum Genet. 2002; 71:715–723. [PubMed: 12235555]

9. Samii M, Matthies C, Tatagiba M. Management of vestibular schwannomas (acoustic neuromas): auditory and facial nerve function after resection of 120 vestibular schwannomas in patients with neurofibromatosis 2. Neurosurgery. 1997; 40:696–705. discussion 705-706. [PubMed: 9092842]

10. Mathieu D, Kondziolka D, Flickinger JC, Niranjan A, Williamson R, Martin JJ, Lunsford LD. Stereotactic radiosurgery for vestibular schwannomas in patients with neurofibromatosis type 2: an analysis of tumor control, complications, and hearing preservation rates. Neurosurgery. 2007; 60:460–468. discussion 468–470. [PubMed: 17327790]

11. Rowe J, Radatz M, Kemeny A. Radiosurgery for type II neurofibromatosis. Prog Neurol Surg. 2008; 21:176–182. [PubMed: 18810217]

12. Phi JH, Kim DG, Chung HT, Lee J, Paek SH, Jung HW. Radiosurgical treatment of vestibular schwannomas in patients with neurofibromatosis type 2: tumor control and hearing preservation. Cancer. 2009; 115:390–398. [PubMed: 19109818]

13. Sobel RA. Vestibular (acoustic) schwannomas: histologic features in neurofibromatosis 2 and in unilateral cases. J Neuropathol Exp Neurol. 1993; 52:106–113. [PubMed: 8440992]

14. Wechsler J, Lantieri L, Zeller J, Voisin MC, Martin-Garcia N, Wolkenstein P. Aberrant axon neurofilaments in schwannomas associated with phacomatoses. Virchows Arch. 2003; 443:768–773. [PubMed: 14508685]

15. Hamada Y, Iwaki T, Fukui M, Tateishi J. A comparative study of embedded nerve tissue in six NF2-associated schwannomas and 17 nonassociated NF2 schwannomas. Surg Neurol. 1997; 48:395–400. [PubMed: 9315139]

16. Jaaskelainen J, Paetau A, Pyykko I, Blomstedt G, Palva T, Troupp H. Interface between the facial nerve and large acoustic neurinomas. Immunohistochemical study of the cleavage plane in NF2 and non-NF2 cases. J Neurosurg. 1994; 80:541–547. [PubMed: 8113868]

17. Sakamoto T, Shirato H, Takeichi N, Aoyama H, Fukuda S, Miyasaka K. Annual rate of hearing loss falls after fractionated stereotactic irradiation for vestibular schwannoma. Radiother Oncol. 2001; 60:45–48. [PubMed: 11410303]

18. Combs SE, Volk S, Schulz-Ertner D, Huber PE, Thilmann C, Debus J. Management of acoustic neuromas with fractionated stereotactic radiotherapy (FSRT): long-term results in 106 patients treated in a single institution. Int J Radiat Oncol Biol Phys. 2005; 63:75–81. [PubMed: 16111574]

19. Ju DT, Lin JW, Lin MS, Lee LM, Tseng HM, Wei CP, Yen CH, Hung CC, Hung KS, Lin CM, Lin TJ, Chiu WT, Tsai JT. Hypofractionated CyberKnife stereotactic radiosurgery for acoustic neuromas with and without association to neurofibromatosis Type 2. Acta Neurochir Suppl (Wien). 2008; 101:169–173.

20. Chang SM, Lamborn KR, Kuhn JG, Yung WK, Gilbert MR, Wen PY, Fine HA, Mehta MP, DeAngelis LM, Lieberman FS, Cloughesy TF, Robins HI, Abrey LE, Prados MD. Neurooncology clinical trial design for targeted therapies: lessons learned from the North American Brain Tumor Consortium. Neuro-Oncology. 2008; 10:631–642. [PubMed: 18559968]

21. Welling DB, Packer MD, Chang LS. Molecular studies of vestibular schwannomas: a review. Curr Opin Otolaryngol Head Neck Surg. 2007; 15:341–346. [PubMed: 17823551]

22. Neff BA, Welling DB, Akhmametyeva E, Chang LS. The molecular biology of vestibular schwannomas: dissecting the pathogenic process at the molecular level. Otol Neurotol. 2006; 27:197–208. [PubMed: 16436990]

23. Caye-Thomasen P, Werther K, Nalla A, Bog-Hansen TC, Nielsen HJ, Stangerup SE, Thomsen J. VEGF and VEGF receptor-1 concentration in vestibular schwannoma homogenates correlates to tumor growth rate. Otol Neurotol. 2005; 26:98–101. [PubMed: 15699727]

24. Scoles DR. The merlin interacting proteins reveal multiple targets for NF2 therapy. Biochim Biophys Acta. 2008; 1785:32–54. [PubMed: 17980164]

25. Hanemann CO. Magic but treatable? Tumours due to loss of merlin. Brain. 2008; 131:606–615. [PubMed: 17940085]

26. Doherty JK, Ongkeko W, Crawley B, Andalibi A, Ryan AF. ErbB and Nrg: potential molecular targets for vestibular schwannoma pharmacotherapy. Otol Neurotol. 2008; 29:50–57. [PubMed: 18199957]

27. Plotkin SR, Halpin CF, McKenna MJ, Batchelor TT, Loeffler JS, Barker FG. Treatment of progressive NF2-related vestibular schwannoma with erlotinib. J Clin Oncol. 2008; 26(15 Suppl 1):100s. (Abstract 2045).

28. Plotkin SR, Singh MA, O'Donnell CC, Harris GJ, McClatchey AI, Halpin C. Audiologic and radiographic response of NF2-related vestibular schwannoma to erlotinib therapy. Nat Clin Pract Oncol. 2008; 5:487–491. [PubMed: 18560388]

29. Plotkin SR, Halpin C, Stemmer-Rachamimoff A, Barker F, Sorensen G, Jain R, Di Tomaso E. Activity of bevacizumab against NF2-related vestibular schwannomas. Neuro-Oncology. 2008; 10:796. [Abstract ET-51].

30. Freidlin B, Simon R. Evaluation of randomized discontinuation design. J Clin Oncol. 2005; 23:5094–5098. [PubMed: 15983399]

31. Rubinstein LV, Korn EL, Freidlin B, Hunsberger S, Ivy SP, Smith MA. Design issues of randomized phase II trials and a proposal for phase II screening trials. J Clin Oncol. 2005; 23:7199–7206. [PubMed: 16192604]

32. Jones, B.; Kenward, MG. Design and analysis of cross-over trials. 2nd edn.. Boca Raton, FL: Chapman & Hall; 2003.

33. Stadler WM. The randomized discontinuation trial: a phase II design to assess growth-inhibitory agents. Mol Cancer Ther. 2007; 6:1180–1185. [PubMed: 17431101]

34. Koutras AK, Mastronikolis NS, Evans TR, Papadeas ES, Makatsoris T, Kalofonos HP. Irreversible ototoxicity associated with the use of erlotinib in a patient with pancreatic cancer. Acta Oncol. 2008; 47:1171–1173. [PubMed: 18615326]

35. Nayfield SG, Gorin MB. Tamoxifen-associated eye disease. A review. J Clin Oncol. 1996; 14:1018–1026. [PubMed: 8622006]

36. Vickers AJ, Ballen V, Scher HI. Setting the bar in phase II trials: the use of historical data for determining "go/no go" decision for definitive phase III testing. Clin Cancer Res. 2007; 13:972–976. [PubMed: 17277252]

37. Buyse M, Thirion P, Carlson RW, Burzykowski T, Molenberghs G, Piedbois P. Relation between tumour response to first-line chemotherapy and survival in advanced colorectal cancer: a meta-analysis. Meta-Analysis Group in Cancer. Lancet. 2000; 356:373–378. [PubMed: 10972369]

38. Paesmans M, Sculier JP, Libert P, Bureau G, Dabouis G, Thiriaux J, Michel J, Van Cutsem O, Sergysels R, Mommen P, Klastersky J. Response to chemotherapy has predictive value for further survival of patients with advanced non-small cell lung cancer. 10 years experience of the European Lung Cancer Working Party. Eur J Cancer. 1997; 33:2326–2332. [PubMed: 9616276]

39. Burzykowski T, Buyse M, Piccart-Gebhart MJ, Sledge G, Carmichael J, Luck HJ, Mackey JR, Nabholtz JM, Paridaens R, Biganzoli L, Jassem J, Bontenbal M, Bonneterre J, Chan S, Basaran GA, Therasse P. Evaluation of tumor response, disease control, progression-free survival, and time to progression as potential surrogate end points in metastatic breast cancer. J Clin Oncol. 2008; 26:1987–1992. [PubMed: 18421050]

40. Johnson KR, Ringland C, Stokes BJ, Anthony DM, Freemantle N, Irs A, Hill SR, Ward RL. Response rate or time to progression as predictors of survival in trials of metastatic colorectal cancer or non-small-cell lung cancer: a meta-analysis. Lancet Oncol. 2006; 7:741–746. [PubMed: 16945769]

41. Brandsma D, Stalpers L, Taal W, Sminia P, van den Bent MJ. Clinical features, mechanisms, and management of pseudoprogression in malignant gliomas. Lancet Oncol. 2008; 9:453–461. [PubMed: 18452856]

42. de Wit MC, de Bruin HG, Eijkenboom W, Sillevis Smitt PA, van den Bent MJ. Immediate post-radiotherapy changes in malignant glioma can mimic tumor progression. Neurology. 2004; 63:535–537. [PubMed: 15304589]

43. Lamborn KR, Yung WK, Chang SM, Wen PY, Cloughesy TF, DeAngelis LM, Robins HI, Lieberman FS, Fine HA, Fink KL, Junck L, Abrey L, Gilbert MR, Mehta M, Kuhn JG, Aldape KD, Hibberts J, Peterson PM, Prados MD. Progression-free survival: an important end point in evaluating therapy for recurrent high-grade gliomas. Neuro-Oncology. 2008; 10:162–170. [PubMed: 18356283]

44. Levin VA, Ictech S, Hess KR. Impact of phase II trials with progression-free survival as end-points on survival-based phase III studies in patients with anaplastic gliomas. BMC Cancer. 2007; 7:106. [PubMed: 17587447]

45. Ballman KV, Buckner JC, Brown PD, Giannini C, Flynn PJ, LaPlant BR, Jaeckle KA. The relationship between sixmonth progression-free survival and 12-month overall survival end points for phase II trials in patients with glioblastoma multiforme. Neuro-Oncology. 2007; 9:29–38. [PubMed: 17108063]

46. Choi H, Charnsangavej C, Faria SC, Macapinlac HA, Burgess MA, Patel SR, Chen LL, Podoloff DA, Benjamin RS. Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumor treated at a single institution with imatinib mesylate: proposal of new computed tomography response criteria. J Clin Oncol. 2007; 25:1753–1759. [PubMed: 17470865]

47. Benjamin RS, Choi H, Macapinlac HA, Burgess MA, Patel SR, Chen LL, Podoloff DA, Charnsangavej C. We should desist using RECIST, at least in GIST. J Clin Oncol. 2007; 25:1760–1764. [PubMed: 17470866]

48. Schuetze SM, Rubin BP, Vernon C, Hawkins DS, Bruckner JD, Conrad EU III, Eary JF. Use of positron emission tomography in localized extremity soft tissue sarcoma treated with neoadjuvant chemotherapy. Cancer. 2005; 103:339–348. [PubMed: 15578712]

49. Gerstner ER, Sorensen AG, Jain RK, Batchelor TT. Advances in neuroimaging techniques for the evaluation of tumor growth, vascular permeability, and angiogenesis in gliomas. Curr Opin Neurol. 2008; 21:728–735. [PubMed: 18989120]

50. Batchelor TT, Sorensen AG, di Tomaso E, Zhang WT, Duda DG, Cohen KS, Kozak KR, Cahill DP, Chen PJ, Zhu M, Ancukiewicz M, Mrugala MM, Plotkin S, Drappatz J, Louis DN, Ivy P, Scadden DT, Benner T, Loeffler JS, Wen PY, Jain RK. AZD2171, a pan-VEGF receptor tyrosine kinase inhibitor, normalizes tumor vasculature and alleviates edema in glioblastoma patients. Cancer Cell. 2007; 11:83–95. [PubMed: 17222792]

51. Sorensen AG, Batchelor TT, Wen PY, Zhang WT, Jain RK. Response criteria for glioma. Nat Clin Pract Oncol. 2008; 5:634–644. [PubMed: 18711427]

52. Abaza MM, Makariou E, Armstrong M, Lalwani AK. Growth rate characteristics of acoustic neuromas associated with neurofibromatosis type 2. Laryngoscope. 1996; 106:694–699. [PubMed: 8656953]

53. Masuda A, Fisher LM, Oppenheimer ML, Iqbal Z, Slattery WH. Hearing changes after diagnosis in neurofibromatosis type 2. Otol Neurotol. 2004; 25:150–154. [PubMed: 15021775]

54. Hajioff D, Raut VV, Walsh RM, Bath AP, Bance ML, Guha A, Tator CH, Rutka JA. Conservative management of vestibular schwannomas: third review of a 10-year prospective study. Clin Otolaryngol. 2008; 33:255–259. [PubMed: 18559034]

55. Massick DD, Welling DB, Dodson EE, Scholfield M, Nagaraja HN, Schmalbrock P, Chakeres DW. Tumor growth and audiometric change in vestibular schwannomas managed conservatively. Laryngoscope. 2000; 110:1843–1849. [PubMed: 11081597]

56. Walsh RM, Bath AP, Bance ML, Keller A, Rutka JA. Consequences to hearing during the conservative management of vestibular schwannomas. Laryngoscope. 2000; 110:250–255. [PubMed: 10680925]

57. Tschudi DC, Linder TE, Fisch U. Conservative management of unilateral acoustic neuromas. Am J Otol. 2000; 21:722–728. [PubMed: 10993466]

58. Sampath P, Holliday MJ, Brem H, Niparko JK, Long DM. Facial nerve injury in acoustic neuroma (vestibular schwannoma) surgery: etiology and prevention. J Neurosurg. 1997; 87:60–66. [PubMed: 9202266]

59. Mautner VF, Baser ME, Thakkar SD, Feigen UM, Friedman JM, Kluwe L. Vestibular schwannoma growth in patients with neurofibromatosis Type 2: a longitudinal study. J Neurosurg. 2002; 96:223–228. [PubMed: 11838794]

60. Slattery WH III, Fisher LM, Iqbal Z, Oppenhiemer M. Vestibular schwannoma growth rates in neurofibromatosis type 2 natural history consortium subjects. Otol Neurotol. 2004; 25:811–817. [PubMed: 15354016]

61. Linskey ME, Lunsford LD, Flickinger JC. Tumor control after stereotactic radiosurgery in neurofibromatosis patients with bilateral acoustic tumors. Neurosurgery. 1992; 31:829–838. discussion 838-829. [PubMed: 1436407]

62. Baser ME, Makariou EV, Parry DM. Predictors of vestibular schwannoma growth in patients with neurofibromatosis Type 2. J Neurosurg. 2002; 96:217–222. [PubMed: 11838793]

63. Kida Y, Kobayashi T, Tanaka T, Mori Y. Radiosurgery for bilateral neurinomas associated with neurofibromatosis type 2. Surg Neurol. 2000; 53:383–389. discussion 389–390. [PubMed: 10825525]

64. Baser ME, Mautner VF, Parry DM, Evans DG. Methodological issues in longitudinal studies: vestibular schwannoma growth rates in neurofibromatosis 2. J Med Genet. 2005; 42:903–906. [PubMed: 15831594]

65. Rowe JG, Radatz MW, Walton L, Soanes T, Rodgers J, Kemeny AA. Clinical experience with gamma knife stereotactic radiosurgery in the management of vestibular schwannomas secondary to type 2 neurofibromatosis. J Neurol Neurosurg Psychiatry. 2003; 74:1288–1293. [PubMed: 12933938]

66. Bederson JB, von Ammon K, Wichmann WW, Yasargil MG. Conservative treatment of patients with acoustic tumors. Neurosurgery. 1991; 28:646–651. [PubMed: 1876241]

67. Battaglia A, Mastrodimos B, Cueva R. Comparison of growth patterns of acoustic neuromas with and without radiosurgery. Otol Neurotol. 2006; 27:705–712. [PubMed: 16868519]

68. Yoshimoto Y. Systematic review of the natural history of vestibular schwannoma. J Neurosurg. 2005; 103:59–63. [PubMed: 16121974]

69. Smouha EE, Yoo M, Mohr K, Davis RP. Conservative management of acoustic neuroma: a meta-analysis and proposed treatment algorithm. Laryngoscope. 2005; 115:450–454. [PubMed: 15744156]

70. Meiteles LZ, Liu JK, Couldwell WT. Hearing restoration after resection of an intracanalicular vestibular schwannoma: a role for emergency surgery? Case report and review of the literature. J Neurosurg. 2002; 96:796–800. [PubMed: 11990824]

71. Niranjan A, Lunsford LD, Flickinger JC, Maitz A, Kondziolka D. Can hearing improve after acoustic tumor radiosurgery? Neurosurg Clin N Am. 1999; 10:305–315. [PubMed: 10099095]

72. Stidham KR, Roberson JB Jr. Hearing improvement after middle fossa resection of vestibular schwannoma. Otol Neurotol. 2001; 22:917–921. [PubMed: 11698819]

73. Inoue Y, Ogawa K, Kanzaki J. Hearing improvement after tumor removal in a vestibular schwannoma patient with severe hearing loss. Eur Arch Otorhinolaryngol. 2003; 260:487–489. [PubMed: 12736745]

74. Fischer G, Costantini JL, Mercier P. Improvement of hearing after microsurgical removal of acoustic neurinoma. Neurosurgery. 1980; 7:154–159. [PubMed: 6968414]

75. Shelton C, House WF. Hearing improvement after acoustic tumor removal. Otolaryngol Head Neck Surg. 1990; 103:963–965. [PubMed: 2126130]

76. Telian SA, Kemink JL, Kileny P. Hearing recovery following suboccipital excision of acoustic neuroma. Arch Otolaryngol Head Neck Surg. 1988; 114:85–87. [PubMed: 3334825]

77. Gray R, Manola J, Saxman S, Wright J, Dutcher J, Atkins M, Carducci M, See W, Sweeney C, Liu G, Stein M, Dreicer R, Wilding G, DiPaola RS. Phase II clinical trial design: methods in translational research from the Genitourinary Committee at the Eastern Cooperative Oncology Group. Clin Cancer Res. 2006; 12:1966–1969. [PubMed: 16609005]

78. Charabi S, Thomsen J, Mantoni M, Charabi B, Jorgensen B, Borgesen SE, Gyldensted C, Tos M. Acoustic neuroma (vestibular schwannoma): growth and surgical and nonsurgical consequences of the wait-and-see policy. Otolaryngol Head Neck Surg. 1995; 113:5–14. [PubMed: 7603722]

79. Luetje CM. Spontaneous involution of acoustic tumors. Am J Otol. 2000; 21:393–398. [PubMed: 10821554]

80. Yasumoto Y, Ito M. Spontaneous regression of a growing vestibular schwannoma. Neurol Med Chir (Tokyo). 2006; 46:601–604. [PubMed: 17185887]

81. Redleaf MI, McCabe BF. Disappearing recurrent acoustic neuroma in an elderly woman. Ann Otol Rhinol Laryngol. 1993; 102:518–520. [PubMed: 8333673]

82. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, Verweij J, Van Glabbeke M, van Oosterom AT, Christian MC, Gwyther SG. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. J Natl Cancer Inst. 2000; 92:205–216. [PubMed: 10655437]

83. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. Cancer. 1981; 47:207–214. [PubMed: 7459811]

84. Macdonald DR, Cascino TL, Schold SC Jr, Cairncross JG. Response criteria for phase II studies of supratentorial malignant glioma. J Clin Oncol. 1990; 8:1277–1280. [PubMed: 2358840]

85. Shah GD, Kesari S, Xu R, Batchelor TT, O'Neill AM, Hochberg FH, Levy B, Bradshaw J, Wen PY. Comparison of linear and volumetric criteria in assessing tumor response in adult high-grade gliomas. Neuro-Oncology. 2006; 8:38–46. [PubMed: 16443946]

86. Stangerup SE, Caye-Thomasen P, Tos M, Thomsen J. The natural history of vestibular schwannoma. Otol Neurotol. 2006; 27:547–552. [PubMed: 16791048]

87. Marshall AH, Owen VM, Nikolopoulos TP, O'Donoghue GM. Acoustic schwannomas: awareness of radiologic error will reduce unnecessary treatment. Otol Neurotol. 2005; 26:512–515. [PubMed: 15891658]

88. Walsh RM, Bath AP, Bance ML, Keller A, Tator CH, Rutka JA. The natural history of untreated vestibular schwannomas. Is there a role for conservative management? Rev Laryngol Otol Rhinol (Bord). 2000; 121:21–26. [PubMed: 10865479]

89. Harris GJ, Plotkin SR, Maccollin M, Bhat S, Urban T, Lev MH, Slattery WH. Three-dimensional volumetrics for tracking vestibular schwannoma growth in neurofibromatosis type II. Neurosurgery. 2008; 62:1314–1319. discussion 1319–1320. [PubMed: 18824998]

90. Committee on Hearing and Equilibrium. guidelines for the evaluation of hearing preservation in acoustic neuroma (vestibular schwannoma). American Academy of Otolaryngology-Head and Neck Surgery Foundation, Inc. Otolaryngol Head Neck Surg. 1995; 113:179–180.

91. Roche PH, Robitail S, Regis J. Two- and three dimensional measures of vestibular schwannomas and posterior fossa— implications for the treatment. Acta Neurochir (Wien). 2007; 149:267–273. discussion 273. [PubMed: 17342379]

92. Slattery WH, Lev MH, Fisher LM, Connell SS, Iqbal Z, Go JL. MRI evaluation of neurofibromatosis 2 patients: a standardized approach for accuracy in interpretation. Otol Neurotol. 2005; 26:733–740. [PubMed: 16015177]

93. Snell JW, Sheehan J, Stroila M, Steiner L. Assessment of imaging studies used with radiosurgery: a volumetric algorithm and an estimation of its error. Technical note. J Neurosurg. 2006; 104:157–162. [PubMed: 16509161]

94. Henson JW, Ulmer S, Harris GJ. Brain tumor imaging in clinical trials. AJNR Am J Neuroradiol. 2008; 29:419–424. [PubMed: 18272557]

95. Babovic-Vuksanovic D, Ballman K, Michels V, McGrann P, Lindor N, King B, Camp J, Micic V, Babovic N, Carrero X, Spinner R, O'Neill B. Phase II trial of pirfenidone in adults with neurofibromatosis type 1. Neurology. 2006; 67:1860–1862. [PubMed: 17035676]

96. Babovic-Vuksanovic D, Widemann BC, Dombi E, Gillespie A, Wolters PL, Toledo-Tamula MA, O'Neill BP, Fox E, Mac-Donald T, Beck H, Packer RJ. Phase I trial of pirfenidone in children with neurofibromatosis 1 and plexiform neurofibromas. Pediatr Neurol. 2007; 36:293–300. [PubMed: 17509460]

97. Barlow JH, Turner AP, Hammond CL, Gailey L. Living with late deafness: insight from between worlds. Int J Audiol. 2007; 46:442–448. [PubMed: 17654086]

98. Petrou S, McCann D, Law CM, Watkin PM, Worsfold S, Kennedy CR. Health status and health-related quality of life preference-based outcomes of children who are aged 7 to 9 years and have

bilateral permanent childhood hearing impairment. Pediatrics. 2007; 120:1044–1052. [PubMed: 17974742]

99. Chia EM, Wang JJ, Rochtchina E, Cumming RR, Newall P, Mitchell P. Hearing impairment and health-related quality of life: the Blue Mountains Hearing Study. Ear Hear. 2007; 28:187–195. [PubMed: 17496670]

100. Dalton DS, Cruickshanks KJ, Klein BE, Klein R, Wiley TL, Nondahl DM. The impact of hearing loss on quality of life in older adults. Gerontologist. 2003; 43:661–668. [PubMed: 14570962]

101. Mulrow CD, Aguilar C, Endicott JE, Velez R, Tuley MR, Charlip WS, Hill JA. Association between hearing impairment and the quality of life of elderly individuals. J Am Geriatr Soc. 1990; 38:45–50. [PubMed: 2295767]

102. Ringdahl A, Grimby A. Severe-profound hearing impairment and health-related quality of life among post-lingual deafened Swedish adults. Scand Audiol. 2000; 29:266–275. [PubMed: 11195947]

103. Vermeire K, Brokx JP, Wuyts FL, Cochet E, Hofkens A, Van de Heyning PH. Quality-of-life benefit from cochlear implantation in the elderly. Otol Neurotol. 2005; 26:188–195. [PubMed: 15793403]

104. Lassaletta L, Castro A, Bastarrica M, de Sarria MJ, Gavilan J. Quality of life in postlingually deaf patients following cochlear implantation. Eur Arch Otorhinolaryngol. 2006; 263:267–270. [PubMed: 16025257]

105. Shin YJ, Fraysse B, Deguine O, Vales O, Laborde ML, Bouccara D, Sterkers O, Uziel A. Benefits of cochlear implantation in elderly patients. Otolaryngol Head Neck Surg. 2000; 122:602–606. [PubMed: 10740189]

106. Metselaar M, Maat B, Krijnen P, Verschuure H, Dreschler WA, Feenstra L. Self-reported disability and handicap after hearing-aid fitting and benefit of hearing aids: comparison of fitting procedures, degree of hearing loss, experience with hearing aids and uni- and bilateral fittings. Eur Arch Otorhinolaryngol. 2009

107. Klop WM, Boermans PP, Ferrier MB, van den Hout WB, Stiggelbout AM, Frijns JH. Clinical relevance of quality of life outcome in cochlear implantation in postlingually deafened adults. Otol Neurotol. 2008; 29:615–621. [PubMed: 18451751]

108. Chisolm TH, Johnson CE, Danhauer JL, Portz LJ, Abrams HB, Lesner S, McCarthy PA, Newman CW. A systematic review of health-related quality of life and hearing aids: final report of the American Academy of Audiology Task Force On the Health-Related Quality of Life Benefits of Amplification in Adults. J Am Acad Audiol. 2007; 18:151–183. [PubMed: 17402301]

109. Morzaria S, Westerberg BD, Anzarut A. Quality of life following ear surgery measured by the 36-item Short Form Health Survey and the Glasgow Benefit Inventory. J Otolaryngol. 2003; 32:323–327. [PubMed: 14974864]

110. Joore MA, Brunenberg DE, Chenault MN, Anteunis LJ. Societal effects of hearing aid fitting among the moderately hearing impaired. Int J Audiol. 2003; 42:152–160. [PubMed: 12705780]

111. Uyama K, Takahashi M, Saito A, Okada Y, Tomizawa I, Kanzaki J. Questionnaire evaluation of balance in the performance of everyday activities after acoustic neuroma surgery. Acta Otolaryngol Suppl. 1991; 487:91–98. [PubMed: 1843593]

112. Myrseth E, Moller P, Wentzel-Larsen T, Goplen F, Lund-Johansen M. Untreated vestibular schwannomas: vertigo is a powerful predictor for health-related quality of life. Neurosurgery. 2006; 59:67–76. discussion 67–76. [PubMed: 16823302]

113. McCombe A, Baguley D, Coles R, McKenna L, McKinney C, Windle-Taylor P. Guidelines for the grading of tinnitus severity: the results of a working group commissioned by the British Association of Otolaryngologists, Head and Neck Surgeons, 1999. Clin Otolaryngol Allied Sci. 2001; 26:388–393. [PubMed: 11678946]

114. Andersson G, Ekvall L, Kinnefors A, Nyberg G, Rask-Andersen H. Evaluation of quality of life and symptoms after translabyrinthine acoustic neuroma surgery. Am J Otol. 1997; 18:421–426. [PubMed: 9233480]

115. Rigby PL, Shah SB, Jackler RK, Chung JH, Cooke DD. Acoustic neuroma surgery: outcome analysis of patient-perceived disability. Am J Otol. 1997; 18:427–435. [PubMed: 9233481]

116. House JW, Brackmann DE. Facial nerve grading system. Otolaryngol Head Neck Surg. 1985; 93:146–147. [PubMed: 3921901]

117. Lassaletta L, Alfonso C, Del Rio L, Roda JM, Gavilan J. Impact of facial dysfunction on quality of life after vestibular schwannoma surgery. Ann Otol Rhinol Laryngol. 2006; 115:694–698. [PubMed: 17044542]

118. Ryzenman JM, Pensak ML, Tew JM Jr. Patient perception of comorbid conditions after acoustic neuroma management: survey results from the acoustic neuroma association. Laryngoscope. 2004; 114:814–820. [PubMed: 15126736]

119. Tufarelli D, Meli A, Alesii A, De Angelis E, Badaracco C, Falcioni M, Sanna M. Quality of life after acoustic neuroma surgery. Otol Neurotol. 2006; 27:403–409. [PubMed: 16639281]

120. Fisher LM, Doherty JK, Lev MH, Slattery WH III. Distribution of nonvestibular cranial nerve schwannomas in neurofibromatosis 2. Otol Neurotol. 2007; 28:1083–1090. [PubMed: 18043434]

121. Gardner G, Robertson JH. Hearing preservation in unilateral acoustic neuroma surgery. Ann Otol Rhinol Laryngol. 1988; 97:55–66. [PubMed: 3277525]

122. Robinson K, Gatehouse S, Browning GG. Measuring patient benefit from otorhinolaryngological surgery and therapy. Ann Otol Rhinol Laryngol. 1996; 105:415–422. [PubMed: 8638891]

123. Hawthorne G, Hogan A. Measuring disability-specific patient benefit in cochlear implant programs: developing a short form of the Glasgow Health Status Inventory, the Hearing Participation Scale. Int J Audiol. 2002; 41:535–544. [PubMed: 12477174]

124. Stewart MG. Outcomes and patient-based hearing status in conductive hearing loss. Laryngoscope. 2001; 111:1–21. [PubMed: 11802001]

125. Zhao F, Bai Z, Stephens D. The relationship between changes in self-rated quality of life after cochlear implantation and changes in individual complaints. Clin Otolaryngol. 2008; 33:427–434. [PubMed: 18983375]

126. Mahmud MR, Khan AM, Nadol JB Jr. Histopathology of the inner ear in unoperated acoustic neuroma. Ann Otol Rhinol Laryngol. 2003; 112:979–986. [PubMed: 14653368]

127. Caye-Thomasen P, Dethloff T, Hansen S, Stangerup SE, Thomsen J. Hearing in patients with intracanalicular vestibular schwannomas. Audiol Neurootol. 2007; 12:1–12. [PubMed: 17033159]

128. Wei BP, Mubiru S, O'Leary S. Steroids for idiopathic sudden sensorineural hearing loss. Cochrane Database Syst Rev. 2006:CD003998. [PubMed: 16437471]

129. Pollock BE. Management of vestibular schwannomas that enlarge after stereotactic radiosurgery: treatment recommendations based on a 15 year experience. Neurosurgery. 2006; 58:241–248. discussion 241–248. [PubMed: 16462477]

130. Hasegawa T, Kida Y, Yoshimoto M, Koike J, Goto K. Evaluation of tumor expansion after stereotactic radiosurgery in patients harboring vestibular schwannomas. Neurosurgery. 2006; 58:1119–1128. discussion 1119–1128. [PubMed: 16723891]

131. Delsanti C, Tamura M, Galanaud D, Regis J. Changing radiological results, pitfalls and criteria of failure. Neurochirurgie. 2004; 50:312–319. [PubMed: 15179284]

132. Nakamura H, Jokura H, Takahashi K, Boku N, Akabane A, Yoshimoto T. Serial follow-up MR imaging after gamma knife radiosurgery for vestibular schwannoma. AJNR Am J Neuroradiol. 2000; 21:1540–1546. [PubMed: 11003293]

133. Yang SY, Kim DG, Chung HT, Park SH, Paek SH, Jung HW. Evaluation of tumour response after gamma knife radiosurgery for residual vestibular schwannomas based on MRI morphological features. J Neurol Neurosurg Psychiatry. 2008; 79:431–436. [PubMed: 17673492]

134. Kondziolka D, Lunsford LD, McLaughlin MR, Flickinger JC. Long-term outcomes after radiosurgery for acoustic neuromas. N Engl J Med. 1998; 339:1426–1433. [PubMed: 9811917]

135. Meijer OW, Weijmans EJ, Knol DL, Slotman BJ, Barkhof F, Vandertop WP, Castelijns JA. Tumor-volume changes after radiosurgery for vestibular schwannoma: implications for follow-up MR imaging protocol. AJNR Am J Neuroradiol. 2008; 29:906–910. [PubMed: 18296549]

136. Brors D, Schafers M, Bodmer D, Draf W, Kahle G, Schick B. Postoperative magnetic resonance imaging findings after transtemporal and translabyrinthine vestibular schwannoma resection. Laryngoscope. 2003; 113:420–426. [PubMed: 12616190]

137. Weissman JL, Hirsch BE, Fukui MB, Rudy TE. The evolving MR appearance of structures in the internal auditory canal after removal of an acoustic neuroma. AJNR Am J Neuroradiol. 1997; 18:313–323. [PubMed: 9111669]

138. Smith M, Castillo M, Campbell J, Pillsbury H, Walters T. Baseline and follow-up MRI of the internal auditory canal after suboccipital resection of acoustic schwannoma: appearances and clinical correlations. Neuroradiology. 1995; 37:317–320. [PubMed: 7666970]

139. Bennett ML, Jackson CG, Kaufmann R, Warren F. Postoperative imaging of vestibular schwannomas. Otolaryngol Head Neck Surg. 2008; 138:667–671. [PubMed: 18439476]

140. Umezu H, Seki Y. Postoperative magnetic resonance imaging after acoustic neuroma surgery: influence of packing materials in the drilled internal auditory canal on assessment of residual tumor. Neurol Med Chir (Tokyo). 1999; 39:141–147. discussion 147–149. [PubMed: 10193147]

141. Hirsh IJ, Davis H, Silverman SR, Reynolds EG, Eldert E, Benson RW. Development of materials for speech audiometry. J Speech Hear Disord. 1952; 17:321–337. [PubMed: 13053556]

142. Tillman, T.; Carhart, R. Northwestern University Auditory test #6; Technical Report SAM-TR-66-55. Brooks Air Force Base, TX: USAF School of Aerospace Medicine; 1966. An expanded test for speech discrimination utilizing CNC monosyllabic words.

143. Causey GD, Hood LJ, Hermanson CL, Bowling LS. The Maryland CNC Test: normative studies. Audiology. 1984; 23:552–568. [PubMed: 6517748]

144. De Cardenas, M.; Marrero, V. Cuaderno de logoaudiometria. Madrid: Universidad Nacional de Educacion a Distancia; 1994.

145. Halpin, C. Measuring audiometric outcomes. In: Shin, J.; Hartnick, G.; Randolph, G., editors. Evidence based otolaryngology. New York: Springer; 2008. p. 227-238.

146. Monsell EM. New and revised reporting guidelines from the Committee on Hearing and Equilibrium. American Academy of Otolaryngology-Head and Neck Surgery Foundation, Inc. Otolaryngol Head Neck Surg. 1995; 113:176–178.

147. Thornton AR, Raffin MJ. Speech-discrimination scores modeled as a binomial variable. J Speech Hear Res. 1978; 21:507–518. [PubMed: 713519]

148. Halpin C, Rauch SD. Using audiometric thresholds and word recognition in a treatment study. Otol Neurotol. 2006; 27:110–116. [PubMed: 16371857]

149. Aronzon A, Ruckenstein MJ, Bigelow DC. The efficacy of corticosteroids in restoring hearing in patients undergoing conservative management of acoustic neuromas. Otol Neurotol. 2003; 24:465–468. [PubMed: 12806300]

150. Nageris BI, Popovtzer A. Acoustic neuroma in patients with completely resolved sudden hearing loss. Ann Otol Rhinol Laryngol. 2003; 112:395–397. [PubMed: 12784975]

151. Gaffney RJ, McShane DP. Bilateral acoustic neurofibromatosis camouflaged by corticosteroid treatment of sudden sensorineural hearing loss. Ir J Med Sci. 1996; 165:151–152. [PubMed: 8824013]

152. Riccardi VM. Mast-cell stabilization to decrease neurofibroma growth. Preliminary experience with ketotifen. Arch Dermatol. 1987; 123:1011–1016. [PubMed: 3115189]

153. Jahrsdoerfer RA, Benjamin RS. Chemotherapy of bilateral acoustic neuromas. Otolaryngol Head Neck Surg. 1988; 98:273–282. [PubMed: 3132678]

154. El-Maraghi RH, Eisenhauer EA. Review of phase II trial designs used in studies of molecular targeted agents: outcomes and predictors of success in phase III. J Clin Oncol. 2008; 26:1346–1354. [PubMed: 18285606]
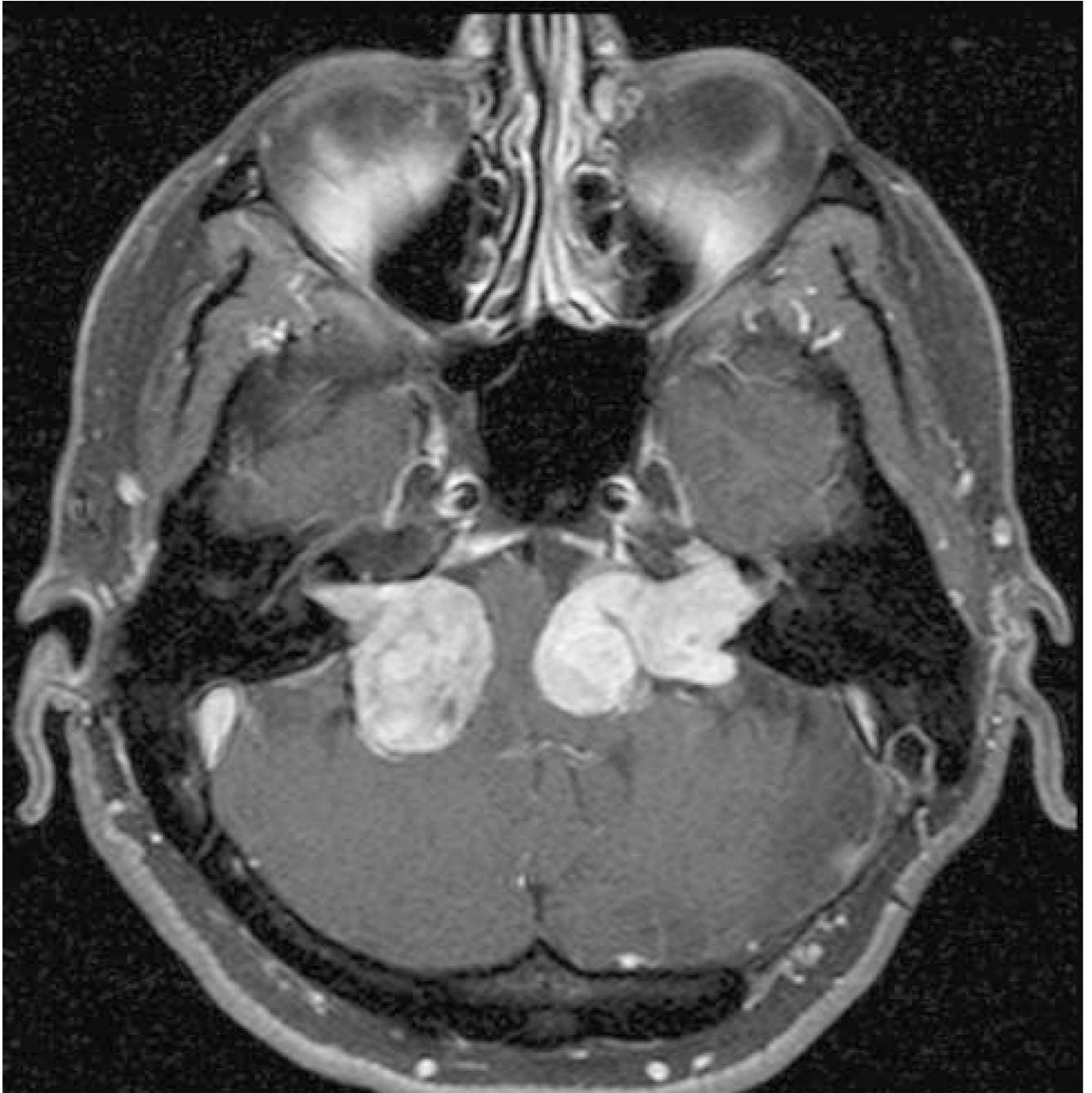
Fig. 1.
Axial gadolinium enhanced T1 weighted MRI scan showing multiple confluent cerebellopontine angle tumors in an NF2 patient
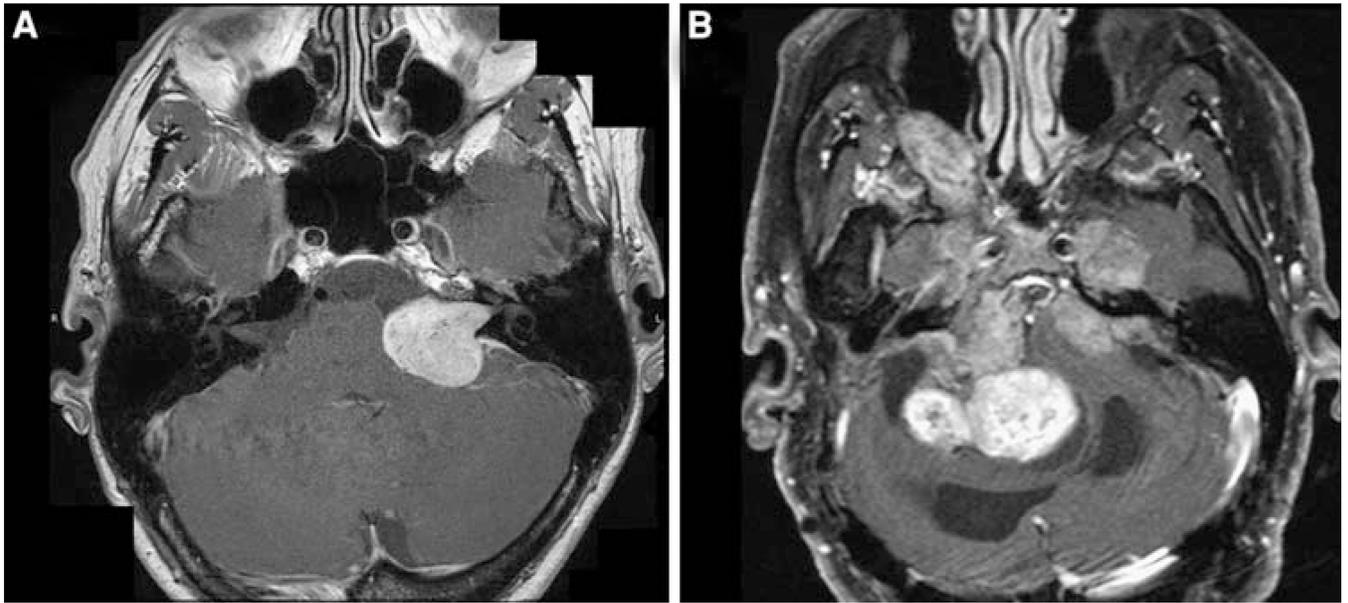
Fig. 2.
Axial gadolinium enhanced T1 weighted MRI scans. **a** Sporadic unilateral VS, showing regular ovoid shape. **b** NF2-related VS showing highly irregular shape
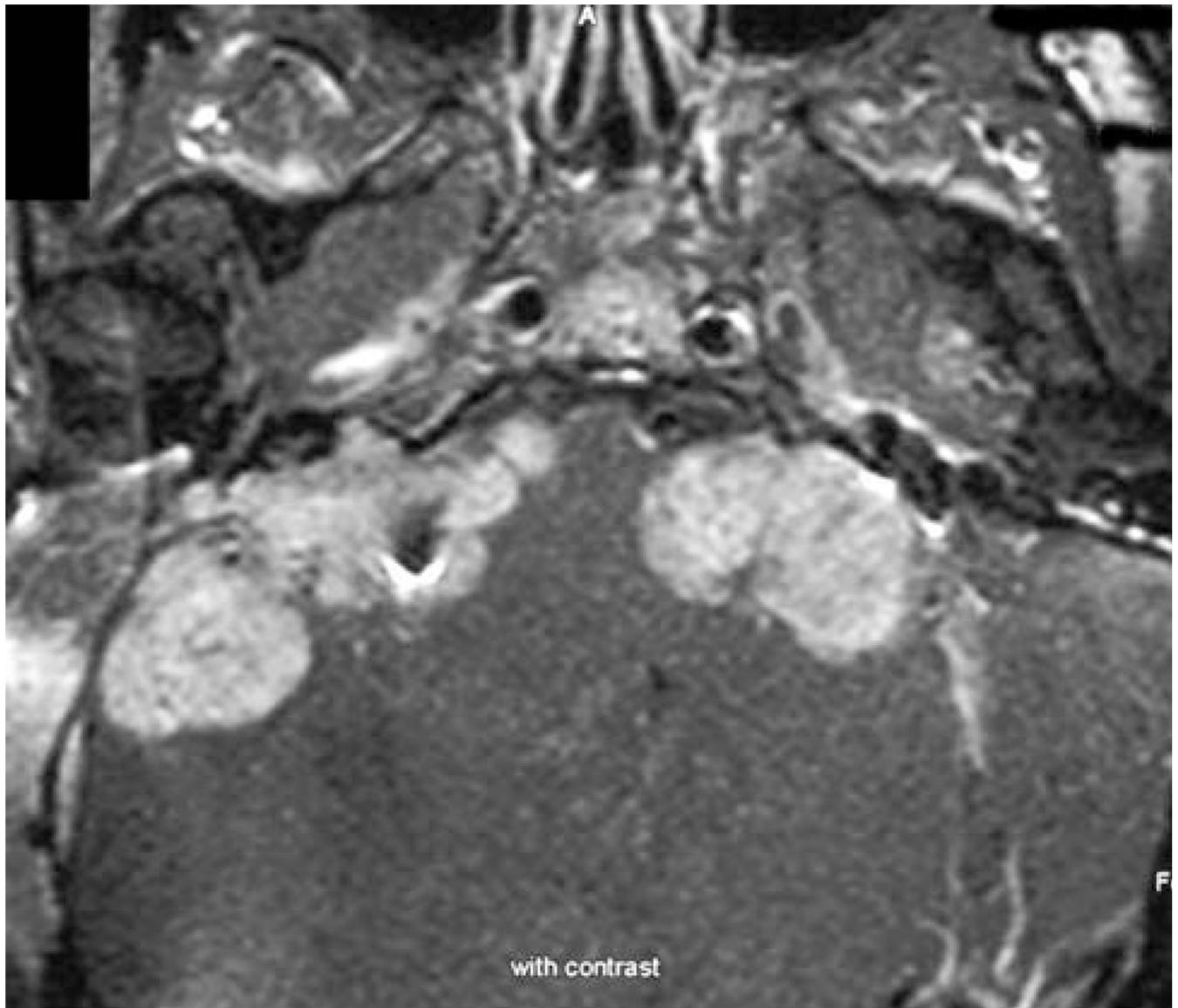
Fig. 3.
Axial gadolinium enhanced T1 weighted MRI scan showing NF2 related VS with multiple abutting tumors, probably both meningiomas and schwannomas of other cranial nerves. Such a tumor would be nonmeasurable by MRI volumetric technique
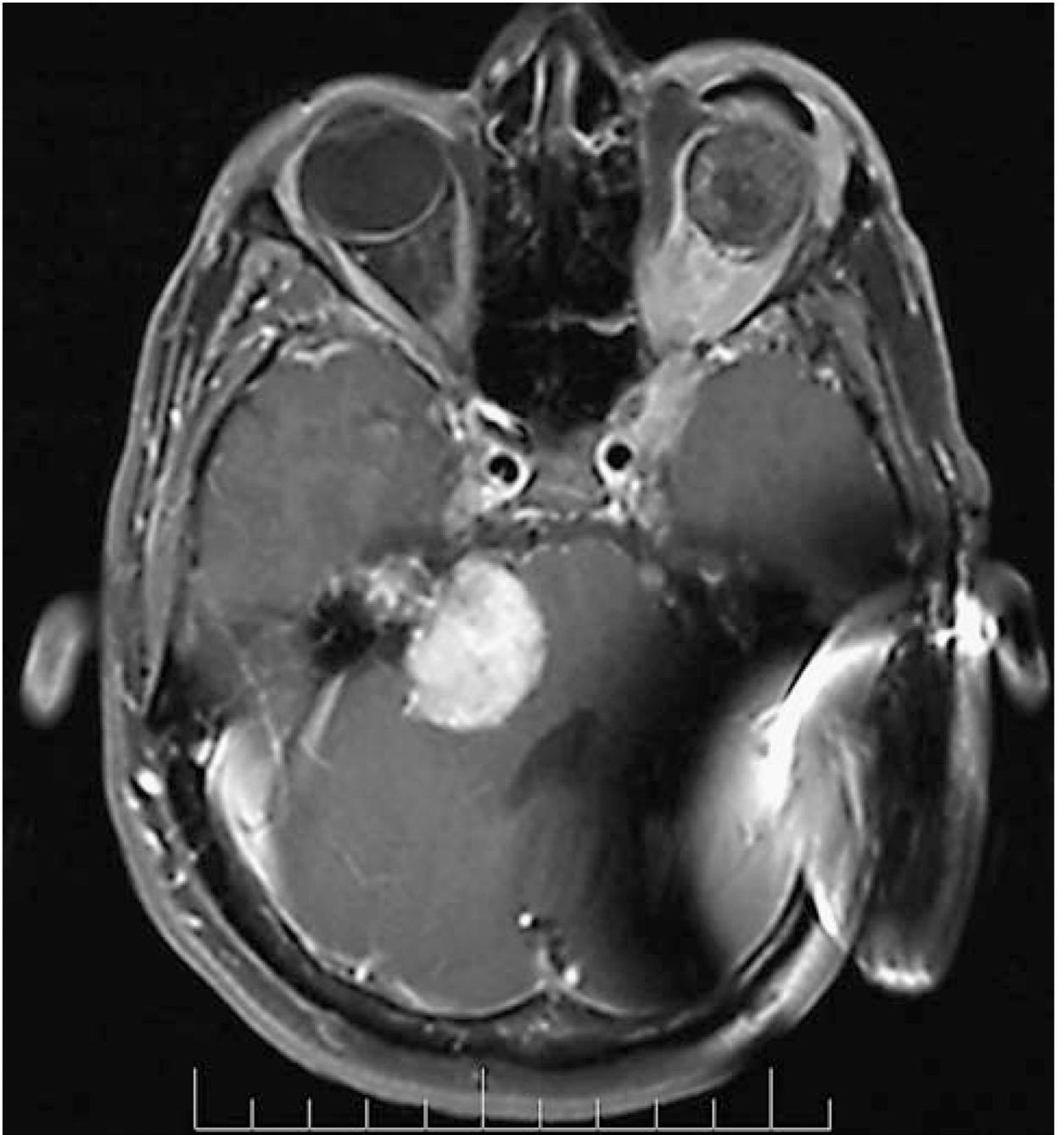
Fig. 4.
Axial gadolinium enhanced T1 weighted MRI scan showing imaging artifact from an auditory brainstem implant device, preventing accurate MRI volumetric measurement of the ipsilateral VS volume

Table 1

Clinical criteria for definition of hearing response based on a 50-word hearing test

Baseline word recognition score (%)	95% critical difference (%)	Hearing response (%)	Progressive disease (%)
0	0–4	6	n/a
2	0–10	12	n/a
4	0–14	16	n/a
6	2–18	20	0
8	2–22	24	0
10	2–24	26	0
12	4–26	28	2
14	4–30	32	2
16	6–32	34	4
18	6–34	36	4
20	8–36	38	6
22	8–40	42	6
24	10–42	44	8
26	12–44	46	10
28	14–46	48	12
30	14–48	50	12
32	16–50	52	14
34	18–52	54	16
36	20–54	56	18
38	22–56	58	20
40	22–58	60	20
42	24–60	62	22
44	26–62	64	24
46	28–64	66	26
48	30–66	68	28
50	32–68	70	30
52	34–70	72	32
54	36–72	74	34
56	38–74	76	36
58	40–76	78	38
60	42–78	80	40
62	44–78	80	42
64	46–80	82	44
66	48–82	84	46
68	50–84	86	48
70	52–86	88	50
72	54–86	88	52
74	56–88	90	54

Baseline word recognition score (%)	95% critical difference (%)	Hearing response (%)	Progressive disease (%)
76	58–90	92	56
78	60–92	94	58
80	64–92	94	62
82	66–94	96	64
84	68–94	96	66
86	70–96	98	68
88	74–96	98	72
90	76–98	100	74
92	78–98	100	76
94	82–98	100	80
96	86–100	n/a	84
98	90–100	n/a	88
100	96–100	n/a	94

Upper and lower limits for the 95% critical differences for percentage scores (second column). Changes in word recognition within the 95% critical difference are not statistically different from the baseline score. In use, third (for improved hearing) or fourth (for decreased hearing) columns are compared to baseline scores in first column. Adapted from Thornton and Raffin [147]

Table 2

Composite response criteria (combining radiographic and hearing response) for phase II studies of NF2-related VS

Radiographic response	Hearing response			
	Major	**Minor**	**Stable**	**Progressive**
Major	Major response	Minor response	Minor response	Progressive disease
Minor	Minor response	Minor response	Stable disease	Progressive disease
Stable	Minor response	Stable disease	Stable disease	Progressive disease
Progressive	Progressive disease	Progressive disease	Progressive disease	Progressive disease

Major radiographic response is defined as 20% decrease in tumor volume compared to baseline; minor radiographic response is defined as 5–19% decrease in tumor volume compared to baseline. Major hearing response is defined as a statistically significant increase in word discrimination score compared to baseline; minor hearing response is defined as a decrease in PTA 10 dB in the setting of stable word recognition scores