VESTIBULAR SCHWANNOMA

OPTIMIZING DIAGNOSIS AND FOLLOW-UP

MAYKE HENTSCHEL

VESTIBULAR SCHWANNOMA

OPTIMIZING DIAGNOSIS AND FOLLOW-UP

Mayke Hentschel

Vestibular Schwannoma – Optimizing diagnosis and follow-up
Mayke Hentschel
ISBN/EAN: 978-90-9032841-6

Cover design: Bregje Jaspers, ProefschriftOntwerp.nl
Lay-out: Bregje Jaspers, ProefschriftOntwerp.nl
Printing: Ipskamp Printing BV, Enschede

The work presented in this thesis was carried out within the Radboud Institute for Health Sciences.

The research was financially supported by ZonMw (the Netherlands Organization for Health Research and Development), project number 837002523.

Printing of this thesis was financially supported by ALK Abello BV, Beter Horen, Dos Medical BV/kno-winkel.nl, EmiD BV, MediTop B.V., Mylan Healthcare B.V., PENTAX Medical Nederland BV, Phonak, Radboud university medical center Nijmegen / Radboud University Nijmegen, Schoonenberg HoorSupport.

VESTIBULAR SCHWANNOMA

OPTIMIZING DIAGNOSIS AND FOLLOW-UP

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op woensdag 1 juli 2020
om 10.30 uur precies

door
Mayke Adriënne Hentschel
geboren op 3 maart 1988
te Leiden

Promotor
Prof. dr. M.M. Rovers

Copromotoren
Prof. dr. H.P.M. Kunst
Dr. S.C.A. Steens

Manuscriptcommissie
Prof. dr. ir. G.A. Zielhuis
Prof. dr. Y. Temel (Universiteit Maastricht)
Dr. E.F. Hensen (Leids Universitair Medisch Centrum)

CONTENTS

CHAPTER 1

Introduction

VESTIBULAR SCHWANNOMAS

Vestibular schwannomas (VSs) are benign tumours that originate from the Schwann cells of the nerve sheath of the eighth cranial nerve; the vestibulocochlear nerve, which transports sensory information on hearing and equilibrium from the inner ear to the brain.[1,2] VSs can be found in the internal auditory canal and/or cerebellopontine angle (CPA) (Figure 1.1). The mean age of diagnosis is 56 years.[3] VSs are relatively rare, with an incidence that is currently estimated at 1.1-3.3 per 100,000 people per year.[3,4] Other types of lesions can also be found in the CPA, but these are more seldom.

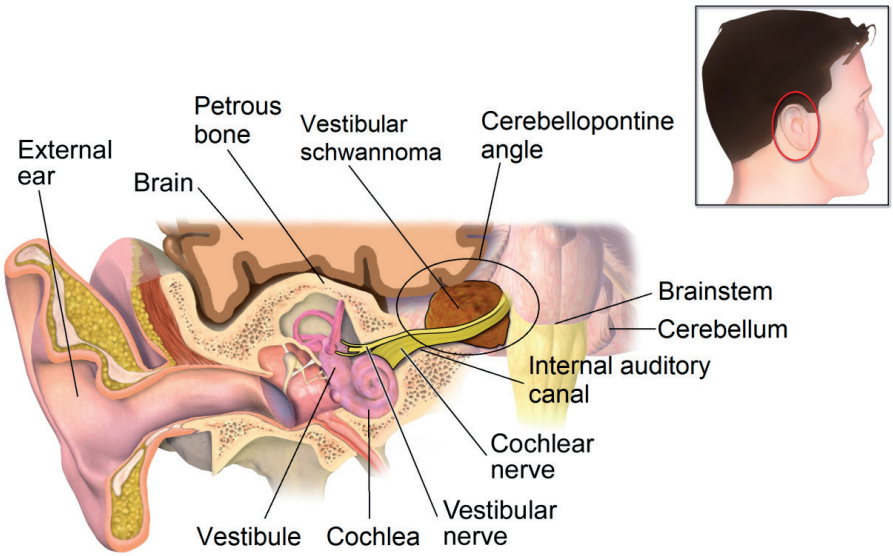


Figure 1.1 Vestibular schwannomas originate from the vestibulocochlear nerve and can be found in the internal auditory canal and/or cerebellopontine angle. Adapted from Blausen.com staff.[5]

SYMPTOMS

Presenting symptoms of patients with a VS usually exist of asymmetrical hearing loss, asymmetrical tinnitus or dizziness complaints (either vertigo or balance problems), i.e. asymmetrical audiovestibular complaints (AAC), which are the result of a malfunctioning vestibulocochlear nerve on the affected side. In a few cases, facial paresis, facial numbness or headaches are present caused by compression of surrounding cranial nerves.[6] AAC are common in every otolaryngology practice, but only in about 3% of patients a VS is the underlying cause.[3,7] Hence, it is a challenge to find out which patients with AAC are suffering from a VS, so the appropriate management strategy can be determined.

DIAGNOSIS

For years clinicians have been searching for ways to detect VSs in patients presenting with AAC. Decades ago, the acoustic neuroma suspicion index was developed to limit the need for, at that time, costly neuro-otological tests.[8] Since its introduction in the 80s Magnetic Resonance Imaging (MRI) has been considered the gold standard to diagnose VS, and both high resolution T2-weighted and contrast enhanced T1-weighted MRI are frequently used (Figure 1.2).[9]

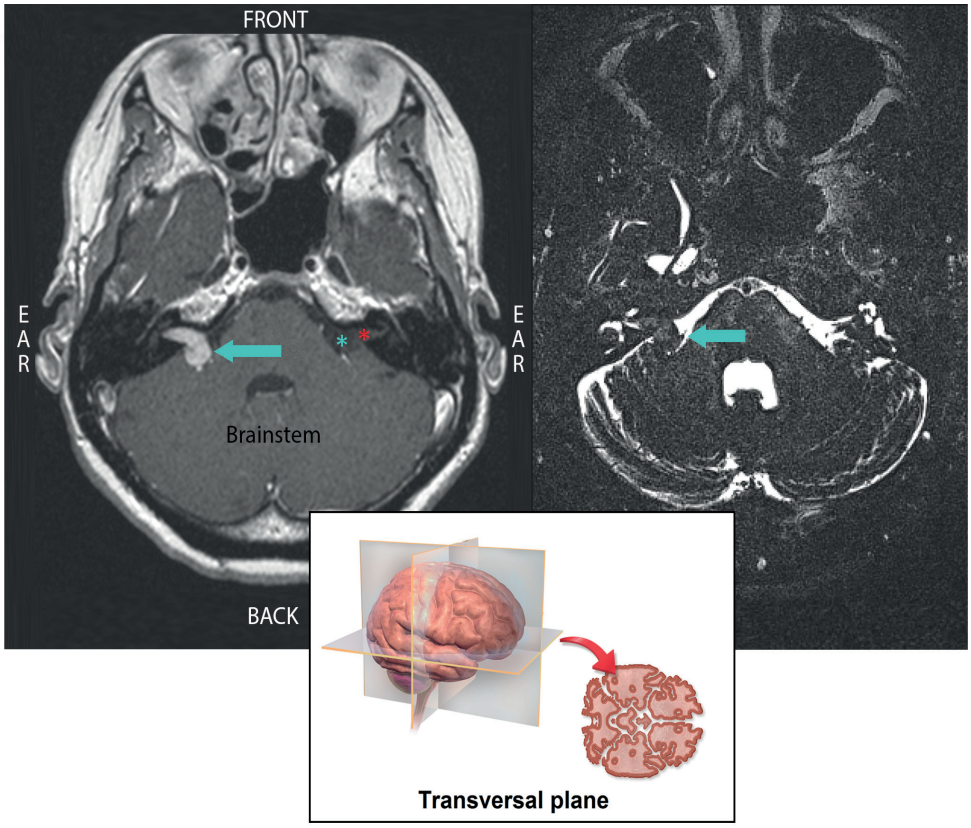


Figure 1.2 Example of a right-sided VS as seen on contrast-enhanced T1-weighted MRI (left) and high resolution T2-weighted MRI (right).

Images are transversal slices through the lower part of the head. The upper side of each image represents the face of the patient, whereas the lower part represents the back of the head. Arrows indicate the location of the VS. The red and green asterisk indicate the contralateral normal internal auditory canal and CPA (without VS), respectively. Image on sectional planes by Blausen.com staff.[5]

MRI, however, is an expensive investigation and has a yield of only 3% in patients with AAC.[3, 7, 10] To decrease the number of negative MRI examinations, clinicians are searching for selection criteria to refer only those patients with a high suspicion of VS for MRI. To this end, numerous tests are available, of which pure-tone audiometry, speech audiometry, vestibular function tests (electronystagmography, video head-impulse-test), and auditory brainstem response are well-known examples. It is, however, unclear which of these can best be used as initial diagnostic tool to select patients for MRI. The current diagnostic strategy is summarized in Figure 1.3.

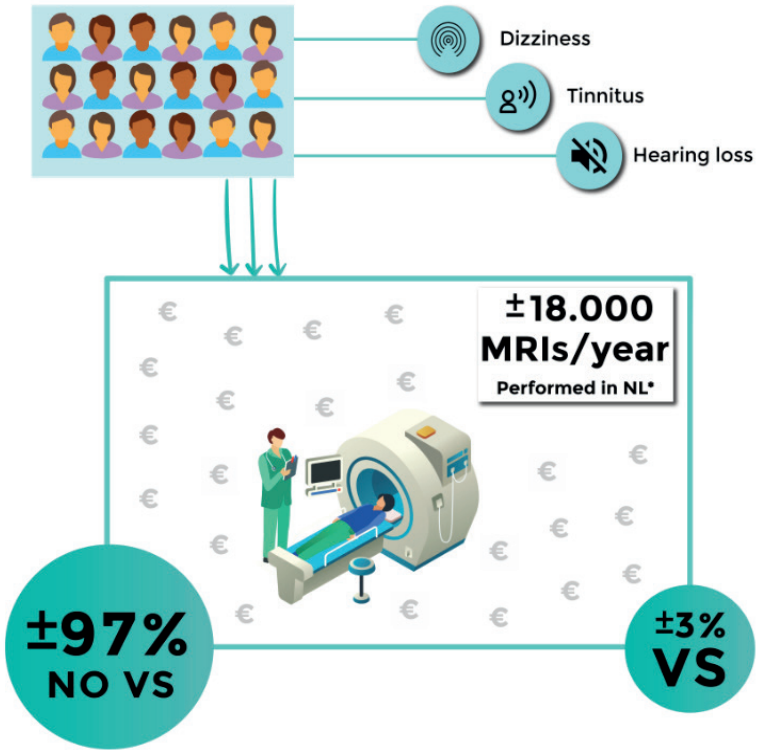


Figure 1.3 Schematic display of diagnostic strategy for VS.

*Calculation based on Kleijwegt et al.[3] MRI image adapted from Blausen Medical Communications, Inc.[11]

DIFFERENTIAL DIAGNOSES

VSs account for the majority (80%) of CPA lesions, followed by meningiomas.[12] More rare types of lesions that can be found in the CPA are arachnoid and epidermoid cysts, lipomas, cholesteatomas, schwannomas originating from the trigeminal or facial nerve, glomus tumours and metastases.[12] The most probable diagnosis is usually established based on MRI findings, since histopathology specimens are difficult to obtain, and surgery is increasingly less performed for most of the aforementioned lesions.

MANAGEMENT

There are several management strategies for VS, varying from a 'wait and scan' (W&S) strategy to active treatment which might consist of radiation therapy (fractionated radiotherapy or stereotactic radiosurgery) or microsurgery. An overview is provided in Figure 1.4.

Wait and scan strategy

MRI has enabled careful monitoring of VSs. As a result, the W&S strategy, in which lesions are being monitored with repeated MRIs over time, has gained popularity. In the Netherlands, this is the preferred strategy in most patients with a newly diagnosed VS.[3] The W&S strategy consists of MRIs acquired at increasing intervals (e.g. in Nijmegen 1, 2, 3, 5, 7, 9, 12, 15, and 20 years after diagnosis, then continuing every 5 years). Growth, usually defined as an increase in tumour diameter of ≥2 mm, is detected in approximately 40-60% of VSs at some point during follow-up, the percentage that is eventually treated is slightly smaller.[13, 14] The possibility to detect even the smallest VSs has led to an increasing proportion of patients being followed in a W&S strategy. Hence, the majority of patients are subject to a lifelong W&S schedule, while their VS remains stable in size and thus untreated. They eventually die with, but not because of their VS.

Treatment

VSs are benign lesions that do not metastasize. However, in a minority of patients expansion of the mass can cause severe problems due to compression of the brainstem. Invasive treatments such as radiation therapy or microsurgery are more and more reserved for these larger lesions, and/or lesions that have been proven to grow during a W&S strategy. In the past decades, treatment goals for the latter patients have shifted from tumour control only, to preservation of function and minimization of morbidity.

Although the name implies it to be a surgical procedure, stereotactic radiosurgery is a highly precise technique of radiation therapy in which a single radiation dose is applied with minimal radiation of adjacent healthy tissues. Stereotactic radiosurgery aims to prevent further growth of the lesion and succeeds to do so in the majority of patients.[15] It can be used for lesions up to 3 cm in size and is increasingly preferred whenever there is an indication for treatment. Fractionated radiotherapy seems to

achieve similar tumour control rates compared to stereotactic radiosurgery, but is less often performed.[16] Symptoms of AAC usually do not improve following radiation therapy and can even be negatively affected by it.[15, 17, 18]

Microsurgery is usually applied to larger size lesions in which severe brainstem compression and/or oedema is present, since radiation therapy could induce further swelling, oedema, and progressive brainstem compression.[19] Multiple surgical approaches can be used (middle fossa, retrosigmoid or translabyrinthine), depending on preference of the surgical team, the patient's hearing level, and tumour size and localization.[20, 21]

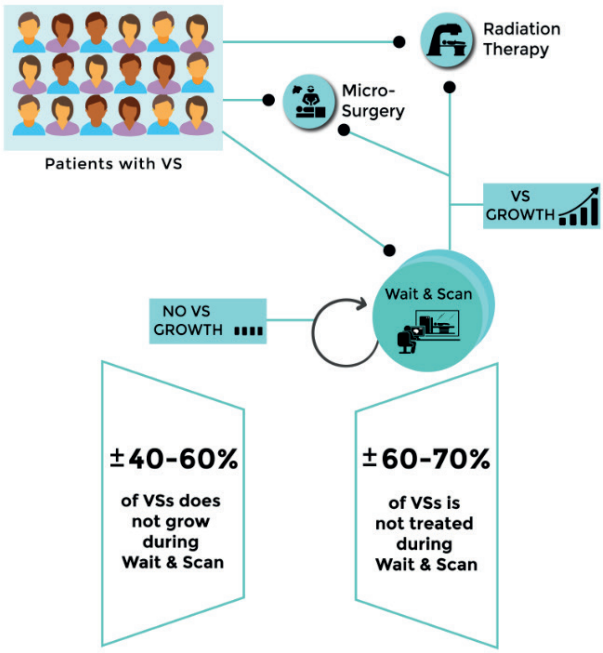


Figure 1.4 Schematic display of vestibular schwannoma management. *Data based on Hunter et al.[13]

(COST-)EFFECTIVENESS

The majority of MRIs to screen for a VS in patients with AAC are 'negative', meaning that they do not reveal any findings that might explain these symptoms. The amount of MRIs could be reduced by improving patient selection prior to MRI, i.e. from all patients presenting with AAC, only those with a high suspicion of VS should be referred for MRI. Once diagnosed, the majority of VS patients undergo lifelong MRIs with, again, repeated 'negative' results, because their VS remains stable in size without

indication for treatment. Subsequently, the amount of MRIs during this follow-up could be reduced by early selection of patients with a high risk of future VS growth. The latter patients should be monitored closely compared to patients with a low risk of future VS growth.

Geographical location seems to influence management of VS.[22] There is a lack of an (inter)national guideline on diagnostic and management strategies for VS. Strategies currently seem to be based on preference and experience of the consulted medical specialist. This may contribute to inconsistencies in diagnostic and managements strategies among specialists and may even cause unnecessary differences in outcome between patients as well as variation in costs.

OBJECTIVES

The overall aims of this thesis are to optimize 1) the (cost-)effectiveness of various diagnostic strategies in patients with AAC, i.e. suspected of a CPA lesion (including VS as most frequent lesion), and 2) the (cost-) effectiveness of the W&S strategy in patients diagnosed with a VS.

More specific objectives comprise:

- To explore what non-imaging strategies are currently being used to diagnose VS, and assess their diagnostic accuracy *(current practice and evidence)*
- To make an international comparison of diagnostic and management strategies for VS *(current practice and evidence)*

Diagnosis of vestibular schwannoma

- To assess potential savings in diagnosis of VS
- To compare the diagnostic accuracy of a limited MRI protocol to the current protocol
- To develop a new strategy to select patients suspected of VS, for MRI

Management of vestibular schwannoma

- To assess quality of life according to tumour size and management strategy in patients with VS
- To model cost-effectiveness of different W&S strategies
- To predict VS growth for patients assigned to a W&S strategy

OUTLINE OF THIS THESIS

In chapter 2, we provide an overview of current diagnostic strategies by performing a systematic review and meta-analysis regarding the diagnostic accuracy of existing non-imaging methods to screen for CPA lesions. Next, we study international practice variations in diagnosis and treatment of VS in chapter 3. The potential savings in diagnosis of VS are assessed in chapter 4 by means of a modelling study.

Chapter 5 and 6 describe the development of new diagnostic strategies, focusing on a limited MRI sequence and a diagnostic rule that can be used to select patients for MRI, respectively.

Quality of life in patients with VS according to lesion size and management strategy is studied in chapter 7, which is important as this information is required for cost-effectiveness studies. Chapter 8 describes a decision analytical modelling study to assess cost-effectiveness of different W&S strategies. The development of a prediction model that can be used to predict VS growth for patients assigned to a W&S strategy is described in chapter 9. A general discussion and summary complete the thesis.

REFERENCES

1. Murray MR, Stout AP. Demonstration of the Formation of Reticulin by Schwannian Tumor Cells in Vitro. 1942;18(4):585-93.
2. Huizing EHS, G.B.; de Vries, N.; Graamans, K.; van de Heyning P. Keel-neus-oorheelkunde en hoofd-halschirurgie. Houten.: Bohn Stafleu van Loghum; 2007.
3. Kleijwegt M, Ho V, Visser O, Godefroy W, van der Mey A. Real Incidence of Vestibular Schwannoma? Estimations From a National Registry. 2016;37(9):1411-7.
4. Kshettry VR, Hsieh JK, Ostrom QT, Kruchko C, Barnholtz-Sloan JS. Incidence of vestibular schwannomas in the United States. 2015;124(2):223-8.
5. Medical gallery of Blausen Medical 2014. Blausen. com staff: WikiJournal of Medicine 1 (2). 2014.
6. Broomfield SJ, O'Donoghue GM. Self-reported symptoms and patient experience: A British Acoustic Neuroma Association survey. 2016;30(3):294-301.
7. Zealley IA, Cooper RC, Clifford KM, Campbell RS, Potterton AJ, Zammit-Maempel I, et al. MRI screening for acoustic neuroma: a comparison of fast spin echo and contrast enhanced imaging in 1233 patients. 2000;73(867):242-7.
8. Windle-Taylor PC, Dhillon RS, Kenyon GS, Morrison AW. Acoustic Neuroma Suspicion Index: an aid to investigation and diagnosis. 1984;94(11 Pt 1):1464-7.
9. Fortnum H, O'Neill C, Taylor R, Lenthall R, Nikolopoulos T, Lightfoot G, et al. The role of magnetic resonance imaging in the identification of suspected acoustic neuroma: a systematic review of clinical and cost effectiveness and natural history. 2009;13(18):iii-iv, ix-xi, 1-154.
10. Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg (verdiepingsmodules) [Internet]. Zorginstituut Nederland. 2016 [cited August 22, 2017]. Available from: https://www.zorginstituutnederland.nl/over-ons/publicaties/publicatie/2016/02/29/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg.
11. Blausen 0654 MRI 01. Blausen Medical Communications, Inc.: Wikimedia Commons; 2017.
12. Moffat DA, Ballagh RH. Rare tumours of the cerebellopontine angle. 1995;7(1):28-41.
13. Hunter JB, Francis DO, O'Connell BP, Kabagambe EK, Bennett ML, Wanna GB, et al. Single Institutional Experience With Observing 564 Vestibular Schwannomas: Factors Associated With Tumor Growth. 2016;37(10):1630-6.
14. Hajioff D, Raut VV, Walsh RM, Bath AP, Bance ML, Guha A, et al. Conservative management of vestibular schwannomas: third review of a 10-year prospective study. 2008;33(3):255-9.
15. Arthurs BJ, Fairbanks RK, Demakas JJ, Lamoreaux WT, Giddings NA, Mackay AR, et al. A review of treatment modalities for vestibular schwannoma. 2011;34(3):265-77.
16. Persson O, Bartek J, Jr., Shalom NB, Wangerid T, Jakola AS, Forander P. Stereotactic radiosurgery vs. fractionated radiotherapy for tumor control in vestibular schwannoma patients: a systematic review. 2017;159(6):1013-21.
17. Pollock BE, Driscoll CL, Foote RL, Link MJ, Gorman DA, Bauch CD, et al. Patient outcomes after vestibular schwannoma management: a prospective comparison of microsurgical resection and stereotactic radiosurgery. 2006;59(1):77-85.

18. Cheng S, Naidoo Y, Da Cruz M, Dexter M. Quality of life in postoperative vestibular schwannoma patients. 2009;119:2252-7.
19. Meijer OW, Weijmans EJ, Knol DL, Slotman BJ, Barkhof F, Vandertop WP, et al. Tumor-volume changes after radiosurgery for vestibular schwannoma: implications for follow-up MR imaging protocol. 2008;29(5):906-10.
20. Chamoun R, MacDonald J, Shelton C, Couldwell WT. Surgical approaches for resection of vestibular schwannomas: translabyrinthine, retrosigmoid, and middle fossa approaches. 2012;33(3):E9.
21. Bennett M, Haynes DS. Surgical approaches and complications in the removal of vestibular schwannomas. 2007;40(3):589-609.
22. Carlson ML, Glasgow AE, Grossardt BR, Habermann EB, Link MJ. Does where you live influence how your vestibular schwannoma is managed? Examining geographical differences in vestibular schwannoma treatment across the United States. 2016;129(2):269-79.

PART I

Current practice and evidence

CHAPTER 2

The diagnostic accuracy of non-imaging screening protocols for vestibular schwannoma in patients with asymmetrical hearing loss and/or unilateral audiovestibular dysfunction: a diagnostic review and meta-analysis

Hentschel M.
Scholte M.
Steens S.
Kunst H.
Rovers M.

Published as:
Clin Otolaryngol. 2017;42(4):815-823. doi: 10.1111/coa.12788.

ABSTRACT

Background

Currently, all patients presenting with asymmetrical sensorineural hearing loss and/or unilateral audiovestibular dysfunction (i.e. tinnitus, dizziness) undergo MRI (magnetic resonance imaging), leading to a substantial amount of MRIs with negative findings as the incidence of vestibular schwannoma in this screening population varies between 1 and 4.7% (i.e. more than 95% of MRIs are negative for vestibular schwannoma).

Objective of review

The aim was to assess the diagnostic accuracy of different non-imaging screening protocols that can be used prior to MRI to select patients at high risk of vestibular schwannoma.

Type of review

Diagnostic review and meta-analysis.

Search Strategy

We systematically searched MEDLINE, Embase and The Cochrane Library as from inception up to 28 July, 2016. We included studies that compared non-imaging screening protocols to MRI as gold reference standard.

Evaluation method

Methodological quality was assessed by two independent reviewers using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool. Data necessary to complete 2×2 tables were obtained, and patient, study, screening and imaging characteristics were extracted. We calculated sensitivity and specificity of all tests and obtained pooled estimates using a bivariate random effects model.

Results

We analysed 12 studies (4969 patients) of poor to moderate quality according to the quality assessment. Most studies tested diagnostic accuracy of multiple screening protocols. Five pure-tone audiometry protocols were studied by multiple authors; pooled estimates for sensitivity ranged from 88% [95% CI 84-91] to 91% [95% CI: 52-99] and specificity from 31% [95% CI: 10-66] to 58% [95% CI: 49-65]. Due to heterogeneity we were unable to pool other tests. In five studies testing auditory brainstem response, sensitivity values ranged from 37% [95% CI: 23-52] to 100% [95% CI: 40-100] and specificity from 57% to 96% [95% CI: 87-100]. Two authors studied pure-tone audiometry shape as a screening test. Presenting symptoms, electronystagmography, caloric irrigation and hyperventilation test were assessed by one study each. All reported low diagnostic accuracy.

Conclusions

All identified studies had a moderate to high risk of bias and none of the currently available non-imaging screening protocols appear to be accurate in detecting vestibular schwannomas.

INTRODUCTION

The incidence of vestibular schwannoma (VS) is estimated at 0.3 to 1.9 per 100 000 individuals per year. [1-4] The majority of patients with a VS present with slowly progressing symptoms, such as asymmetrical hearing loss, vestibular complaints and/or asymmetrical tinnitus. [5, 6] These complaints are common and not specific for VS patients. If pure-tone audiometry (PTA) confirms asymmetrical sensorineural hearing loss, electronystagmography reveals vestibular asymmetry, and/or the patient reports asymmetrical tinnitus, a magnetic resonance imaging (MRI) examination will follow to exclude the presence of VS.

Currently, T1-weighted MRI using gadolinium-based contrast (GdT1-MRI) is the gold standard in the diagnostic work up of these patients. However, some hospitals also perform T2-weighted MRI (T2-MRI), either as additional examination next to GdT1-MRI, but more frequently replacing GdT1-MRI. Whenever sufficient visibility of the cochlea, labyrinth and the VIIth and VIIIth cranial nerve in the cerebellopontine angle and internal auditory canal is achieved by high-resolution T2-MRI, it probably has comparable diagnostic accuracy to GdT1-MRI in detecting VS. [5]

Currently, all patients presenting with asymmetrical sensorineural hearing loss and/or unilateral audiovestibular dysfunction (i.e. tinnitus, dizziness) undergo MRI, leading to a substantial amount of MRIs with negative findings as the incidence of VS in this screening population varies between 1 and 4.7% (i.e. more than 95% of MRIs are negative for VS). [7-10] Based on Dutch cost data, a screening MRI costs €206 per patient at risk of VS. [4, 11] If we could reduce the amount of patients needing MRI by improving selection criteria (i.e. reduce the amount of MRIs with negative findings), this would lead to a reduction in unnecessary examinations and costs. Using a potential new diagnostic strategy, high risk patients would still undergo MRI to confirm the diagnosis. Based on MRI findings tumour location and size can be evaluated to determine further action.

The aim of this diagnostic review and meta-analysis was to assess the diagnostic accuracy of different non-imaging screening protocols that can be used to identify patients with VS.

METHODS

Searches

We systematically searched MEDLINE (OvidSP), Embase (OvidSP) and the Cochrane Library from inception up to 28 July, 2016 for studies testing non-imaging screening protocols to identify patients with VS. The search query combined synonyms for MRI, asymmetrical sensorineural hearing loss, unilateral audiovestibular dysfunction and VS (see Document S1 for the complete search strategy). We also performed a reference and related article search. Duplicate articles were manually filtered using

the bibliographic EndNote database, version X5 (Thomas Reuters, New York City, NY, USA). There was no limitation in publication year or status, nor in language.

Study selection

We included studies with a diagnostic study design, presenting original study data, both prospective and retrospective. They had to compare at least one non-imaging screening protocol to MRI findings, in patients with asymmetrical sensorineural hearing loss and/or unilateral audiovestibular dysfunction, considered at risk of VS. We included all types of screening protocols (e.g. using symptoms, PTA, or other diagnostic tests). Both GdT1-MRI and T2-MRI were used for this review. Furthermore, they had to provide sufficient data to construct a 2x2 contingency table. We excluded opinion articles, animal studies, (systematic) reviews, case reports, and studies using other imaging techniques than MRI as reference test. Two reviewers (MH and MS) independently assessed the eligibility of the identified articles. Any disagreements were resolved by discussion with a third reviewer (MR).

Data extraction and quality assessment

Two authors (MH and MS) extracted data independently using a predefined form, including data on patient, study, screening and imaging characteristics. Patient characteristics comprised age, gender and presenting symptoms. Study characteristics included sample size and study type. Screening characteristics comprised incidence of VS, index tests and target disease. Imaging characteristics included MRI field strength, sequence and assessment. We contacted authors of different articles to obtain absolute data not displayed in their publications, or for clarification (see Document S2 for an overview of missing data, assumptions and translations).

Next, we assessed the risk of bias and applicability of the studies using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) scoring system. [12] This is a validated tool for assessment of the methodological quality and applicability of diagnostic accuracy studies. Four domains are distinguished: (1) patient selection, which describes the process of patient selection and included patients; (2) index test, describing the test under study and how it was conducted and interpreted; (3) reference standard, describing the used reference standard and how it was conducted and interpreted; and (4) flow and timing, describing the flow of patient in- and exclusion and the time interval between index test and reference standard. Two independent reviewers (MH and MS) performed the quality assessment. Any disagreements were resolved by discussion with a third reviewer (MR).

Analysis

We summarized data from each study in 2×2 tables of true positive, false positive, true negative, and false negative values and calculated sensitivity and specificity values. Authors of studies that did not report all sufficient data were asked to provide additional information. To calculate sensitivity and specificity measurements at study level, we used Review Manager (RevMan) software, version 5.3 (The Nordic Cochrane Centre, The Cochrane Collaboration, 2014). To show variation and explore heterogeneity for

sensitivity and specificity, we drew forest plots. If possible we provided pooled estimates of sensitivity and specificity with 95% confidence intervals (95% CIs). We used R version 3.2.3 (The R Foundation for Statistical Computing, Vienna, Austria, 2015) using the mada package to carry out the meta-analyses. We wanted to study heterogeneity in more detail, by performing sensitivity analyses of clinically relevant covariates (age, VS size and location, and MRI field strength and sequence).

RESULTS

Literature search

Our search yielded 1788 unique records, of which 60 remained after screening titles and abstracts (see also Figure 2.1). We reviewed full-texts of these studies for eligibility and excluded 48. Studies were excluded because of their study design (case reports, editorials/comments, reviews, guidelines), the same data were used in another study, there was lack of a control group or reference test, or because they had a different domain, determinant or outcome. Finally, 12 studies were included, covering a total of 4969 patients. We did not retrieve additional items after screening references and related articles.

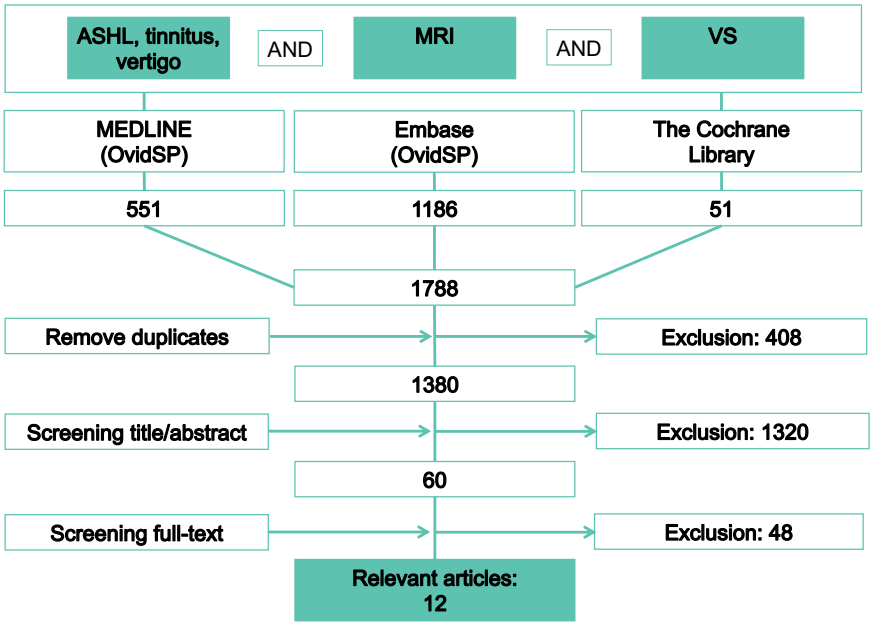


Figure 2.1 Flowchart of search and selection of studies.

Table 2.1 Study characteristics of included studies.

Study	N	Incidence of VS (n(%))	Study type	Symptoms	Target disease	Index test(s)	Reference test		
							Field strength (Tesla)	Sequence	Assessment by
Chatrath 2008	420	?	CC	• Asymmetric tinnitus • Subjective unilateral HL (no conductive HL) • ASHL ≥15 dB at 0.25-8 kHz • Symmetrical SHL ≥40 dB average at 0.5-2 kHz in patients ≤ 65 years • Episodic positional or spontaneous vertigo ≥ 6 months.	CPA lesion	• Prediction model	?	?	?
Cheng 2012	1751	131 (7.5)	C	• SHL	Acoustic tumour*	• 15 PTA protocols	?	T2 or GdT1	Consultant radiologist + 2nd assessment
Cueva 2004	312	31 (9.9)	NRT	• ASHL (≥15 dB in ≥2 frequencies, but ≤ 70 dB hearing loss between 2-4 kHz) • ≥15% asymmetry on SDS.	Causative lesion†	• ABR	?	GdT1	Neuroradiologist
Gimsing 2010	425	199 (46.8)	CC	• ASHL • Tinnitus • Vertigo	VS	• PTA shapes • 8 PTA protocols	?	?	?
Mandala 2013	102	49 (48.0)	CC	• Unilateral SHL	VS	• HVT • ABR • Caloric irrigation	?	GdT1	?
Obholzer 2004	128	36 (28.1)	CC	• AHL +/- tinnitus • Vertigo • Trigeminal paraesthesia	VS	• 9 PTA protocols	?	?	Radiologist
Raber 1997	48	?	C	• AHL, reliable waveforms on ABR • UAD	CPA tumour	• ABR	0.5 / 1.5	GdT1	?
Rupa 2003	90	4 (5.6)	C	• AHL (>15 dB at ≥2 frequencies at 0.25-8 kHz) • Asymmetric tinnitus	VS	• ABR	?	GdT1	?
Saliba 2011	212	84 (39.6)	C	• ASHL of ≥ 10 dB at any frequency • Asymmetry of ≥15% on SDS	VS	• 9 PTA protocols	?	?	Radiologist

Scherler 1995	100	50 (50.0)	CC	?	VS	• PTA protocol • ENG • ABR	?	GdT1	?
Suzuki 2010	500	13 (2.6)	C	• ASHL(>15 dB at any frequency 0.5- 4 kHz)	VS	• PTA shapes	1.5	T2	Radiologist / otolaryngologist
Vandervelde 2009	881	12 (1.4)	C	• AHL +/- AD • Central vestibular findings	VS	• Presenting symptoms	1.5	T2 +/- GdT1	Radiologist / neuroradiologist

*Acoustic tumour: Vestibular schwannoma and meningioma

†Causative lesions: 24 VS, 7 other lesions

?: Unclear/unknown, +/-: either with or without

ABR: Auditory brainstem response audiometry, (A)(S)HL: (asymmetric) (sensorineural) hearing loss, ENG: Electronystagmography, C: cohort, CC: case-control, CPA: cerebellopontine angle, GdT1: T1 with administration of gadolinium, HVT: Hyperventilation test, NRT: non-randomised trial, PTA: Pure-tone audiometry, SDS: Speech discrimination score, (U)AD: (unilateral) audiovestibular dysfunction, VS: vestibular schwannoma

Study characteristics

Table 2.1 provides an overview of the patients included in the different studies. Studies providing information on sex included 806 (50.7%) men and 784 (49.3%) women. [13-17] The mean age reported varied from 45 to 57 years. [13-17] The incidence of VS ranged from 1.4 to 39.6% in the included cohort studies. [13, 16-20] Most other studies were case-control studies. Several authors studied multiple screening protocols in their study population. [14, 15, 17, 18, 21, 22] Six studies used PTA as index test, comprising 3116 patients. [14, 17-19, 21, 22] Five of these [14, 17, 18, 21, 22] tested specific PTA protocols (hearing thresholds at different frequencies) that had been published before in 2616 patients, and two [14, 19] tested different PTA shapes (description of the PTA shape without providing specific thresholds, e.g. flat, mountain, etc.) using 925 patients. ABR was tested in five studies in a total of 652 patients. [13, 15, 16, 22, 23] Hyperventilation test, electronystagmography, caloric irrigation, presenting symptoms and a prediction model were tested in one study each, comprising 102, 100, 102, 881 and 420 patients, respectively. [15, 20, 22, 24]

Quality assessment

Overall, the quality of included studies was low to moderate (Figure 2.2). Nine studies found a high risk of bias in ≥1 domain. For the domains patient selection, index test, reference standard, and flow and timing, a high or unclear risk of bias was scored for 9, 11, 12 and 5 studies, respectively. The risk of bias in index and reference test categories was generally scored unclear, because there was uncertainty about blinding, and MRI sequence and/or field strength were not provided. For the domains patient selection, index test and reference standard, respectively, only 1, 1 and 2 studies scored high on concerns regarding applicability. Thus, for the majority of the articles there was low concern that applicability of the articles did not fit the review question. No studies were excluded from the analysis based on the quality assessment.

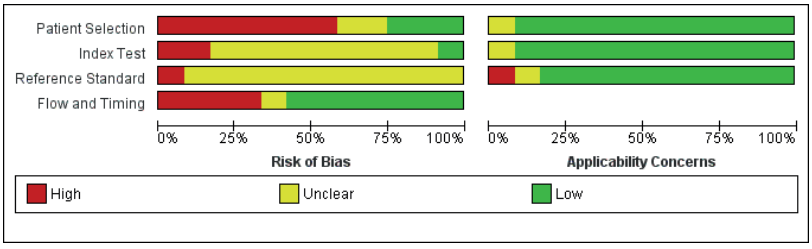


Figure 2.2 Overview QUADAS-2 results: Risk of bias and concerns regarding applicability.

Diagnostic accuracy of PTA, ABR, and other tests

Table 2.2 provides an overview of asymmetrical sensorineural hearing loss definitions of the different PTA protocols. Forest plots representing the diagnostic accuracy of each protocol as tested in the individual studies can be found in Supplemental Figure 2.1. Sensitivities were high for all PTA protocols. Specificities however, did not exceed 70%. Five PTA protocols were studied by multiple authors in studies with similar incidences of VS; pooled estimates for sensitivity and specificity are displayed in Table 2.3. [14, 17, 21] Due to heterogeneity, we were unable to pool other tests.

Table 2.2 Overview of pure-tone audiometry protocols.

Protocol name	Definition of ASHL	Studied by
AAO-HNS	Average ≥ 15 dB at frequencies 0.5- 3 kHz.	Cheng 2012, Obholzer 2004, Saliba 2011
AMCLASS	≥ 10 dB at ≥ 2 frequencies, or ≥ 15 dB at any single frequency.	Saliba 2011
AMCLASS-A-Urben	≥ 10 dB at two adjacent frequencies.	Cheng 2012
AMCLASS-B-Urben	≥ 15 dB at any frequency.	Cheng 2012, Obholzer 2004†
Cueva	≥ 15 dB at two adjacent frequencies.	Cheng 2012, Obholzer 2004†
	≥ 15 dB at ≥ 2 frequencies, or ≥ 15% asymmetry in SDS.	Saliba 2011
Dawes and Jeannon	> 19 dB at two adjacent frequencies, unilateral tinnitus, Ménière's disease or sudden deafness.	Gimsing 2010
DOH‡	≥ 20 dB at any frequency between 0.5- 4 kHz.	Cheng 2012, Obholzer 2004, Saliba 2011, Gimsing 2010
Mangham	Average ≥ 10 dB at frequencies 1- 8 kHz.	Cheng 2012, Gimsing 2010
Nashville	≥ 15 dB at any frequency between 0.5- 4 kHz.	Cheng 2012, Obholzer 2004, Saliba 2011*
Obholzer 1§	If better ear average < 31 dB hearing loss: > 15 dB at adjacent frequencies 0.25- 8 kHz. Otherwise: Average > 20 dB at adjacent frequencies 0.25- 8 kHz.	Gimsing 2010, Obholzer 2004†
Obholzer 2§	If better ear average ≤ 30 dB hearing loss: Average ≥ 15 dB at frequencies 0.25- 8 kHz. If better ear average > 30 dB hearing loss: Average ≥ 20 dB at frequencies 0.25- 8 kHz.	Cheng 2012, Obholzer 2004†
Oxford	Average ≥ 15 dB at frequencies 0.5- 8 kHz.	Cheng 2012, Obholzer 2004, Saliba 2011*
Rule 3000	≥ 15 dB at 3 kHz.	Cheng 2012, Saliba 2011
Rule 4000	≥ 20 dB at 4 kHz.	Cheng 2012
Scherler	≥ 25 dB at ≥ 2 frequencies.	Scherler 1995
Schlauch and Levine	Average ≥ 20 dB at frequencies 1- 8 kHz.	Cheng 2012
Seattle¶	Males: Average > 19 dB at frequencies 1- 8 kHz . Females: > 19 dB at 4 kHz.	Gimsing 2010
	Average ≥ 15 dB at frequencies 1- 8 kHz.	Cheng 2012, Gimsing 2010, Obholzer 2004, Saliba 2011
Sheppard	Average ≥ 15 dB at frequencies 0.25- 8 kHz.	Cheng 2012
Sunderland	Age ≤ 70 years: Average > 14 dB at frequencies 0.25- 8 kHz or normal hearing with unilateral tinnitus or canal paresis. (age >70 years not screened).	Gimsing 2010
	≥ 20 dB at two adjacent frequencies.	Cheng 2012, Obholzer 2004, Saliba 2011

Table 2.2 Continued

Welling	> 15 dB at any frequency between 0.5- 4 kHz, SDS asymmetry >20%, or unilateral tinnitus.	Gimsing 2010

* For unknown reason the results of Oxford and Nashville protocols are combined
† Does not use a name for the protocol, but uses identical criteria
‡ Same protocol as Nouraei
§ Obholzer protocol has a slightly different definition in the original articles, therefore we named them Obholzer 1 and 2
¶ Same protocol as Hunter
ASHL: Asymmetrical sensorineural hearing loss, SDS: Speech discrimination score

Table 2.3 Pooled estimates of five PTA protocols.

PTA protocol	**Pooled estimate**	
	Sensitivity [95% CI]	Specificity [95% CI]
AAO-HNS [17, 21]	0.909 [0.842-0.949]	0.575 [0.494-0.652]
DOH [14, 17, 21]	0.881 [0.839-0.913]	0.313 [0.097-0.661]
Seattle [14, 17, 21]	0.892 [0.799-0.945]	0.438 [0.258-0.636]
Obholzer 1 [14, 21]	0.911 [0.517- 0.990]	0.394 [0.234-0.58]
Sunderland [17, 21]	0.907 [0.442-0.992]	0.397 [0.05-0.892]

PTA: pure-tone audiometry, CI: confidence interval

Sensitivity of PTA shape as a screening test ranged from 0% [95% CI: 0-25] for low-frequency and mountain-shaped loss to 76% [95% CI: 69-82] for sloping loss and specificity from 29% [95% CI: 23-36] to 99% [95% CI: 96-100] for a sloping and peak loss, respectively. [14, 19]

The sensitivity and specificity values of the five studies that compared ABR results to MRI ranged from 37% [95% CI: 23-52] to 100% [95% CI: 40-100] and from 57% [95% CI: unknown] to 96% [95% CI: 87-100], respectively. Table 2.4 provides an overview of the results. [13, 15, 16, 22, 23]

Table 2.4 Overview and results of auditory brainstem response.

Study	Definition of positive ABR	No response on ABR	Sensitivity [95% CI]	Specificity [95% CI]
Cueva 2004	Interaural latency difference of wave V > 0.2 ms, abnormal absolute wave V latency, or absent/distorted waveform morphology.	Excluded	71% [52-86]	74% [68-79]
Mandala 2013	Increased interpeak I-III and/or I-V latencies.	Excluded	37% [23-52]	96% [87-100]
Raber 1997	Not specified.	Excluded	75%	57%
Rupa 2003	Increased interpeak latencies (I-III of ≥ 2.5 ms, III-V of ≥ 2.3 ms, I-V of ≥ 4.4 ms), interaural latency difference of ≥ 0.3 ms, poor waveform morphology and replicability or absent response despite normal/mildly elevated audiometric thresholds.	Excluded	100% [40-100]	62% [49-73]
		Included	67% [22-96]	69% [58-79]
Scherler 1995	Increased I-V interpeak latency and/or absolute wave V latency and/or interaural latency difference of wave V. Values of > 2 standard deviations of the mean of a normal population are abnormal.	?	94% [83-99]	86% [73-94]

?: Unclear/unknown
CI: Confidence interval
ABR: Auditory brainstem response, ms: millisecond

The sensitivity of different combinations of presenting symptoms varied, with the highest sensitivity of 58% [95% CI: 28-85] for symptoms of hearing loss. Reported specificities were higher, ranging from 63% [95% CI: 60-66] for hearing loss only, to 99% [95% CI: 98-100] for the combination of hearing loss, pulsatile tinnitus, dizziness and other symptoms. [20]

The sensitivity and specificity of electronystagmography were 70% [95% CI: 55-82] and 48% [95% CI: 34-63], 43% [95% CI: 29-58] and 91% [95% CI: 79-97] for caloric irrigation and 65% [95% CI: 50-78] and 98% [95% CI: 90-100] for HVT, respectively. [15, 22] Chatrath et al. (2008) developed a prediction model containing PTA thresholds and presence of vertigo. Sensitivity ranged from 80% to 95% [95% CIs: unknown] and specificity from 74% to 38% [95% CIs: unknown] depending on the chosen cut-off value. [24]

Sensitivity analysis

We were unable to explore heterogeneity in more detail, because of the limited amount and quality of data.

DISCUSSION

Summary of main results

This systematic review and meta-analysis of the diagnostic accuracy of different non-imaging screening protocols to identify patients with VS revealed large ranges in sensitivity and specificity values of diagnostic tests. On average, sensitivity values were quite high. Specificity values on the other hand were low. Results of the pooled estimates of 5 PTA protocols revealed a sensitivity of 88% [95% CI 84-91] to 91% [95% CI: 52-99] and a specificity from 31% [95% CI: 10-66] to 58% [95% CI: 49-65]. Most studies were of low to moderate quality.

Strengths and limitations

This is the first systematic review and meta-analysis studying different screening protocols for patients at risk of VS. As others have previously studied one screening method, we gathered evidence of multiple screening methods in this review. [5] The major strength of this diagnostic review is that we made no restrictions to the search in terms of publication year, status or in language. Furthermore, we were able to provide some pooled estimates of relatively large study populations that tested the same PTA protocols.

Some potential limitations should also be discussed. First, we were not able to obtain pooled estimates of all tests, because incidence rates among studies differed or protocols and/or thresholds of the described tests were too heterogeneous. Pooling of these studies would result in unreliable estimates. [25] Second, the included studies often had a lack of important details. For example, for most studies it was unclear whether results of the index test were interpreted without knowledge of the reference standard, therefore the risk of biased results was unclear. Furthermore, details about MRI (field strength, sequence and assessment) were lacking in most studies, therefore it was not always possible to assess its ability to correctly detect lesions. Third, we performed one overall search for all tests instead of separate searches for each index test. We do, however, believe we have identified all available evidence, as we did not encounter additional items during screening of references and related articles. Fourth, we intended to study heterogeneity in more detail by performing sensitivity analyses of the clinically relevant covariates (age, VS size, location of VS and MRI field strength and sequence). Unfortunately, this was not possible due to the limited amount and quality of data. Finally, only one study looked at tinnitus as an isolated symptom. It did not provide enough data to draw final conclusions regarding the question whether asymmetrical tinnitus without asymmetry on PTA might also be an indication for MRI.[20] It therefore remains a challenge to select high risk patients from this particular group.

Implications for clinical practice

It would be ideal to use a relatively simple, low-cost test as a first screening tool for VS, such as PTA. Most screening protocols published in this review make use of PTA as a screening tool. [14, 17-19, 21, 22] Based on the pooled results from this meta-analysis, AAO-HNS, Obholzer 1 and Sunderland protocols, appeared to

have highest sensitivities in detecting VS, but all three have poor specificities. [14, 17, 21] With a lack of better alternatives, these tests could potentially lead to a reduction in the number of patients needing MRI. Patients with a positive non-imaging screening test should still undergo MRI to confirm the diagnosis and to get information on tumour size and location, i.e. to make well informed decisions regarding treatment. However, when using these protocols up to 9% of VSs would be missed (false negative rate). At this moment we cannot assess what consequences false negative results have on clinical outcome and costs, therefore it is difficult to assess the impact of a false negative rate of 9%. Moreover, one needs to consider the large confidence intervals around the point estimates of sensitivity and specificity rates, and the low to moderate quality of the included studies. This uncertainty could also explain why none of the PTA protocols are currently being used in an international guideline for VS screening, despite their fairly high sensitivity values. Before MRI became the standard, ABR was frequently used to detect VS. The ability of ABR to detect small intracanalicular lesions is limited and ABR cannot be performed in patients with severe hearing loss of more than about 70 dB, while these patients are part of the screening population. [5, 13, 15] This results in a relatively small population in which ABR could be considered a reliable method for VS screening. Moreover, ABR is a relatively costly and time-consuming investigation and its reported diagnostic accuracy varies among included studies, therefore we find it an unsuitable screening tool for VS. [5, 13, 15, 16, 22, 23]

Implications for research

This review shows that reliable screening protocols are currently not available. If one aims to detect every single case of VS, every patient with the slightest asymmetrical sensorineural hearing loss and/or unilateral audiovestibular dysfunction should undergo MRI. However, when we look at the amount of negative MRIs in the population screened for VS (more than 95% negative for VS), we consider this a great waste of resources. Based on Dutch cost prices, an average of €177 (approximately US$198) could be saved per patient at risk of VS, when we would be able to avoid all MRIs negative for pathology. [11] With the current discussions about the use of resources and overdiagnosis in mind, we believe there is a need for more reliable screening protocols to identify VS in patients with asymmetrical sensorineural hearing loss and/or unilateral audiovestibular dysfunction.[26] It would also be worthwhile to investigate which VSs will require future treatment and focus on diagnosing these VSs in particular, because it is known that two thirds of VSs in a wait and scan policy remain stable for years without a need for treatment. [27, 28]

Our results confirm that none of the currently available non-imaging screening protocols appear to be accurate in detecting VS.

ACKNOWLEDGEMENTS

The authors thank Gerjon Hannink for his help with the analyses, Motokazu Yanagi and Queeny Tsang for help with translations, and Steen Gimsing for his devoted support. This study was funded by a Health Care Efficiency Research grant by The Netherlands Organization for Health Research and Development (ZonMw). Conflicts of interest: none declared. The review protocol was submitted to PROSPERO (available at: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42015025618)

REFERENCES

1. Howitz M.F., Johansen C., Tos M. et al. Incidence of Vestibular Schwannoma in Denmark, 1977-1995. *Am J Otol*. 2000;21:690-694.
2. Lin D., Hegarty J.L., Fischbein N.J. et al. The Prevalence of "Incidental" Acoustic Neuroma. *Arch Otolaryngol Head Neck Surg*. 2005;131:241-244.
3. Propp J.M., McCarthy B.J., Davis F.G. et al. Descriptive Epidemiology of Vestibular Schwannomas. *Neuro Oncol*. 2006;8:1-11.
4. Stangerup S.E., Tos M., Thomsen J. et al. True Incidence of Vestibular Schwannoma? *Neurosurgery*. 2010;67:1335-1340.
5. Fortnum H., O'Neill C., Taylor R. et al. The Role of Magnetic Resonance Imaging in the Identification of Suspected Acoustic Neuroma: A Systematic Review of Clinical and Cost Effectiveness and Natural History. *Health Technol Assess*. 2009;13(iii-iv, ix-xi):1-154.
6. Godefroy W.P., Kaptein A.A., Vogel J.J. et al. Conservative Treatment of Vestibular Schwannoma: A Follow-up Study on Clinical and Quality-of-Life Outcome. *Otol Neurotol*. 2009;30:968-974.
7. Annesley-Williams D.J., Laitt R.D., Jenkins J.P. et al. Magnetic Resonance Imaging in the Investigation of Sensorineural Hearing Loss: Is Contrast Enhancement Still Necessary? *J Laryngol Otol*. 2001;115:14-21.
8. Mahrous A.K. Kalepu R. Positive Findings on Mri in Patients with Asymmetrical Snhl. *Eur Arch Otorhinolaryngol*. 2008;265:1471-1475.
9. Newton J.R., Shakeel M., Flatman S. et al. Magnetic Resonance Imaging Screening in Acoustic Neuroma. *Am J Otolaryngol*. 2010;31:217-220.
10. Zealley I.A., Cooper R.C., Clifford K.M. et al. MRI Screening for Acoustic Neuroma: A Comparison of Fast Spin Echo and Contrast Enhanced Imaging in 1233 Patients. *Br J Radiol*. 2000;73:242-247.
11. Zorginstituut Nederland (2015). Guideline for Cost Research in Health Care; Available at: https://www.zorginstituutnederland.nl/binaries/content/documents/zinl-www/documenten/publicaties/overige-publicaties/1602-richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg/1602-richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorgRichtlijn+voor+het+uitvoeren+van+economische+evaluaties+in+de+gezondheidszorg.pdf.
12. Whiting P.F., Rutjes A.W., Westwood M.E. et al. Quadas-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med*. 2011;155:529-536.
13. Cueva R.A. Auditory Brainstem Response Versus Magnetic Resonance Imaging for the Evaluation of Asymmetric Sensorineural Hearing Loss. *Laryngoscope*. 2004;114:1686-1692.
14. Gimsing S. Vestibular Schwannoma: When to Look for It? *J Laryngol Otol*. 2010;124:258-264.
15. Mandala M., Giannuzzi A., Astore S. et al. Hyperventilation-Induced Nystagmus in Vestibular Schwannoma and Unilateral Sensorineural Hearing Loss. *Eur Arch Otorhinolaryngol*. 2013;270:2007-2011.
16. Rupa V., Job A., George M. et al. Cost-Effective Initial Screening for Vestibular Schwannoma: Auditory Brainstem Response or Magnetic Resonance Imaging? *Otolaryngol Head Neck Surg*. 2003;128:823-828.
17. Saliba I., Bergeron M., Martineau G. et al. Rule 3,000: A More Reliable Precursor to Perceive Vestibular Schwannoma on Mri in Screened Asymmetric Sensorineural Hearing Loss. *Eur Arch Otorhinolaryngol*. 2011;268:207-212.

18. Cheng T.C. Wareing M.J. Three-Year Ear, Nose, and Throat Cross-Sectional Analysis of Audiometric Protocols for Magnetic Resonance Imaging Screening of Acoustic Tumors. *Otolaryngol Head Neck Surg*. 2012;146:438-447.

19. Suzuki M., Hashimoto S., Kano S. et al. Prevalence of Acoustic Neuroma Associated with Each Configuration of Pure Tone Audiogram in Patients with Asymmetric Sensorineural Hearing Loss. *Ann Otol Rhinol Laryngol*. 2010;119:615-618.

20. Vandervelde C. Connor S.E. Diagnostic Yield of Mri for Audiovestibular Dysfunction Using Contemporary Referral Criteria: Correlation with Presenting Symptoms and Impact on Clinical Management. *Clin Radiol*. 2009;64:156-163.

21. Obholzer R.J., Rea P.A. Harcourt J.P. Magnetic Resonance Imaging Screening for Vestibular Schwannoma: Analysis of Published Protocols. *J Laryngol Otol*. 2004;118:329-332.

22. Scherler M. Bohmer A. Diagnostic Value of Clinical Tests in the Diagnosis of Acoustic Neuromas. [German]. *HNO*. 1995;43:487-491.

23. Raber E., Dort J.C., Sevick R. et al. Asymmetric Hearing Loss: Toward Cost-Effective Diagnosis. *J Otolaryngol*. 1997;26:88-91.

24. Chatrath P., Frosh A., Gore A. et al. Identification of Predictors and Development of a Screening Protocol for Cerebello-Pontine Lesions in Patients Presenting with Audio-Vestibular Dysfunction. *Clin Otolaryngol*. 2008;33:102-107.

25. Leeflang M.M. (2014) Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy. *Clin Microbiol Infect*. 20**,** 105-113.

26. Glasziou P., Moynihan R., Richards T. et al. Too Much Medicine; Too Little Care. *BMJ*. 2013;347**:** f4247.

27. Ferri G.G., Modugno G.C., Pirodda A. et al. Conservative Management of Vestibular Schwannomas: An Effective Strategy. *Laryngoscope*. 2008;118:951-957.

28. Hajioff D., Raut V.V., Walsh R.M. et al. Conservative Management of Vestibular Schwannomas: Third Review of a 10-Year Prospective Study. *Clin Otolaryngol*. 2008;33:255-259.

SUPPLEMENTAL CONTENT

Supplement 2.1: Search strategy.

Search performed on July 28, 2016.

MEDLINE (OvidSP)

1. hearing disorders/ or hearing loss, unilateral/
2. ((hearing adj3 disorder?) or (hearing adj3 loss adj3 unilateral) or (sensorineural adj3 hearing adj3 loss)).tw.
3. (hearing adj3 impair*).tw,kw.
4. Tinnitus/
5. tinnitus.tw,kw.
6. dizziness/ or vertigo/
7. (dizziness or vertigo).tw,kw.
8. ((balance adj3 impair*) or (vestibular adj3 impairment) or (vestibular adj3 disorder) or (vestibular adj3 dysfunction)).tw,kw.
9. or/1-8
10. Magnetic Resonance Imaging/
11. (mri or (magnetic adj3 resonance) or "mr scan" or nmr).tw.
12. (mri or (magnetic adj3 resonance) or "mr scan" or nmr).kw.
13. Diagnostic Imaging/
14. or/10-13
15. 9 and 14
16. exp Neuroma, Acoustic/
17. (acoustic adj3 tumo?r?).tw,kw.
18. (acoustic adj3 l??sion?).tw,kw.
19. (acoustic adj3 schwannoma?).tw,kw.
20. (vestibular adj3 schwannoma?).tw,kw.
21. (vestibular adj3 tumo?r?).tw,kw.
22. (acoustic adj3 neuri?????oma?).tw,kw.
23. (CPA adj3 tumo?r?).kw,tw.
24. (cpa adj3 l??sion?).tw,kw.
25. (cerebello?pontine angle adj3 tumo?r?).tw,kw.
26. (cerebello?pontine angle adj3 l??sion?).tw,kw.
27. or/16-26
28. 15 and 27

Embase (OvidSP)

1. exp unilateral hearing loss/
2. exp perception deafness/
3. *hearing impairment/
4. ((hearing adj3 loss) or (perception adj3 deafness) or (hearing adj3 impair*)).tw,kw.
5. exp tinnitus/
6. tinnitus.tw,kw.
7. *dizziness/
8. exp vestibular disorder/
9. exp balance impairment/
10. imbalance.ti.
11. ((balance adj3 impair*) or vertigo or (vestibular adj3 impairment) or (vestibular adj3 disorder) or dizziness or (vestibular adj3 dysfunction)).tw,kw.
12. or/1-11
13. exp nuclear magnetic resonance imaging/
14. *diagnostic imaging/
15. (mri or (magnetic adj3 resonance) or "mr scan" or nmr).tw,kw.
16. or/13-15
17. 12 and 16
18. exp acoustic neurinoma/
19. ((vestibular adj3 schwannoma?) or (acoustic adj3 tumo?r?) or (acoustic adj3 neur?????oma?) or (vestibular adj3 tumo?r?) or (acoustic adj3 schwannoma?) or (acoustic adj3 l??sion?) or (CPA adj3 tumo?r?) or (CPA adj3 l??sion?)).tw,kw.
20. (cerebello?pontine angle adj3 tumo?r?).kw,tw.
21. (cerebello?pontine angle adj3 l??sion?).kw,tw.
22. or/18-21
23. 17 and 22

The Cochrane Library

#1 MeSH descriptor: [Tinnitus] this term only
#2 tinnitus:ti,ab,kw
#3 MeSH descriptor: [Hearing Loss, Unilateral] this term only
#4 (unilateral near/3 hearing near/3 loss):ti,ab,kw
#5 (sensorineural near/3 hearing near/3 loss):ti,ab,kw
#6 (hearing near/3 impairment):ti,ab,kw
#7 MeSH descriptor: [Hearing Loss, Sensorineural] this term only
#8 MeSH descriptor: [Vertigo] this term only
#9 vertigo:ti,ab,kw
#10 {or #1-#9}
#11 MeSH descriptor: [Magnetic Resonance Imaging] this term only
#12 (magnetic near/3 resonance near/3 imag*):ti,ab,kw
#13 (nmr near/3 imag*):ti,ab,kw
#14 {or #11-#13}
#15 #10 and #14

Supplement 2.2: Overview of missing data, assumptions and translations.

Chatrath et al. 2008
The study stated the number of CPA lesions and controls of the total study population (n=369), but did not provide the number of patients and controls used for split-half analysis. Therefore, we were not able to extract absolute data.

Cheng et al. 2012
This study described two types of sensitivity and specificity. We used the control group consisting of non-acoustic pathology and radiological normal cases.

Cueva et al. 2004
This study had an incomplete list of references. We obtained the missing references after contacting one of the authors.

Gimsing et al. 2010
This study presented screening protocols of which the numbers were not exact enough to obtain a 2x2 table. After contacting one of the authors, we received the correct study results. Therefore, results shown in this paper do not correspond with the original article.

Katsura et al. 2004
Japanese translator consulted.

Mandala et al. 2013
We chose not to present the results of bedside examination by Mandala et al. They did not state how these tests were performed, nor did they state any cut-off values.

Obholzer et al. 2004
The study described a large set of PTA results. For this review we only used results of PTA protocols, which have been previously described in literature. Some of the results presented by Obholzer et al. made use of the same audiological criteria as other protocols, although they were not identified as such in the study.

Raber et al. 1997
We only present results of patients that underwent MRI in the study of Raber et al. This study stated sensitivity and specificity rates for a subset of the population without providing population details. We were not able to construct a 2x2 table from this study, therefore we present the results without confidence intervals.

Ruckenstein et al. 1996

We excluded this reference because one of the authors confirmed the study used the same study population as Cueva et al.

Rupa et al. 2003

We reported the results of Rupa et al. both ignoring and taking into account (as false negative) the 'no response' group.

Saliba et al. 2011

This study reported discrepancies in statements about the total study population. We decided that n=212 was most likely the correct study population. This study also combined the Oxford and Nashville protocol for unknown reasons, therefore we excluded those protocols from our analysis. Moreover, specificity of the AAO-HNS protocol was not exact enough to draw absolute numbers from, therefore we rounded down the true negative rate (69 true negatives, 59 false positives, sensitivity 54%).

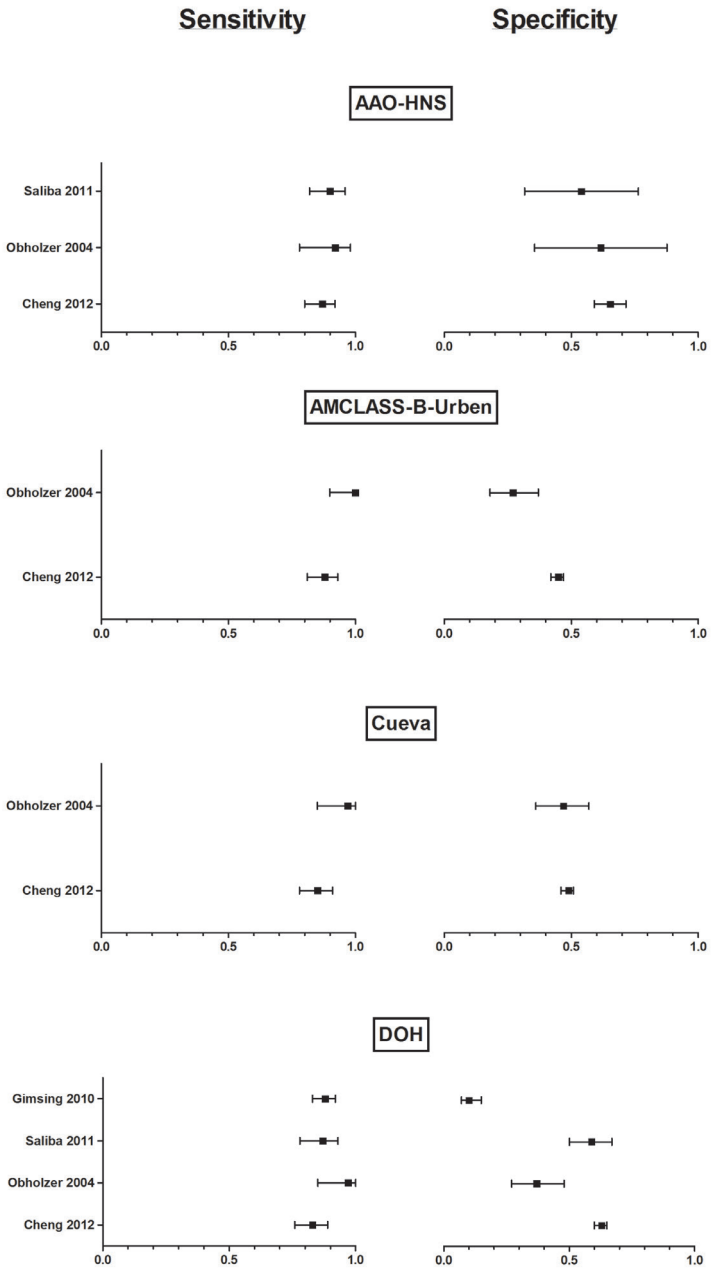
Scherler et al. 1995

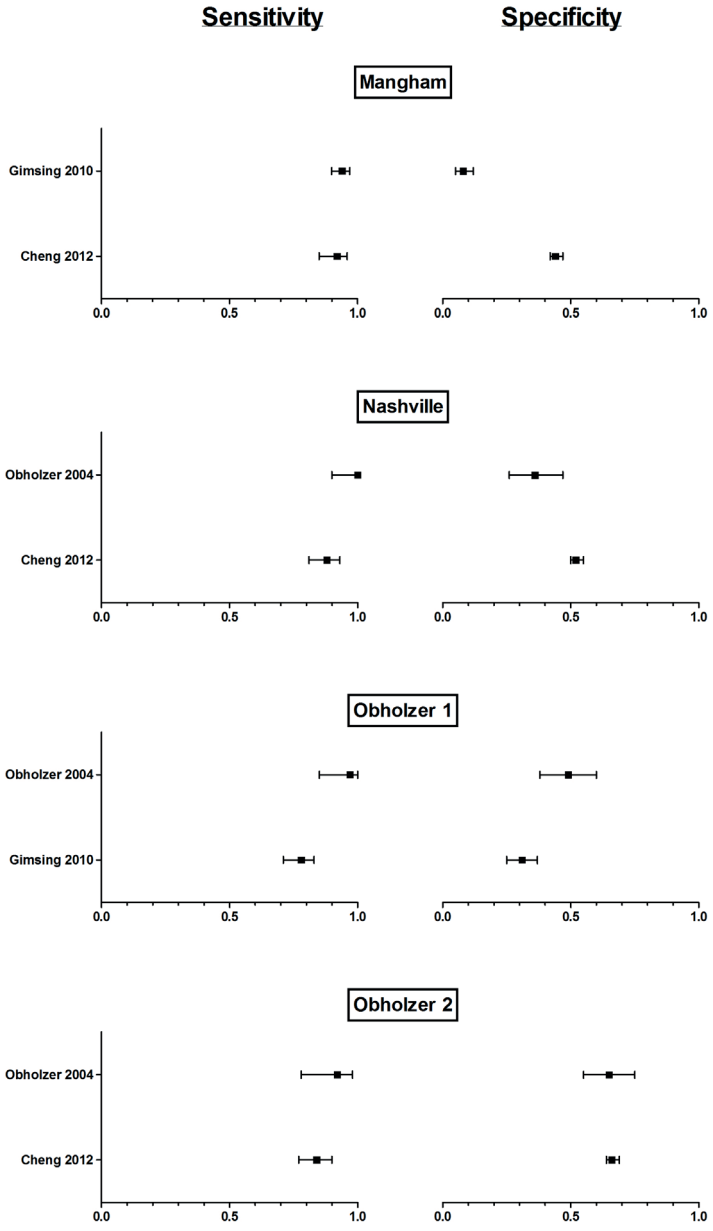
The sensitivity rate for the ENG turned out not to be exact enough to draw absolute numbers from. Therefore, we rounded down the amount of true positive ENG examinations (35 true positives, 15 false negatives, sensitivity 70%), preventing overestimation of sensitivity.

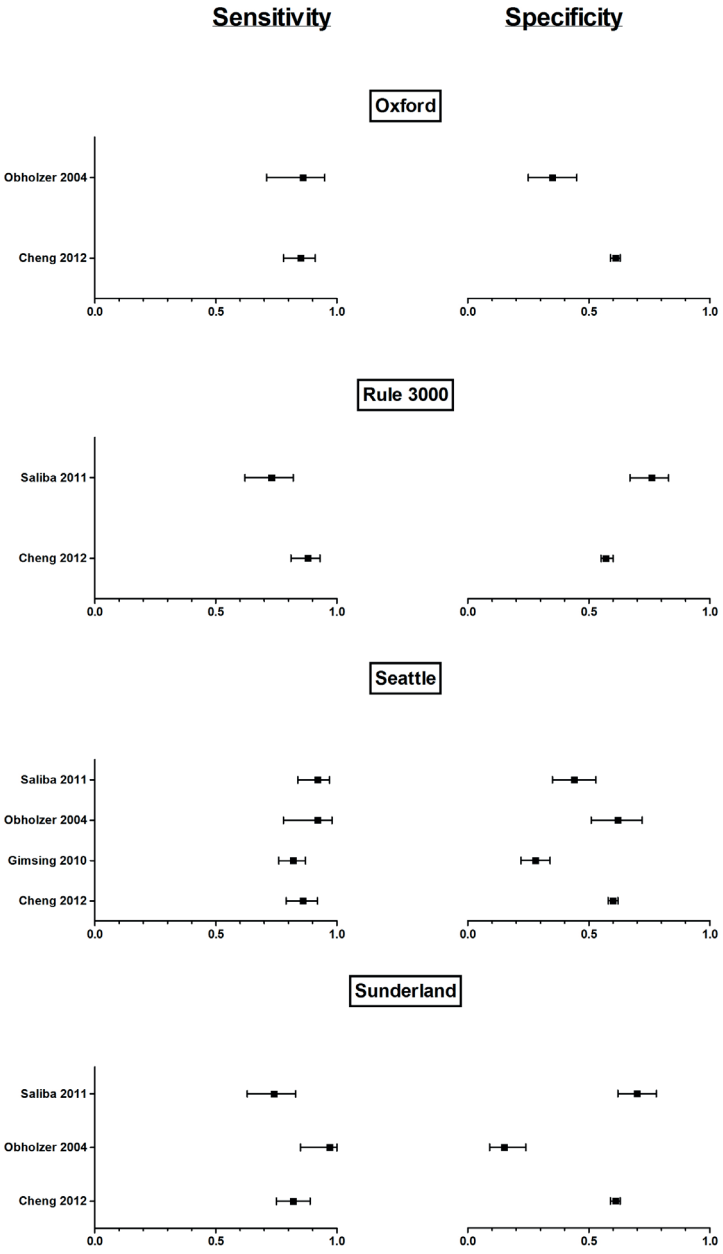
Xu et al. 2005

Chinese translator consulted.

Supplemental Figure 2.1 Forest plots of sensitivity and specificity of different pure-tone audiometry protocols per study. Each square indicates the point estimate, whereas the horizontal line indicates the 95% confidence region.




Sensitivity
Specificity
Mangham
Gimsing 2010
Cheng 2012
0.0
0.5
1.0
Nashville
Obholzer 2004
Cheng 2012
Obholzer 1
Obholzer 2004
Gimsing 2010
Obholzer 2
Obholzer 2004
Cheng 2012


Sensitivity
Specificity
Oxford
Obholzer 2004
Cheng 2012
Rule 3000
Saliba 2011
Cheng 2012
Seattle
Saliba 2011
Obholzer 2004
Gimsing 2010
Cheng 2012
Sunderland
Saliba 2011
Obholzer 2004
Cheng 2012
0.0
0.5
1.0

CHAPTER 3

An international comparison of diagnostic and management strategies for vestibular schwannoma

Hentschel M.
Rovers M.
Markodimitraki L.
Steens S.
Kunst H.

Published as:
Eur Arch Otorhinolaryngol. 2019;276(1):71-78. doi: 10.1007/s00405-018-5199-6.

ABSTRACT

Objective

To compare international diagnostic and management strategies for vestibular schwannoma (VS).

Methods

A web-based questionnaire was sent to 130 otolaryngologists, mainly identified through the European Skull Base Society. It contained questions on general information including guideline usage as well as questions on diagnosis (focussing on selection of patients for MRI) and management of VS, including case scenarios. Descriptive statistics were reported.

Results

Thirty-six otolaryngologists working in 11 different countries completed the questionnaire (response rate: 28%). Guidelines for diagnosis and management of VS are used by 44% and 42% of respondents, respectively. In the diagnostic strategy for VS, different types and combinations of audiovestibular function tests are used when deciding whether a patient should undergo an MRI. Respondents apply 18 different definitions of asymmetrical hearing loss. Variation was also apparent from reported considerations on management of VS. Most respondents (84%) prefer a wait and scan strategy in case of a small intrameatal VS (Koos 1). Variety in management strategies increases for patients with a medium to large sized VS (Koos 2, 3 and 4). The details of each management strategy (wait and scan, microsurgery, stereotactic radiosurgery and fractionated radiotherapy) also differ among respondents.

Conclusions

A large variation in diagnostic and management strategies for VS was identified between respondents. More evidence and/or consensus seem warranted to reduce uncertainties for patients, and differences in outcome and costs that might result from the variety of strategies currently being applied.

INTRODUCTION

Patients with a vestibular schwannoma (VS) usually present with symptoms of (asymmetrical) sensorineural hearing loss, tinnitus, vertigo and/or disequilibrium. Magnetic resonance imaging (MRI) is considered the gold standard to diagnose VS and is performed whenever there is a high suspicion of VS in patients with aforementioned symptoms.[1] It is, however, a challenge to determine which patients should undergo MRI, the reported yield of diagnostic MRIs being approximately 3%.[2, 3] There are several tests available that can help to determine whether a patient should be referred for MRI. Pure tone audiometry is usually the first step in the diagnostic process and is sometimes followed by other audiovestibular function tests, such as auditory brainstem response, speech perception tests and electronystagmography.[4, 5] Although numerous studies examined the effectiveness of these audiovestibular function tests in selecting patients for MRI,[1, 5-7] there seems no consensus regarding their role in everyday practice.

Apart from the variability in diagnostic strategies, there are multiple management strategies available for VS, consisting of microsurgical resection, radiation therapy (fractionated radiotherapy, stereotactic radiosurgery) or "wait and scan" (W&S, observation with serial imaging aiming to detect tumour growth). Over the past years the W&S strategy has gained popularity in Europe and the US.[8, 9] Treatment is increasingly being reserved for patients with a large size and/or growing VS. Because the natural growth pattern of VS is variable and unpredictable, [10, 11] it is a challenge to determine the time interval between MRIs in the W&S strategy as well as indications for, and type of treatment.

An international guideline concerning the diagnosis and management of VS is lacking. Specialists seem to counsel their patients based on personal preference and experience.[12] A lack of guidelines prescribing the appropriate use of various strategies may contribute to inconsistencies in care delivery among specialists. These, in turn, may lead to uncertainties for patients, unnecessary differences in outcome between patients and unnecessary variation in costs associated with care. For these reasons, it is important to identify practice variations and possibilities to further improve healthcare.

In this study we aimed to investigate variation in diagnostic and management strategies for VS across countries, and explore determinants of such variation.

METHODS

Study design and population

To obtain information regarding the current diagnostic and management strategies for VS, an online questionnaire was sent to 102 otolaryngologists working in 15 different countries, that were registered in the European Skull Base Society database. We additionally invited 28 otolaryngologists whose contact

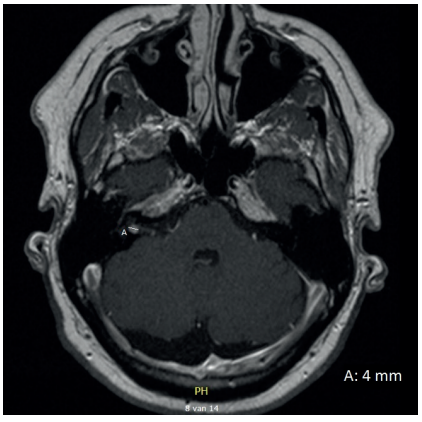
details were acquired through hospital websites or by personal acquaintance (excluding the authors' hospitals). The questionnaire was distributed using 'Castor EDC'[13] in January 2017. We sent two reminders in a time span of two weeks.

Questionnaire

The questionnaire consisted of three main sections. The first section contained questions on general topics, e.g. patient volumes and guideline usage. The second section focused on audiovestibular function tests and investigated based on what parameters respondents refer a patient for MRI. In the third section different management strategies for VS were addressed, including the proportion of patients being assigned to each strategy, variables used when considering management strategies, and their conduct. It included several case scenarios describing VSs of increasing size (see Figure 3.1). Respondents chose the preferred management option for each case, which enabled us to further explore their considerations and assess impact on individual patients. The questionnaire was tested prior to its distribution, and adapted accordingly to the comments received by two otolaryngologists, a professor in evidence-based surgery, a radiologist specialized in neuro- and head & neck radiology, and a junior researcher in the field of VS.

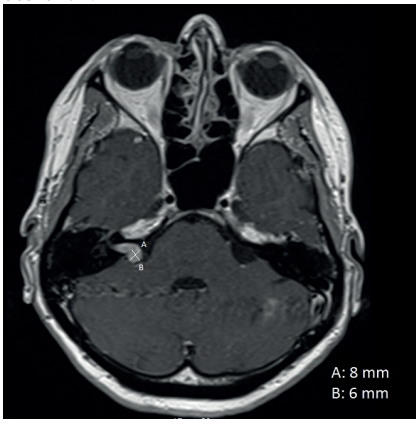
Case 1

Scenario I:



A female patient of 50 years old in overall good health has symptoms of hearing loss and unilateral tinnitus in the right ear. Audiometry shows asymmetrical sensorineural hearing loss. You decide to screen this patient for VS by acquiring an MRI. It shows an intrameatal VS (Koos grade 1), 4 mm maximum diameter.

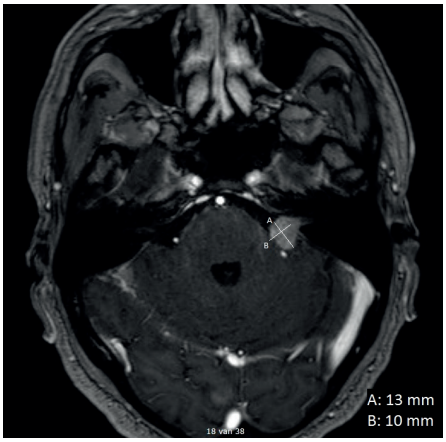
Scenario II:



W&S → Scenario IIA: Same case, follow up after two years, hearing loss has worsened and MRI reveals significant growth of the VS.
No W&S → Scenario IIB: Same case, different maximum diameter.
In both scenarios IIA and IIB MRI shows a VS extending in the CPA without contacting the brainstem *(Koos grade 2)*.

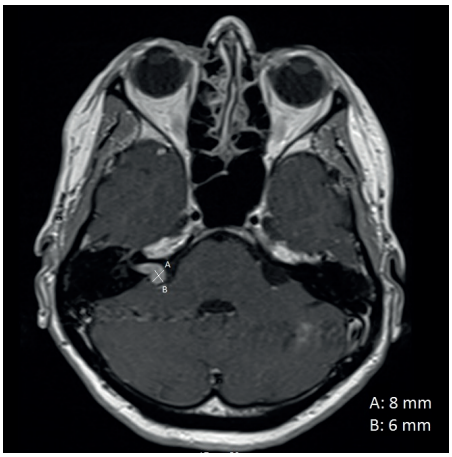
Case 2

Scenario I:



A male patient of 65 years old in overall good condition has symptoms of hearing loss, a sense of pressure in the left ear and vertigo symptoms. Audiometry shows asymmetrical sensorineural hearing loss. MRI reveals a 13 mm maximum diameter VS, contacting the brainstem (Koos grade 3).

Scenario II:



W&S → Scenario IIA: Same case, follow up after three years. The patient complaints of facial numbness and MRI reveals significant growth.
No W&S → Scenario IIB: Same case, but additional complaints of facial numbness. MRI shows a VS with a different maximum diameter. In both scenarios IIA and IIB, MRI shows a lesion exerting pressure on the brainstem *(Koos grade 4)*.

Figure 3.1 Case scenarios and MRI images used in the questionnaire.

Data analysis

Diagnostic and management strategies were compared between respondents and countries, by providing descriptive statistics. Percentages and medians or means were calculated where applicable. Data were analysed per item, so the denominator varied per question due to missing data. All analyses were performed using IBM SPSS Statistics for Windows, version 22 (IBM Corp., Armonk, N.Y., USA).

RESULTS

Respondents

Of the 130 addressed, 36 otolaryngologists from 11 different countries returned the questionnaire (response rate 28%) (Supplemental table 3.1), visited by a median of 85 (n=32, range 3-300) patients with newly diagnosed VSs in 2016. All respondents (n=36) indicated to work at a university hospital and the mean experience in working with VS patients was 16 years (SD 7.3).

The majority of respondents (56%) never uses a guideline, neither for diagnosis nor management. Despite the availability of guidelines in the Netherlands (asymmetrical hearing loss and tinnitus)[14] and the United Kingdom (diagnosis and management)[9], heterogeneity was found in responses from these countries.

Thirty respondents (88%) participate in multidisciplinary meetings to discuss VS cases, while the remaining 4 (12%) do not have such meetings.

Diagnostic strategies

All respondents have MRI available, and most (79%) use contrast-enhanced MRI to diagnose VS. There was more variety regarding the use of audiovestibular function tests to select patients for MRI. Most respondents (n=31, 94%) use pure-tone audiometry in the diagnostic process, the remaining two respondents solely relying on auditory brainstem response. We investigated the definition of asymmetrical hearing loss that respondents apply when referring patients for MRI. The minimum asymmetry of hearing loss being used ranged from 5 to 30 dB, 20 dB being mostly used (n=11, 36%). Most respondents (n=27, 87%) define asymmetry as an absolute difference between ears at specific frequencies, while 4 (13%) calculate a mean. However, how many and which frequencies are considered as well as their mutual relation (adjacent or non-adjacent) varies. Respondents provided 18 different definitions of asymmetrical hearing loss, including 12 different combinations of frequencies that are considered important. The most frequently reported definition (n=6, 19%) uses an absolute asymmetry of 10 dB as threshold, followed by an absolute asymmetry of 20 dB (n=4, 13%).

Twenty-four respondents (73%) usually perform speech audiometry to decide if a patient should undergo MRI. Most (n=14, 58%) use speech reception thresholds.

Moreover, respondents explained to look for the roll-over phenomenon as well as discrepancy with pure-tone audiometry measurements.

Twenty-one respondents (64%) order an MRI for patients with unilateral tinnitus as only symptom. Its duration varied (0-11 months), but most respondents apply a minimum of 3 months (n=6, 29%). Twenty-one respondents (64%) perform electronystagmography in case of vertigo and 16 respondents (49%) use auditory brainstem response.

Additional audiovestibular function tests consisting of (video) head impulse tests and/or vestibular-evoked myogenic potentials are used by 4 and 3 respondents, respectively.

Management strategies

Most respondents (n=15, 42%) deem tumour size the most important variable when deciding about management strategies, followed by cerebellopontine angle size/intracranial space (n=10, 28%). VS size is measured using dimensional and volumetric measurements by 27 (84%) and 5 (16%) respondents, respectively. Twenty respondents (63%) use a specific threshold in tumour size/volume when considering treatment, varying from 15 to 30 mm, a majority applying 20 mm (n=9, 45%). Five respondents (16%) base their decision to proceed to treatment on tumour size/volume only, while most (n=15, 47%) also consider other variables (e.g. brainstem contact and/or compression, patient characteristics and/or symptoms, and/or tumour growth).

In 2016, the proportion of VS patients assigned to each management strategy at time of diagnosis varied (Figure 3.2). Both microsurgery and W&S were applied in all participating centres in varying proportions. Twenty-one respondents (70%) assigned a majority of patients to W&S, while only 5 (17%) did so for microsurgery. Stereotactic radiosurgery was prescribed by more respondents (n=22, 73%) than fractionated radiotherapy (n=11, 37%).

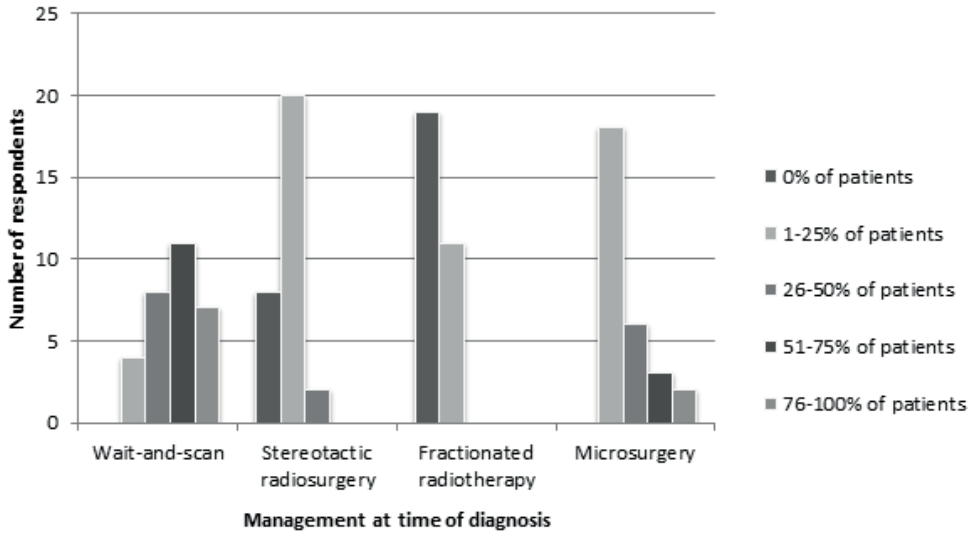


Figure 3.2 Fraction of VS patients assigned to management strategies at time of diagnosis, reported by 30 respondents. y-axis: number of respondents, x-axis: different management strategies at time of diagnosis.

Wait and scan

Figure 3.3 provides a schematic overview of W&S strategies being used. Strategies vary in timing of the first MRI following diagnosis, interval periods and total duration of the observational period. Respondents quit the W&S strategy when a patient reaches a specific age (75 and 80 years), after a specified period (4-21 years), or continue lifelong. Most (n=18, 56%) define significant tumour growth as an increase in diameter of ≥2 mm, while others apply an increase of ≥1 mm (22%) or a volume increase of ≥10% (9%). Ten respondents (31%) consider tumour growth a strict indication for treatment, whereas 69% also take other variables into account (e.g. tumour size, symptoms and/or patient age, health and/or preference).

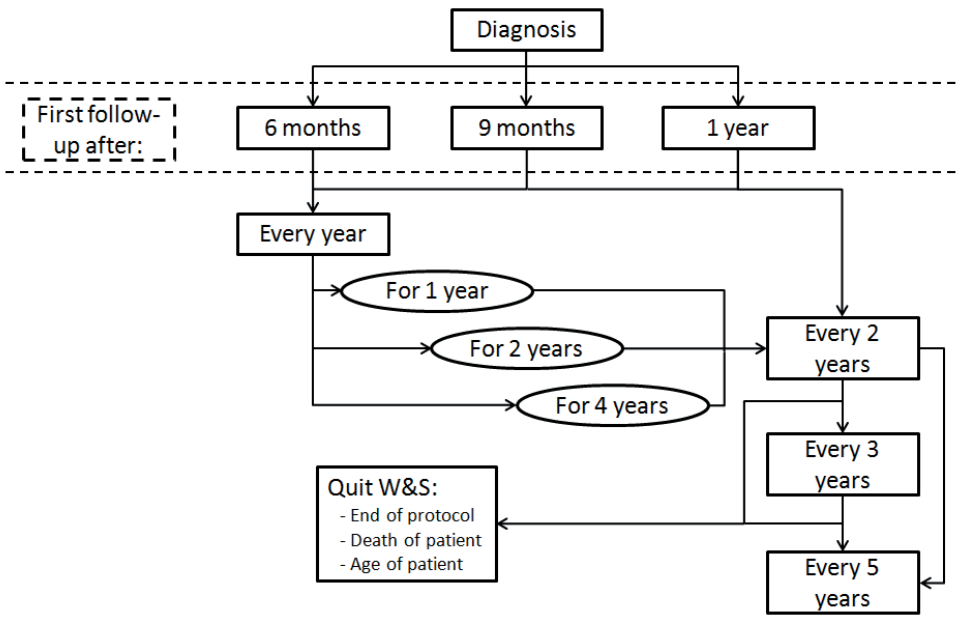


Figure 3.3 Display of variation in applied wait and scan (W&S) strategies.

Radiation therapy

Of the respondents prescribing stereotactic radiosurgery, most use Gamma Knife (n=11, 52%). Fractionated radiotherapy treatment plans were reported by 10 respondents and contain up to 30 fractions, with total dosages of 12 to 60 Gy.

Microsurgery

The microsurgical approach most frequently used by respondents is translabyrinthine (n=20, 65%), followed by suboccipital (n=6, 19%) and middle fossa (n=5, 16%). Whenever the facial nerve is difficult to recognise intra-operatively, most respondents (n=29, 94%) opt for incomplete VS removal in order to reduce the risk of facial nerve injury.

Case scenarios

Table 3.1 displays the management strategies chosen for the different case scenarios (from Figure 3.1). There is much agreement about conservative management for small intrameatal VSs, W&S being the most popular choice (84%) for Case 1 – Scenario I. Even though all respondents indicated to allocate part of their patients to W&S, 5 (16%) preferred treatment over W&S for this scenario, motivated by the patient's age (50 years) and chances of hearing preservation. For VSs extending in the CPA but not contacting the brainstem (Case 1 – Scenario II) diversity in management choices becomes more apparent. For the VS that had grown during W&S (Case 1–Scenario IIA) some respondents would let the patient choose between W&S and treatment, and/or different treatment modalities. For the VS making brainstem contact (Case 2 – Scenario I), preferred management options consisted mostly of W&S or microsurgery. For the largest VS causing brainstem compression (Case 2 – Scenario II) differences were reduced again, microsurgery being most popular. The case scenarios confirm that fractionated radiotherapy is an uncommon treatment option for VS. Remarkably, some respondents working in the same centre chose different management options for the same case scenario. It seemed that otolaryngologists from high volume centres (>100 new VS patients per year) preferred W&S over treatment for Case 2 - Scenario I. For the other case scenarios, we could not identify a difference between low- and high-volume centres.

Table 3.1 Management strategies chosen for two patients with different clinical characteristics. Scenarios were adapted to choices made in scenario I. W&S: wait and scan, RS: stereotactic radiosurgery, RT: fractionated radiotherapy, MS: microsurgery.

	Case 1			**Case 2**		
Management strategy	Scenario I N=32 (n(%))	Scenario IIA n=27* (n(%))	Scenario IIB n=5** (n(%))	Scenario I N=32 (n(%))	Scenario IIA n=16* (n(%))	Scenario IIB n=16**† (n(%))
W&S	27 (84)	4 (15)	-	16 (50)	1 (6)	-
RS	-	11 (41)	-	1 (3)	4 (25)	-
RT	1 (3)	-	-	1 (3)	1 (6)	-
MS	4 (13)	12 (44)	5 (100)	14 (44)	10 (63)	14 (88)

* Respondents that had chosen a W&S strategy in scenario I.

** Respondents that had chosen to proceed to treatment in scenario I.

† Two respondents that completed Scenario I did not complete this question.

Between-country comparisons

In the six countries with multiple responders, we identified a lot of heterogeneity making a comparison difficult. Some consistency was found in the UK, where most respondents use unenhanced MRI and all consider unilateral tinnitus (of various durations) as only symptom an indication for MRI. Respondents that mostly use the middle fossa approach work in Germany. Based on the case scenarios it seems that respondents working in Germany, the USA and France more often proceed to microsurgery compared to other countries.

DISCUSSION

Our questionnaire regarding diagnostic and management strategies for VS identified and explored variations in clinical practice. Less than half of respondents use a guideline for diagnosis and management. Respondents apply many different strategies to select patients for MRI and apply different thresholds for treatment. The case scenarios emphasized the impact this has on choices for VS management in individual patients.

Our results are in agreement with a study that reported on variations in disease presentation and initial management of small to medium sized VSs in the USA.[12] The study described place of residence as a stronger predictor for choice of management strategy than a patient's age or VS size.[12] Next to referral patterns and availability of care, the authors contribute this to provider or institutional preference.[12] Naturally, in the current study, the proportion of patients assigned to each management strategy can be partially attributed to differences in local availability of care and referral patterns. However, our study also points out the various thresholds that otolaryngologists apply for diagnosis and management of VS, which will not depend on latter factors.

To our knowledge this is the first study to provide an overview of international differences in diagnosis and management of VS. The reliability of our data is dependent on accurate reporting of each otolaryngologist, which seems fair. However, selection of respondents cannot be precluded. Our results may therefore not comprise all strategies that are currently being used for diagnosis and management of VS. However, most included respondents work in high-volume centres and are responsible for consulting a substantial amount of VS patients. Furthermore, the variation might only further increase rather than becoming less after including more otolaryngologists. Preferably we would have achieved a higher response rate. Some non-responders motivated why they did not participate (i.e. retirement, no otolaryngologist, working in a private practice). Of the invited people, there was a larger proportion of professors that did not respond. Furthermore, we could not identify any differences between responders and non-responders. We deliberately limited our study population to otolaryngologists as we wanted to use one general questionnaire for both diagnosis and management of VS. Neurosurgeons compose an important link in the management of VS patients. However, otolaryngologists will be aware of the management of patients in their clinic/region, although they might not be treating them in person.

The lack of uniformity in diagnosis and management of VS is emphasized by the current study. The reported diagnostic work-up showed great variation, remarkably even by respondents from the same centre. The case scenarios revealed that this variation has an even greater impact on management strategies for individual patients than we had expected. Every otolaryngologist should realise that a patient might be managed differently elsewhere, even by a colleague within their own centre. We believe this information should not be neglected during patient counselling.

The current study cannot confirm whether different strategies also lead to differences in patient outcomes and costs. However, considering the extent of differences this does seem inevitable. The same patient might or might not be selected to undergo an MRI when visiting a different otolaryngologist. It is also, for example, known that facial nerve and hearing outcomes differ following radiation therapy and microsurgery.[15] It should be noted that experience with a certain treatment modality affects the results achieved. What sources of information underlie current strategies is largely unclear. Factors such as the training or reimbursements received by participants, participants' age, access to diagnostic means and treatment and involvement in research projects might influence choices made. It seems that more evidence and/or consensus could reduce uncertainties for patients as well as potential differences in outcome and associated costs. Based on the currently available evidence, it is difficult to state what diagnostic strategy should be followed.[4] Moreover, there is a lack of evidence (functional outcomes and quality of life) on different treatment modalities stratified according to VS size. Combined with the current variety in applied strategies, this constitutes a challenge to implement an (inter)national guideline. Therefore, we encourage the exchange of evidence and considerations between otolaryngologists in order to try to reach consensus on a(n) (inter)national level.

CONCLUSION

There is a high variability regarding diagnostic and management strategies for VS between otolaryngologists across countries. Otolaryngologists working in this field should realise these differences exist in order to optimize patient counselling. Further exploration of this variability may provide opportunities to synthesize the available evidence and to discuss possibilities to reach consensus.

ACKNOWLEDGEMENTS

We would like to thank all otolaryngologists who participated in this study by answering our questionnaire. Furthermore, we would like to thank the European Skull Base Society for helping us to get into contact with otolaryngologists active in the field of VS. Conflicts of interest: none.

REFERENCES

1. Fortnum H, O'Neill C, Taylor R, et al. The Role of Magnetic Resonance Imaging in the Identification of Suspected Acoustic Neuroma: A Systematic Review of Clinical and Cost Effectiveness and Natural History. *Health Technol Assess.* 2009; 13**:** iii-iv, ix-xi, 1-154.
2. Kleijwegt M, Ho V, Visser O, et al. Real Incidence of Vestibular Schwannoma? Estimations from a National Registry. *Otol Neurotol.* 2016; 37**:** 1411-1417.
3. Zealley IA, Cooper RC, Clifford KM, et al. Mri Screening for Acoustic Neuroma: A Comparison of Fast Spin Echo and Contrast Enhanced Imaging in 1233 Patients. *Br J Radiol.* 2000; 73**:** 242-247.
4. Hentschel M, Scholte M, Steens S, et al. The Diagnostic Accuracy of Non-Imaging Screening Protocols for Vestibular Schwannoma in Patients with Asymmetrical Hearing Loss and/or Unilateral Audiovestibular Dysfunction: A Diagnostic Review and Meta-Analysis. *Clin Otolaryngol.* 2017; 42**:**
5. von Kirschbaum C Gurkov R Audiovestibular Function Deficits in Vestibular Schwannoma. *Biomed Res Int.* 2016; 2016**:** Article ID 4980562.
6. Metselaar M, Demirtas G, van Immerzeel T, et al. Evaluation of Magnetic Resonance Imaging Diagnostic Approaches for Vestibular Schwannoma Based on Hearing Threshold Differences between Ears: Added Value of Auditory Brainstem Responses. *Otol Neurotol.* 2015; 36**:** 1610-1615.
7. Koors PD, Thacker LR Coelho DH Abr in the Diagnosis of Vestibular Schwannomas: A Meta-Analysis. *Am J Otolaryngol.* 2013; 34**:** 195-204.
8. Carlson ML, Habermann EB, Wagie AE, et al. The Changing Landscape of Vestibular Schwannoma Management in the United States--a Shift toward Conservatism. *Otolaryngol Head Neck Surg.* 2015; 153**:** 440-446.
9. Saeed SR, Suryanarayanan R, Dezso A, et al. Vestibular Schwannoma Management: Current Practice Amongst Uk Otolaryngologists--Time for a National Prospective Audit. *Ann R Coll Surg Engl.* 2006; 88**:** 490-495.
10. Stangerup SE Caye-Thomasen P Epidemiology and Natural History of Vestibular Schwannomas. *Otolaryngol Clin North Am.* 2012; 45**:** 257-268, vii.
11. Kirchmann M, Karnov K, Hansen S, et al. Ten-Year Follow-up on Tumor Growth and Hearing in Patients Observed with an Intracanalicular Vestibular Schwannoma. *Neurosurgery.* 2017; 80**:** 49-56.
12. Carlson ML, Glasgow AE, Grossardt BR, et al. Does Where You Live Influence How Your Vestibular Schwannoma Is Managed? Examining Geographical Differences in Vestibular Schwannoma Treatment across the United States. *J Neurooncol.* 2016; 129**:** 269-279.
13. Ciwit BV. Castor Edc, version 1.2. Amsterdam, The Netherlands.
14. Perceptieve Slechthorendheid Bij Volwassenen. Available at: https://richtlijnendatabase.nl/richtlijn/perceptieve_slechthorendheid_bij_volwassenen/perceptieve_slechthorendheid_-_startpagina.html. Accessed July 30, 2017.
15. Pollock BE, Driscoll CL, Foote RL, et al. Patient Outcomes after Vestibular Schwannoma Management: A Prospective Comparison of Microsurgical Resection and Stereotactic Radiosurgery. *Neurosurgery.* 2006; 59**:** 77-85.

SUPPLEMENTAL CONTENT

Supplemental table 3.1 Countries included in the study.

Selected countries	Number of otolaryngologists	Number of responses
Austria	1	-
Belgium	3	3
Denmark	1	1
France	6	3
Germany	70	10
Italy	6	1
New Zealand	1	1
Norway	1	-
Poland	4	1
Spain	1	-
Sweden	3	1
Switzerland	1	-
The Netherlands	11	4
United Kingdom	11	7
United States of America	10	4
Total	130	36

PART II

Diagnosis of vestibular schwannoma

CHAPTER 4

Potential savings in the diagnosis of vestibular schwannoma

Scholte M.
Hentschel M.
Hannink G.
Kunst H.
Steens S.
Rovers M.
Grutters J.

Published as:
Clin Otolaryngol. 2018 Feb;43(1):285-290. doi: 10.1111/coa.13326.

ABSTRACT

Magnetic resonance imaging (MRI) is used to screen patients at risk for vestibular schwannoma (VS). These MRIs are costly and have an extremely low yield; only 3% of patients in the screening population has an actual VS. It might be worthwhile to develop a test to predict VS, and refer only a subset of all patients for MRI.

Objective

To examine the potential savings of such a hypothetical diagnostic test before MRI.

Design

We built a decision analytical model of the diagnostic strategy of VS. Input was derived from literature and key opinion leaders. The current strategy was compared to hypothetical new strategies, assigning MRI to: (1) all patients with pathology, (2) all patients with important pathology and (3) only patients with VS. This resulted in potential cost savings for each strategy. We conducted a budget impact analysis to predict nationwide savings for the Netherlands and the United Kingdom (UK), and a probabilistic sensitivity analysis to address uncertainty.

Result

Mean savings ranged from €256 (95%CI €250 - €262) or approximately US$284 (95%CI US$277 - US$291) per patient for strategy 1 to €293 (95%CI €290 - €296) or approximately US$325 (95%CI US$322 - US$328) per patient for strategy 3. Future diagnostic strategies can cost up to these amounts per patient and still be cost saving. Annually, for the Netherlands €2.8 to €3.2 million could be saved and €10.8 to €12.3 million for the UK.

Conclusions

The model shows that substantial savings could be generated if it is possible to further optimise the diagnosis of VS.

INTRODUCTION

Vestibular schwannomas (VSs) are benign, slow growing tumours originating from Schwann cells of the vestibular part of the eighth cranial nerve.[1] They represent 6% of all intracranial tumours.[2] Patients with sporadic/unilateral VS most commonly present between their 40s and 60s, some with small intracanalicular tumours and others with larger extrameatal tumours expanding into the cerebellopontine angle (CPA).[3] Patients typically present with asymmetric sensorineural hearing loss (ASHL), asymmetric tinnitus and/or vertigo.[4]

Currently, magnetic resonance imaging (MRI) is used to screen for VS in patients visiting the otolaryngology clinic with either ASHL confirmed by pure tone audiometry (PTA), complaints of vertigo with asymmetry on electronystagmography (ENG) or asymmetric tinnitus. T1 weighted MRI of the CPA with contrast is considered the gold standard for detecting VS, and is often used in combination with high-resolution T2 of the CPA and T2 weighted MRI of the whole brain to screen for other pathology. [5,6] Since the incidence of VS is only 3% in the screening population, MRI has a very low yield.[7] Taking into account other pathologies (e.g. meningioma, demyelinating disease, etc.) that can be found through MRI in this population, approximately 14% of all MRI scans will show abnormalities.[7] The other 86% of MRIs are negative and can be considered an inefficient use of resources, and may cause unnecessary uncertainty for the patient.[8,9]

More efficient strategies could potentially avoid unnecessary MRIs in the diagnosis of VS and thus potentially increase cost-effectiveness. Before research is initiated and money is spent to develop these new strategies, it is important to assess the potential benefits these strategies might have. In this cost analysis, we examined the potential savings that new diagnostic strategies could bring, compared to the current diagnostic strategy. These analyses are performed using a decision analytical model, which compares the expected costs of potential new and current diagnostic strategies.

METHODS

We developed a decision tree of the diagnostic process, as is shown in Figure 4.1. A decision tree is a way to model decisions and their possible consequences. A fictional cohort of patients can be sent through the model. The target population for this model comprised patients at risk of VS. Patients were considered at risk when they reported asymmetric audiovestibular symptoms such as ASHL (confirmed by PTA), asymmetric tinnitus and/or vertigo (confirmed by asymmetry on ENG). The model makes use of input derived from literature and expert opinion.

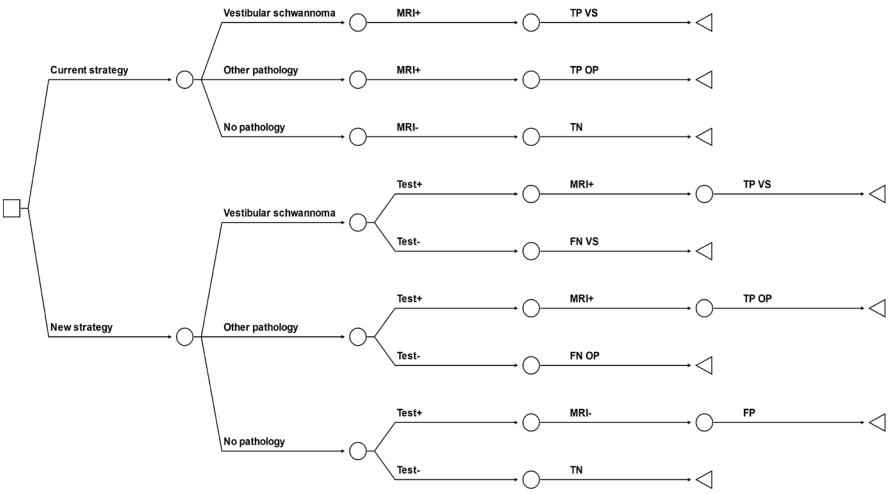


Figure 4.1 Decision tree of the diagnostic pathway. Current strategy: all patients receive MRI. New strategy: only patients with a positive test receive MRI. Other pathology was switched to other important pathology in this model for analysis of strategy 3; only VS and OIP received MRI. Non-important pathology was considered as no pathology for this analysis. FN, false negative; FP, false positive; MRI+, pathology is found on MRI; MRI-, no pathology is found on MRI; OP, other pathology; TN, true negative; TP, true positive; VS, vestibular schwannoma.

In this model, the current diagnostic strategy was compared to hypothetical new diagnostic strategies to compare the expected costs of these strategies and to determine the maximal potential cost savings of new strategies. In the current diagnostic strategy, all patients at risk of VS undergo a T2 weighted MRI examination, and in many hospitals a T1 weighted MRI with and without contrast, combined with one hospital visit. In the hypothetical new diagnostic strategies, a hypothetical test is placed before MRI investigation, making it possible to refer only a subset of patients for MRI. Patients are only referred for MRI when the hypothetical test is positive (Figure 4.1).

Possible outcomes of the diagnostic MRI were VS, other pathology (OP), other important pathology (OIP, which is a subset of OP), or no pathology. We distinguished VS patients from patients with OP and other important pathology, because this allowed us to conduct separate analyses for VS and the combination of VS and OP or OIP. Other pathology we considered as important comprised: CPA tumours, other CPA pathology (cysts, aneurysms, cholesterol granulomas, and cholesteatomas), neurofibromatosis, other malignancies of the head, arterial or venous abnormalities with consequences for treatment, congenital labyrinthine dysplasia, and demyelinating disease. Based on Dawes et al. other pathologies were considered not important, since no action was required after diagnosis.[7] Also infarction was considered a non-important pathology. In the last step of the diagnostic process, a consultation in a tertiary hospital to determine course of action for all patients with VS or OIP was included.

Strategies

We compared three hypothetical new strategies with the current diagnostic strategy (Table 4.1). We varied the population that received MRI: all patients at risk, VS and OP, VS and OIP, and VS only. We assumed that all patients with VS were correctly diagnosed in each strategy (i.e., no false positives and false negatives) and all patients were eligible for MRI. Costs of general practitioner consultations and extra diagnostic tests (PTA, ENG, etc.) were not included, since the target population comprised patients who already received general practitioner consultation and these diagnostic tests.

Table 4.1 Potential new diagnostic strategies.

Strategy	Description
Current diagnostic strategy	
1	Current diagnostic strategy: all patients at risk for VS received MRI.
Potential new strategies	
2	All patients with VS or OP received MRI.
3	Only patients with VS or OIP received MRI.
4	Only patients with VS received MRI.

MRI, magnetic resonance imaging; OIP, other important pathology; OP, other pathology; VS, vestibular schwannoma.

Probabilities

Probabilities were used to guide patients through the model, i.e. at each junction the cohort of patients was divided in two or more groups based on probabilities. When possible, we derived probabilities from literature through targeted systematic review (Table 4.2). When no evidence was available in literature, we asked key opinion leaders to provide probabilities. All expert based values were confirmed by at least two key opinion leaders (clinicians). Important probabilities were derived from Dawes et al. who reported MRI findings from a series of 1139 patients at risk of VS. The overall probability to have pathology was 14.1%, from which we distinguished the probability to have VS (3.0%). Of the remaining 11.1% with other pathology, 20.5% was considered to be important.[7]

Table 4.2 Probabilities and costs used in the decision analytical model.

Parameter	Estimated value (sensitivity analysis parameters)	Additional assumptions	Source
Probabilities			
Vestibular schwannoma	0.030 (N 1139, α 34, β 1105)		Dawes et al. 2000[7]
Other pathology	0.111 (N 1139, α 127, β 1012)	Other pathology = number of patients-number of negative scans-number of VS 1139-978-34=127	Dawes et al. 2000[7]
Other important pathology	0.023 (N 1139, α 26, β 1113)	Other important pathology=other pathology-non-important pathology 127-101=26	Dawes et al. 2000[7]
Initial referral rate to general hospitals	0.850	Patients with VS or OIP were subsequently referred to a tertiary hospital to determine course of action.	Guideline for costing research[10]
Reassessment of diagnostic MRI after consultation in a general hospital	0.225	A radiologist in a tertiary hospital reassessed the scans.	Expert opinion
Costs			
Consultation – tertiary hospital	€163		Guideline for costing research[10]
Consultation – general hospital	€80		Guideline for costing research[10]
MRI brain†	€206	Mostly comprised both T1 with and without contrast enhancement, as well as T2 MRI.	Guideline for costing research[10]
Reassessment of MRI	€87		Health care administration Radboudumc

† The guideline for costing research does not differentiate between costs of specific MRI sequences for the brain (e.g., MRI of the cerebellopontine angle).
MRI, magnetic resonance imaging; OIP, other important pathology; VS, vestibular schwannoma.

Costs

We performed the cost analysis from a healthcare perspective, restricted to direct diagnostic costs. Other costs such as out of pocket expenses, travel costs and treatment costs were not included. Costs were assessed in Euros (€) and based on the 2014 price level. For the main results, we converted Euros to US Dollars with an exchange rate of 1.109 on October 24, 2016, to increase readability. If possible, we derived unit costs (average costs incurred by producing one unit of a good or service) from the Dutch guideline for costing research.[10] Key costs in the diagnostic process were MRI scans (€206 / US$228) and consultations (€163 / US$181 for tertiary hospitals and €80 / US$89 for general hospitals).[10] To determine the maximum extra costs of a potential future diagnostic strategy at which it could still be cost-saving, we did not include potential costs of the new diagnostic test. All costs and probabilities are shown in Table 4.2.

Analysis

A hypothetical cohort of patients was sent through the model to determine mean costs of diagnosis for the current diagnostic strategy and potential new diagnostic strategies. Next, for each strategy we calculated the potential mean savings per patient: we subtracted costs of each potential new diagnostic strategy from costs of the current diagnostic strategy. To assess the potential annual savings of these strategies, we performed a budget impact analysis for the Netherlands and the United Kingdom (UK). We based our budget impact analysis on population estimates of the Netherlands and the UK, an incidence of VS of 19.4 per million in the general population and an incidence of VS 3% in the screening population [3, 7].

To address uncertainty, we conducted a probabilistic sensitivity analysis using Monte Carlo Simulation (repeated random sampling) with 1000 samples. Distributions were estimated for the probability of VS, the probability of other pathology, and the probability of other important pathology. This sensitivity analysis quantifies the level of confidence of the model's conclusions. Based on the simulation, 95% Confidence Intervals of mean costs per patient and savings were calculated using the percentile method in Excel 2007 (Microsoft).

RESULTS

The results of the decision analytical model show that the mean expected costs per patient in the current diagnostic strategy are €399 (95%CI €398 - €401) or US$442 (95%CI US$359 - US$445). In Figure 4.2 the mean diagnostic costs per patient and the mean savings per patient are stated per strategy, where the red area shows the mean costs per patient and green the mean savings per patient. A change of the initial referral rate to 100% referral to general hospitals results in expected mean savings per patient of €23 (95%CI €23 - €24) or US$26 (95%CI US$26 - US$27).

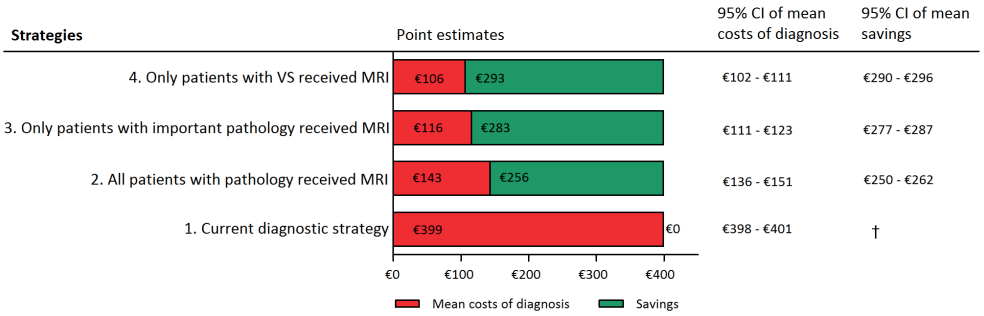


Figure 4.2 Mean costs of diagnosis and savings.
† Current diagnostic strategy, no savings. CI, confidence interval; OIP, other important pathology; OP, other pathology; VS, vestibular schwannoma.

Savings are highest in the strategy where only patients with VS receive MRI (#3). With this strategy mean expected savings per patient are €293 (95%CI €290 - €296) or US$325 (95%CI US$322 - US$328) and the cost of diagnosis will be €106 (95%CI €102 - €111) or US$118 (95%CI US$113 - US$123) per patient. It is important to keep in mind that patients with other, also important, pathology are missed in this strategy. In the strategy where besides patients with VS, also patients with other important pathology are diagnosed (#2), mean savings per patient are €283 (95%CI €277 - €287) or US$314 (95%CI US$307 - US$318) and the cost of diagnosis is €116 (95%CI €111 - €123) or US$129 (95%CI US$123 - US$136). If it is possible to distinguish between patients with and without pathology, as described in strategy 2, mean savings per patient will be €256 (95%CI €250 - €262) or US$284 (95%CI US$277 - US$291). The cost of diagnosis will be €143 (95%CI €136 - €151) or US$159 (95%CI US$151 - US$167).

Table 4.3 shows the results of the budget impact analysis. Annually, for the Netherlands €2.8 to €3.2 million (US$3.1 to US$3.5 million) could be saved and €10.8 to €12.3 million (US$12.0 to US$13.6 million) for the UK.

Table 4.3 Budget impact analysis.

	The Netherlands	United Kingdom
Annual costs (in millions)		
1. Current strategy	€4.4	€16.8
Annual potential savings per strategy (in millions)		
2. Patients with VS and OP receive MRI	€2.8	€10.8
3. Patients with VS and OIP receive MRI	€3.1	€11.9
4. Only patients with VS receive MRI	€3.2	€12.3

MRI, magnetic resonance imaging; OIP, other important pathology; OP, other pathology; VS, vestibular schwannoma.

DISCUSSION

Synopsis of key findings

The results of our model regarding the potential savings in the diagnostic pathway of VS showed that in the diagnosis of VS, €293 (95%CI €290 - €296) or US$325 (95%CI US$322 - US$328) per patient could potentially be saved when all MRIs negative for VS are avoided. A strategy where other important pathologies are also detected results in mean expected savings per patient of €283 (95%CI €277 - €287) or US$314 (95%CI US$307 - US$318). Annually, this implies that up to €3.2 million (US$ 3.5 million) can be saved in the Netherlands, and up to €12.3 million (US$ 13.6 million) in the UK.

Strengths and limitations

The major strength of this study is that it is the first study to assess the potential benefits of hypothetical diagnostic strategies for VS. This information can aid better informed decision making in research and development and can prevent investments in lower value strategies. There are multiple studies which investigate the role of MRI in the diagnostic pathway of VS and searched for an effective diagnostic algorithm to reduce the number of negative MRIs. However, none of these studies quantified the savings that might be achieved when the current diagnostic strategy would be improved.[11-22] Our results show the merits of future research to focus on optimizing these diagnostic algorithms. Furthermore, most diagnostic VS studies do not mention patients with pathology other than VS.[11,12,14-22] In our diagnostic model, patients at risk of VS who turn out to have other relevant pathology were incorporated, because these pathologies are also being detected in the current diagnostic strategy. The incorporation of other pathology allowed us to not only evaluate the diagnostic cost of detecting VS patients, but also of patients with other pathology.

Some potential limitations should also be discussed. First, our model solely used costs to calculate possible savings, whereas effects (and costs) of (missed) diagnosis were not incorporated. Currently, literature about the consequences of false negative results is lacking. We assumed that the sensitivity of MRI in the current strategy and the sensitivity of the test in new diagnostic strategies was 100%, i.e. effects would be similar for both strategies. In reality, the sensitivity of potential new tests will probably be lower than 100% (i.e., VSs and other pathologies could be missed in the future). These VSs and other pathologies might incur extra costs when they are diagnosed at a later stage, therefore the potential savings in the diagnostic strategy might be overestimated. However, we expect there will still be savings despite incorporation of costs of missed VSs, since few VSs grow rapidly. [23,24] The consequences of missing other pathology are also unclear, but could have a large impact (e.g. malignancy of the head). Since the extra costs of including OIP in the diagnosis are approximately €10 per patient, it seems worthwhile to further investigate this strategy.

Second, we assumed that the new diagnostic test does not incur extra costs. The potential value of a new strategy making use of existing tests, such as a diagnostic algorithm, could approach our current

findings. However, when an entirely new diagnostic test is used, for example a new device, extra costs will be incurred. The potential value of a new strategy will then likely be lower than current findings. If the costs of a new diagnostic test do not exceed current savings per patient this test could still be cost-saving.

Third, costs of the first steps in the diagnostic process (e.g. consultation general practitioner and diagnostic tests such as PTA and ENG) were not captured in our model, since we assumed that the potential new test would take place just before MRI. When a new test is developed that allows the general practitioner to diagnose VS, making PTA and ENG redundant in the new diagnostic strategy, the potential value of such a new test is even higher than our reported estimates.

Fourth, all costs are based on Dutch healthcare prices, and may therefore slightly differ from other countries. The same applies to probabilities based on expert opinion, which can differ between hospitals and countries. We expect small differences in exact costs for other countries, but a similar trend. Given the detailed presentation of the model and its input parameters, those interested can assess the transferability of the results to their specific situation.

Implications for clinical practice

Two important implications for daily clinical practice and research follow from the results of this study. First, although the proposed strategies in this study are hypothetical, they provide valuable insight in the potential benefits that new diagnostic strategies might have. Since calculated savings per patient are over €250 and annual savings range between €2.8 and €3.2 million for the Netherlands and €10.8 to €12.3 million for the UK, there seems to be room for the development of innovative strategies that reduce the number of MRIs.

Second, potential savings were highest when solely VS is detected. However, this does not incorporate potential extra costs of missed diagnoses. Missing other important pathology will likely have a large impact on costs, effects and treatment options, but most of all on the patient (e.g. spread of malignancies of the head) and this pathology should therefore not be missed. Hence, a potential new diagnostic strategy where all important pathology is detected will likely provide most value for money. We believe that when research is initiated into more effective strategies for diagnosing VS, these should also put effort into diagnosing other important pathology.

In conclusion, the model shows that substantial savings could be generated if we would be able to further optimise the diagnosis of patients suspected of VS.

CONFLICTS OF INTEREST

There are no conflicts of interest.

ACKNOWLEDGEMENTS

The authors thank Jef Mulder for his valuable input and expert opinion. This study was funded by a Health Care Efficiency Research grant by The Netherlands Organization for Health Research and Development (ZonMw).

REFERENCES

1. Fortnum, H., O'Neill, C., Taylor, R., et al. The role of magnetic resonance imaging in the identification of suspected acoustic neuroma: a systematic review of clinical and cost effectiveness and natural history. *Health Technol Assess. 2009;13*(18):iii-iv, ix-xi, 1-154.
2. Lin, D., Hegarty, J. L., Fischbein, N. J., et al. The prevalence of "incidental" acoustic neuroma. *Arch Otolaryngol Head Neck Surg.* 2005;*131*(3);241-244.
3. Stangerup, S. E., Tos, M., Thomsen, J., & Caye-Thomasen, P. True incidence of vestibular schwannoma? *Neurosurgery.* 2010;*67*(5):1335-1340.
4. Matthies, C., & Samii, M. Management of 1000 vestibular schwannomas (acoustic neuromas): clinical presentation. *Neurosurgery.* 1997;*40*(1):1-9.
5. Press, G. A., & Hesselink, J. R. MR imaging of cerebellopontine angle and internal auditory canal lesions at 1.5 T. *AJR Am J Roentgenol.* 1988;1*50*(6):1371-1381.
6. Sriskandan, N., & Connor, S. E. The role of radiology in the diagnosis and management of vestibular schwannoma. *Clin Radiol.* 2011;*66*(4):357-365.
7. Dawes, P. J., Mehta, D., & Arullendran, P. Screening for vestibular schwannoma: magnetic resonance imaging findings and management. *J Laryngol Otol.* 2000;*114*(8):584-588.
8. Moynihan, R., & Smith, R. Too much medicine? *BMJ.* 2002;*324*(7342):859-860.
9. Glasziou, P., Moynihan, R., Richards, T., et al. Too much medicine; too little care. *BMJ.* 2013;*347*:f4247.
10. Zorginstituut Nederland. Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg, February 29, 2016 version. Available at: https://www.zorginstituutnederland.nl/over-ons/publicaties/publicatie/2016/02/29/richtlijn-voor- het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg (in Dutch). Accessed March 16, 2016.
11. Chatrath, P., Frosh, A., Gore, A., et al. Identification of predictors and development of a screening protocol for cerebello-pontine lesions in patients presenting with audio-vestibular dysfunction. *Clin Otolaryngol.* 2008;*33*(2):102-107.
12. Cheng, T. C., & Wareing, M. J. Three-year ear, nose, and throat cross sectional analysis of audiometric protocols for magnetic resonance imaging screening of acoustic tumours. *Otolaryngol Head Neck Surg.* 2012;*146*(3):438-447.
13. Cueva, R. A. Auditory brainstem response versus magnetic resonance imaging for the evaluation of asymmetric sensorineural hearing loss. *Laryngoscope.* 2004;*114*(10):1686-1692.
14. Gimsing, S. Vestibular schwannoma: when to look for it? *J Laryngol Otol.* 2010; *124*(3):258-264.
15. Mandala, M., Giannuzzi, A., Astore, S., et al. Hyperventilation-induced nystagmus in vestibular schwannoma and unilateral sensorineural hearing loss. *Eur Arch Otorhinolaryngol.* 2013;*270*(7):2007-2011.
16. Obholzer, R. J., Rea, P. A., & Harcourt, J. P. Magnetic resonance imaging screening for vestibular schwannoma: analysis of published protocols. *J Laryngol Otol.* 2004;*118*(5):329-332.
17. Saliba, I., Bergeron, M., Martineau, G., et al. Rule 3,000: a more reliable precursor to perceive vestibular schwannoma on MRI in screened asymmetric sensorineural hearing loss. *Eur Arch Otorhinolaryngol.* 2011;*268*(2):207-212.
18. Scherler, M., & Bohmer, A. The value of clinical examination methods in diagnosis of acoustic neuroma. *HNO.* 1995;*43*(8):487-491.
19. Suzuki, M., Hashimoto, S., Kano, S., et al. Prevalence of acoustic neuroma associated with each

configuration of pure tone audiogram in patients with asymmetric sensorineural hearing loss. *Ann Otol Rhinol Laryngol.* 2010;*119*(9):615-618.

20. Vandervelde, C., & Connor, S. E. Diagnostic yield of MRI for audiovestibular dysfunction using contemporary referral criteria: correlation with presenting symptoms and impact on clinical management. *Clin Radiol.* 2009;*64*(2):156-163.

21. Raber, E., Dort, J. C., Sevick, R., et al. Asymmetric hearing loss: toward cost-effective diagnosis. *J Otolaryngol.* 1997;*26*(2):88-91.

22. Rupa, V., Job, A., George, M., et al. Cost-effective initial screening for vestibular schwannoma: auditory brainstem response or magnetic resonance imaging? *Otolaryngol Head Neck Surg.* 2003;*128*(6):823-828.

23. Ferri, G. G., Modugno, G. C., Pirodda, A., et al. Conservative management of vestibular schwannomas: an effective strategy. *Laryngoscope.* 2008;*118*(6):951-957.

24. Hajioff, D., Raut, V. V., Walsh, R. M., et al. Conservative management of vestibular schwannomas: third review of a 10-year prospective study. *Clin Otolaryngol.* 2008;*33*(3):255-259.

CHAPTER 5

Diagnostic accuracy of high resolution T2-weighted MRI versus contrast enhanced T1-weighted MRI to screen for cerebellopontine angle lesions in symptomatic patients

Hentschel M.
Kunst H.
Rovers M.
Steens S.

Published as:
Clin Otolaryngol. 2018;43(3):805-811. doi: 10.1111/coa.13051.

ABSTRACT

Objective

To evaluate diagnostic accuracy of high resolution T2-weighted MRI (T2w) for detecting cerebellopontine angle (CPA) lesions compared to a combined protocol including gadolinium enhanced T1-weighted MRI (GdT1w).

Setting

Department of Radiology & Nuclear Medicine, Radboud University Medical Centre, Nijmegen, the Netherlands.

Participants

A random sample of MRIs from 350 patients (700 CPAs) with asymmetrical audiovestibular complaints was used, acquired between 2013-2016.

Main outcome measures

Sensitivity, specificity, positive and negative predictive values of T2w results compared to GdT1w and, in patients with any suggestion of CPA pathology, to the complete examination (T1w, GdT1w and T2w). Inter-rater agreement between an experienced neuroradiologist and a less experienced observer was calculated.

Results

Results of 678 CPAs in 340 patients were analysed. On T2w the neuroradiologist identified all 27 lesions >2 mm in size out of a total of 30 CPA lesions (sensitivity: 90% [95%CI:73.5-97.9%]). Negative predictive value reached 99.5% [95%CI:98.7-99.9]. One missed lesion of 2 mm would have been detected in clinical practice, as this was one of 14 patients for which additional GdT1w would have been ordered based on T2w alone, increasing sensitivity to 93% [95%CI:77.9-99.2%] and negative predictive value to 99.7% [95%CI:98.9-100%]. Inter-rater agreement for T2w was 98% [95%CI:96.4-98.8].

Conclusion

T2w has a very high diagnostic accuracy for the presence of CPA lesions in patients with asymmetrical audiovestibular complaints. However, in a screening protocol with T2w only, smallest vestibular schwannomas as well as rare differential diagnoses that probably only would be detected on GdT1w may remain unnoticed.

INTRODUCTION

In a patient with asymmetrical hearing loss, unilateral tinnitus and/or vestibular symptoms, the otolaryngologist will order an MRI examination to evaluate the presence of a cerebellopontine angle (CPA) lesion. The most common CPA lesion is a vestibular schwannoma (VS). However, with an incidence of VS of around 1 to 3.3 per 100,000 people[1-4], the overall yield of these MRIs is 3-4%.[2, 5] International guidelines discussing the optimal MRI protocol in these patients are lacking. Currently, MRI protocols to screen for CPA lesions in patients with asymmetrical audiovestibular complaints differ per hospital, either consisting of high resolution T2-weighted MRI (T2w) alone, or a combination with contrast enhanced T1-weighted MRI (GdT1w). Including GdT1w increases scan time compared to T2w only. Moreover, an intravenous catheter and administration of a contrast agent are required, resulting in discomfort for patients, a risk of allergic reactions, and nephrogenic systemic fibrosis or accumulation in the brain of certain types of gadolinium contrast agents, and increased costs for society (e.g. Dotarem® (Gadoterate meglumine, Guerbet, Villepinte, France) has a list price (commercial price) of about €80 or approximately $86 per 15 ml vial).[6-8] Studies reporting on diagnostic accuracy of MRI protocols only consisting of T2w, on the other hand, are either outdated (lacking current state-of-the-art image quality), contain small patient samples and/or mainly positive MRI findings, or are conducted in academic setting where specialized neuroradiologist assess images.[9-13] The consequences of usage of a protocol consisting of T2w only in current clinical practice remains unclear. Our aim was to evaluate the diagnostic accuracy of T2w for the presence of CPA lesions in comparison to a combined protocol including GdT1w in a large cohort of patients with asymmetrical audiovestibular complaints, using standard MRI machines and observers with different levels of experience.

MATERIALS AND METHODS

This study was performed alongside a diagnostic (cost-)effectiveness study in which we aim to optimize the diagnostic strategy, including imaging, in patients with asymmetrical audiovestibular complaints. The need for informed consent was waived by the medical ethics committee of our institution, because of the size of the study population and retrospective nature of the study.

Population

In our tertiary hospital (Radboud University Medical Centre, Nijmegen, the Netherlands) it is policy to perform a screening MRI of the CPA region in patients with asymmetrical tinnitus, asymmetrical sensorineural hearing loss (confirmed by pure-tone audiometry), and/or vestibular symptoms (in conjunction with asymmetrical findings on electronystagmography). We used a random sample of 350 patients aged ≥16 years, for which an otolaryngologist had ordered a "screening MRI CPA", i.e. 700 CPAs. All MRIs were acquired between January 2013 and April 2016. This recent time period was chosen to ascertain state-of-the-art quality of the included images, representative for current clinical practice.

Data collection

Most MRI images were acquired using a Siemens Avanto (Siemens Healthcare GmbH, Erlangen, Germany) with a field strength of 1.5 Tesla. Eight patients were scanned using a Siemens TrioTim with 3 Tesla field strength. For each patient both axial GdT1w of the CPA region (Avanto and TrioTim standard parameters: slice thickness 2 and 2 mm, spacing 2.2 and 2.2 mm, TR 400-500 and 700-795 ms, TE 17 and 11 ms, flip angle 90° and 120°, respectively) and 3D TSE T2w (Avanto and TrioTim standard parameters: slice thickness 0.5 and 0.5 mm, TR 1500 and 1500 ms, TE 296-297 and 301 ms, flip angle 170° and 170°, respectively) were obtained in the same session. The screening MRI of the CPA additionally includes a 5 mm T2w of the whole brain, however abnormal findings with a location other than the CPA were outside the scope of the current study. MRI images of the selected patients were pseudo-anonymized (coded) and imported in a specially developed survey environment (Cirrus, a scoring and viewing platform produced in-house, based on MeVisLab (MeVis Medical Solutions AG, Germany)).

Two observers, a radiologist specialized in neuro- and head & neck radiology with 6 years of experience [SS], and a medical doctor without formal education in neuroradiology [MH] assessed all images independently of each other using a predefined form. The less experienced observer was instructed on CPA anatomy and pathology by the neuroradiologist before the onset of the study. Both observers were blinded to any patient characteristics, clinical information, and original reports during data collection. The presence (yes or no) and most probable diagnosis of lesions located in the CPA (including internal auditory canal and/or vestibulocochlear system) were noted, first for all consecutive T2w, followed by all consecutive GdT1w images without access to T2w diagnoses and images, in separate sessions. In all patients in which one or both observers had reported a lesion, an equivocal finding or had expressed doubts about a diagnosis based on T2w and/or GdT1w, the same neuroradiologist additionally assessed the combination of T1w, GdT1w and T2w to establish a final diagnosis. In case T2w and GdT1w were independently assessed as being normal by both observers, the chance of finding pathology on the combination of GdT1w, T2w and T1 was considered negligible. Following assessment of each T2w image, the neuroradiologist additionally registered whether he would have wanted the patient to be recalled for additional GdT1w, because of doubts about T2w diagnosis, if an imaging policy with T2w only would have been standard. This provides insight in the number and type of patients that would need to return for acquisition of GdT1w in a screening policy consisting of T2w only.

If quality of the images was judged insufficient to reliably assess one or both CPAs by one or both observers (e.g. due to movement artefacts) this CPA was excluded. Maximal tumour diameter in any direction was measured for detected lesions (including internal auditory canal portion).

Analyses

All statistical analyses were performed in R version 3.2.3 (The R Foundation for Statistical Computing, Vienna, Austria, 2015), using the 'epiR' and 'rel' packages. Presence or absence of a CPA lesion was used as outcome, irrespective of the type of lesion. T2w findings (index test) were initially compared to GdT1w

alone (reference test). In case of positive, equivocal or doubtful findings on T2w and/or GdT1w, the combination of T1w, GdT1w and T2w (reference test) was additionally assessed to confirm the diagnosis. We summarized this in 2 by 2 tables and calculated sensitivity, specificity and positive and negative predictive values. Moreover, we calculated the proportion of patients that would have been recalled by the neuroradiologist for additional GdT1w images if T2w only was standard imaging policy.

Since CPA lesions are rare findings, even in symptomatic patients such as used in this study, we used Gwet's AC1 as coefficient to evaluate inter-rater agreement for T2w. In contrary to intraclass correlation coefficient or Cohen's Kappa, this outcome value is less affected by heterogeneity in the study population. Moreover, it adjusts for the fact that assessors may agree, even when giving random values as results.[14, 15] Gwet's AC1 can range from 0 to 1; where 1 indicates perfect reliability and 0 no reliability at all.

RESULTS

Of the 700 CPAs in 350 patients, 22 CPAs in 12 patients were excluded: in 7 patients the whole set of images was not available in the Cirrus system and in 3 patients at least one of the observers reported insufficient quality of the images (the experienced observer reported insufficient quality of 1 T2w and 1 GdT1w sequence, and in 1 patient both T2w and GdT1w were reported insufficient by both observers), resulting in exclusion of both CPAs, and in 2 patients unilateral artefacts (identified by both observers) resulted in exclusion of one CPA. Results of 678 CPAs in 340 patients remained for further analysis.

Diagnoses

Based on the reference test, the neuroradiologist ultimately identified 30 CPA lesions (4.4% of CPAs) in 28 patients (8.2% of patients). VS as most probable diagnosis occurred most often, with 16 and 10 VSs diagnosed on the right and left side, respectively, of which 2 patients had a bilateral VS. There was 1 patient with a meningioma and 1 with an arachnoid cyst in the right and left CPA, respectively. One patient had a lesion of unknown origin on the left side and 1 had an absent signal on T2w in the basal and 2nd turn of the right cochlea of which the origin could not be defined.

Diagnostic accuracy

A 2 by 2 table containing the results of the neuroradiologist, as well as sensitivity and specificity and positive and negative predictive values with corresponding 95% CIs is displayed in Table 5.1. The neuroradiologist identified all 27 CPA lesions >2 mm (90%) correctly on T2w alone when compared to the 30 CPA lesions identified on the reference test.

There were 14 patients (4%) for which the neuroradiologist would have ordered additional GdT1w to confirm/exclude the presence of a CPA lesion with more certainty after assessment of T2w, i.e. these patients would have been recalled in clinical practice with an imaging policy of T2w only. One of the

3 false negative observations on T2w would be additionally detected in clinical practice after being recalled for GdT1w (Table 5.1). Including GdT1w results in the index test in these 14 selected patients increased diagnostic accuracy: 1 false negative diagnosis was reclassified as true positive and 1 false positive diagnosis was reclassified as true negative, finally resulting in 2 false negative and 2 false positive diagnoses. Hence, the number of correctly diagnosed CPA lesions (sensitivity) increased to 28 (93.3% [95%CI: 77.9-99.2%]) and the number of correctly classified negative MRIs (specificity) to 646 (99.7% [95%CI: 98.9-100.0%]), resulting in a positive and negative predictive value of 93.3% [95%CI: 77.9-99.2%] and 99.7% [95%CI: 98.9-100%], respectively. In the remaining 326 patients (95.9%) GdT1w was not deemed necessary, because the neuroradiologist indicated to be certain about diagnoses based on T2w alone. Inter-rater agreement (absence/presence of CPA lesion) on T2w between both observers, expressed as Gwet's AC1 constant, reached 0.976 [95%CI: 0.964-0.988].

CPA lesions

Maximal lesion diameter varied from 2-23 mm (including internal auditory canal portion). All lesions >2 mm were correctly identified on T2w alone by the neuroradiologist. The 3 lesions that were missed on T2w alone compared to the reference test were 2 mm and located distally in the internal auditory canal. The less experienced observer did not identify these 3 lesions either. T2w and GdT1w images of the 3 lesions are provided in Figures 5.1a and b. For the patient shown in Figure 5.1a the neuroradiologist indicated T2w should be followed by additional GdT1w, i.e. the patient would have been recalled. This lesion would have been detected on the combination of sequences and thus would not have been missed in clinical practice. Figure 5.1b displays the 2 lesions that were missed on T2w compared to the reference test. None of the latter 2 lesions are visible on T2w in retrospect. All 3 false positive lesions on T2w were located in the cochlea, of which an example is displayed in Figure 5.2. In one of these patients, following T2w assessment, the neuroradiologist indicated the patient should be recalled for additional GdT1w, i.e. the presence of a lesion would eventually have been excluded on the combination of sequences. Another lesion in the CPA was correctly detected by both observers based on T2w (true positive), but was classified as a VS and appeared to be a meningioma on GdT1w (Supplemental figure 5.1).

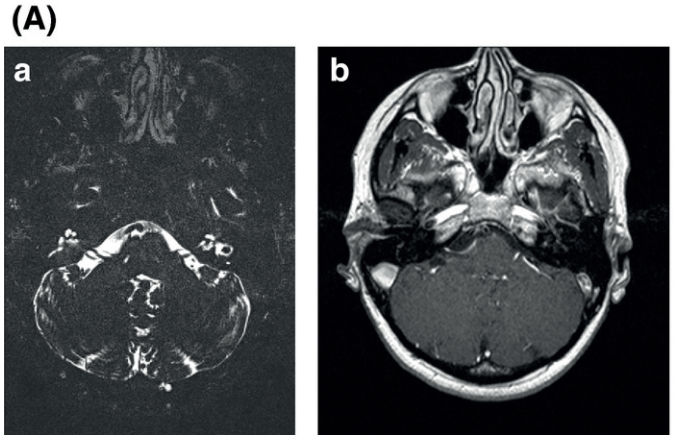


Figure 5.1A Example of an initially false negative diagnosis of a left sided intracanalicular VS of 2 mm based on T2w (a) compared to GdT1w (b). Findings on T2w would have been a reason to request additional GdT1w, so this lesion would eventually have been diagnosed correctly.

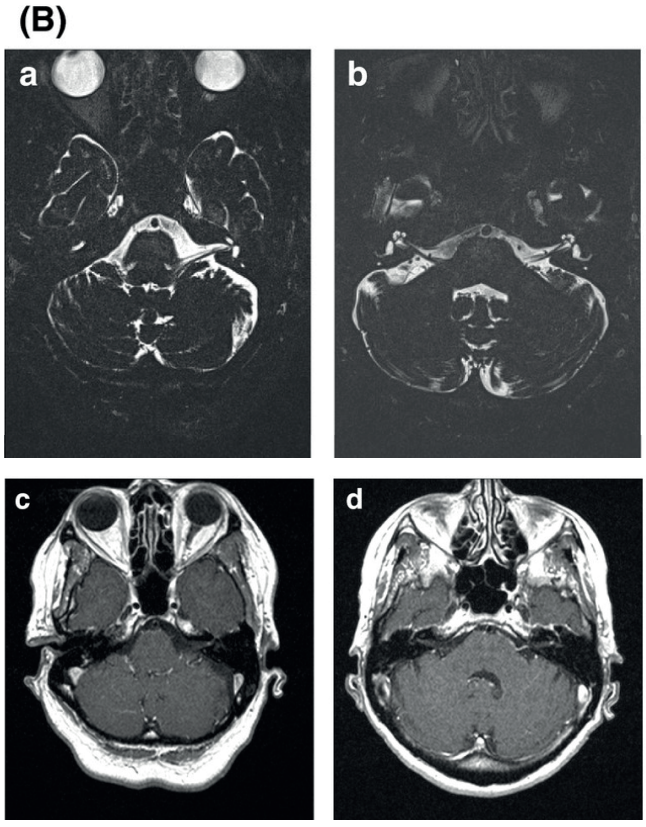


Figure 5.1B Example of two false negative diagnoses of left sided intracanalicular VSs of 2 mm based on T2w (a and b) compared to GdT1w (c and d). Even in retrospect, these lesions could not be detected on T2w.

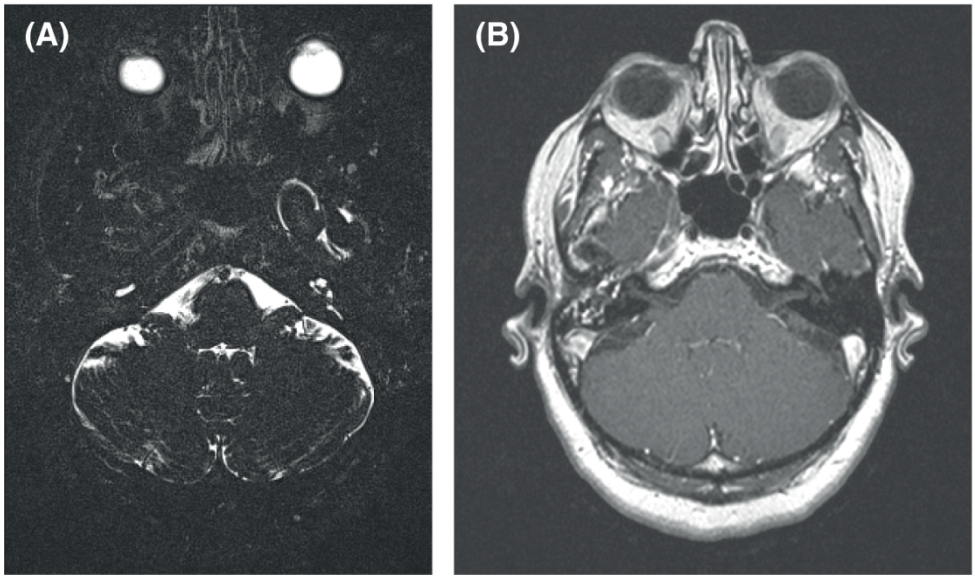


Figure 5.2 Example of a false positive finding in the left cochlea diagnosed as VS based on T2w (A), while GdT1w revealed no enhancing structures (B). Findings on T2w would have been a reason for additional GdT1w, so the absence of a VS would ultimately have been correctly confirmed.

DISCUSSION

Our results show a very high negative predictive value for the presence of CPA lesions in patients with asymmetrical audiovestibular complaints using T2w alone, and a very high inter-rater agreement between an experienced and less experienced observer. A protocol consisting of T2w only would result in 2 (out of 30) VSs being missed, having a maximal diameter of 2 mm.

Previous studies

Our results are not completely in agreement with Fortnum et al. who also compared diagnostic accuracy of high resolution T2w to GdT1w in diagnosing VS. They pooled estimates of 11 papers published between 1996 and 2001 and reported sensitivity and specificity values of 98% and 96%, respectively.[16] Specificity in the current study was higher with 99.5%, while sensitivity was lower with 90.0%. Possible explanations are the choice not to blind for clinical details[9, 10, 13, 17], usage of different sequences[18, 19] and field strength[10, 19], and a different prevalence[10, 13, 17-20] and size[9, 10, 17] of CPA lesions in some of the studies included in the meta-analysis by Fortnum et al.[16] In the current study lesions not identified on T2w were 2 mm in size. In several studies within the meta-analysis diameter of the smallest detected lesions was ≥3 mm.[9, 10, 17] Possibly, lesions of 2 mm did not occur in these studies or were not detected.

The prevalence of CPA lesions in 8.2% of patients (including VSs in 7% of all patients) in our study was relatively high, which influences positive and negative predictive values.[21] The higher incidence of CPA lesions in our study as compared to literature is probably caused by the fact that we selected our sample based on MRI coding. Due to the tertiary study setting, our sample might include some patients with a CPA lesion referred from another hospital, in which a new screening MRI was acquired in our institution.

Strengths and limitations

We carefully chose the time frame to guarantee similar and state-of-the-art quality MRIs, which makes results representative for current clinical practice. The number of included CPAs was high. It is unlikely that insufficient GdT1w quality in few patients has introduced bias; the index test (T2w images) was first scored in a separate session and could not be linked to a particular GdT1w image during scoring. Only 1 patient was excluded, because of insufficient quality of T2w (assessed by experienced observer).

The current study was performed in a tertiary hospital, in which radiologists reporting on CPA screening studies are specialized in neuro- and head & neck radiology. However, the majority of patients with asymmetrical audiovestibular complaints will be screened in a general hospital, and images will usually be assessed by general radiologists who have less exposure and experience. We have shown that an experienced and less experienced observer have high agreement with respect to T2w assessment. Agreement between a neuroradiologist and general radiologist is expected to be at least as sufficient and probably even higher. Moreover, we used Gwet's AC1 to determine the agreement between assessors, because it is less affected by the prevalence of a disease, which is relatively low in case of CPA lesions.[14]

Some limitations should also be discussed. First, clinical information was not included in the analysis. In practice, such information will be provided and will focus the radiologist's attention to one CPA in particular. This might have caused an underestimation of diagnostic accuracy and an overestimation of the simulated recall rate for additional GdT1w of 4%. It should be noted that a radiologist should always assess both CPAs, since incidentalomas are not uncommon in the CPA.[22] Second, we have used the presence/absence of a CPA lesion as outcome, instead of the (most probable) type of lesion. Both observers misclassified one lesion for VS instead of meningioma (Supplemental figure 5.1). Consequences of such misdiagnoses are limited, since management strategies are usually identical.[23] Third, GdT1w is often considered gold standard to diagnose CPA lesions.[16] Some types of lesions are, however, better visualized using T2w images. False positive diagnoses on T2w (compared to GdT1w) are not necessarily absent in truth. Both observers identified an arachnoid cyst on T2w images but missed the same lesion on the GdT1w sequence (Supplemental figure 5.2). According to clinical practice, we included additional T1w and T2w images to GdT1w as reference standard, whenever T2w or GdT1w showed any (suspicion of) abnormality.

Clinical implications

Based on the neuroradiologist's assessment of T2w alone, 14 out of 340 patients would have been recalled for additional GdT1w. After including results of additional GdT1w examinations in the index test for the latter patients, as would occur in clinical practice, 2 lesions of 2 mm would have been missed. In other words, to prevent these 2 CPA lesions of 2 mm from being missed, 326 patients have to undergo GdT1w, with accessory costs (€26,080 or approximately US$ 27,929 when accounting for the official price of €80 per unit Dotarem®), discomfort and risks. In clinical practice, patients with lesions ≤2 mm are managed using a wait and scan policy, instead of invasive treatments such as radiotherapy or microsurgery.[24] Of the VS patients in a wait and scan policy, no more than 40% require treatment at some point.[25-28] Cisternal extension and a large tumour diameter at time of diagnosis have been mentioned as significant predictors for growth[26, 29], which are not applicable to the small missed lesions in this study. On the contrary, at the moment there is no safety net for patients who have a small missed lesion that may require treatment once growth occurs later in life. As it appears that hearing loss deteriorates faster in case of growing intracanalicular tumours as compared to non-growing lesions and the contralateral ear[27, 28], thorough patient instructions and a relatively simple test such as pure-tone audiometry which is already routinely acquired in these patients could possibly identify such patients at a later time and so may serve as a safety net. The general practitioner could potentially play a larger role in such a scenario. This should be further investigated (including analyses on cost-efficiency) and translated into a guideline for usage in clinical practice.

Decisions about management also determine further steps regarding imaging. Given the results of this study and above-named arguments, specifically the high negative predictive value of T2w alone for VSs >2 mm, one could argue to restrain from further action when T2w is normal. Whenever a CPA lesion is evident, and the patient will be assigned to a wait and scan policy, additional GdT1w may not be required at all, especially when the lesion is typical for VS. Omitting GdT1w in follow-up of VSs was recently shown to be cost-effective and is policy in our hospital.[30] MRI policies for postsurgical patients were outside the scope of the current study.

We have shown that T2w alone has a high sensitivity and specificity for the presence of CPA lesions. Although intracochlear VSs were not encountered in our study population, we know from experience that even small intracochlear VSs can be detected by scrutinizing a high-quality T2w.[31] When in doubt patients can be recalled for additional GdT1w. Another differential diagnosis that justifies additional GdT1w is labyrinthitis, which can be visualized as a filling defect on T2w.[31] Intralabyrinthine haemorrhage can easily be diagnosed with aid of native T1w, which we would recommend to acquire in all patients.[31] Diagnoses such as Lyme or tuberculosis cannot be established based on T2w alone, but are seldom and symptoms are rarely restricted to audiovestibular complaints.[31, 32] GdT1w is also required to detect leptomeningeal metastases, usually prevalent in patients with a known history of malignancy, although these lesions are also seldom restricted to the CPA.[11, 31] One should note that we did not encounter any patients with granulomatous, inflammatory or other infectious pathology in our unselected study

population, nor any patient with leptomeningeal metastases. A large series of GdT1w images of the CPA studied by Dawes et al. only revealed 2 cases of labyrinthine inflammation out of 1139 scans (0.18%), which did not influence management.[5] This indicates the rareness of differential diagnoses that cannot be seen on T2w. However, if one considers a policy of screening with T2w alone in patients with audiovestibular complaints, GdT1w should still be included when screening patients with a broader variety of symptoms and in those with a known history of inflammatory, infectious or malignant disorders.[11] Preferably, differences in MRI protocols between hospitals are eliminated by establishment of a guideline weighing benefits of improved diagnostic accuracy of GdT1w against its disadvantages of increased scan time, costs and potential harmful effects.

In conclusion, T2w alone has a high diagnostic accuracy for detection of CPA lesions >2 mm. Further research is needed focusing on a (clinical) safety net for the few patients with very small lesions being missed on T2w alone and patients with rare differential diagnoses. At least, GdT1w should always be considered in patients with a broader variety of symptoms, or a known history of inflammatory, infectious or malignant disorders.

ACKNOWLEDGEMENTS

This study was funded by a Health Care Efficiency Research grant by The Netherlands Organization for Health Research and development (ZonMw). The need for informed consent was waived by the medical ethics committee of our institution, because of the size of the study population and retrospective nature of the study. The authors thank Sven Lafebre for designing the Cirrus software.

REFERENCES

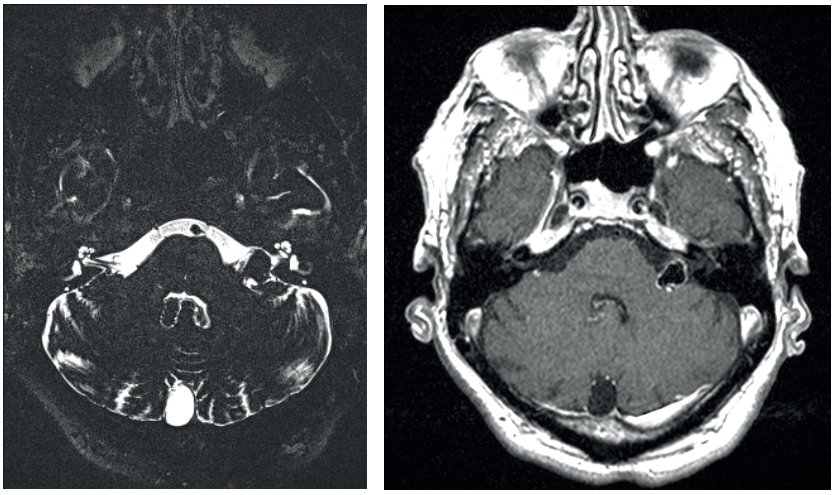
1. Kshettry VR, Hsieh JK, Ostrom QT, Kruchko C, Barnholtz-Sloan JS. Incidence of vestibular schwannomas in the United States. JNO. 2015;124(2):223-8. doi: 10.1007/s11060-015-1827-9.
2. Kleijwegt M, Ho V, Visser O, Godefroy W, van der Mey A. Real Incidence of Vestibular Schwannoma? Estimations From a National Registry. Otol Neurotol. 2016;37(9):1411-7. doi: 10.1097/MAO.0000000000001169.
3. Stangerup SE, Tos M, Thomsen J, Caye-Thomasen P. True incidence of vestibular schwannoma? Neurosurgery. 2010;67(5):1335-40. doi: 10.1227/NEU.0b013e3181f22660.
4. Evans DG, Moran A, King A, Saeed S, Gurusinghe N, Ramsden R. Incidence of vestibular schwannoma and neurofibromatosis 2 in the North West of England over a 10-year period: higher incidence than previously thought. Otol Neurotol. 2005;26(1):93-7.
5. Dawes PJ, Mehta D, Arullendran P. Screening for vestibular schwannoma: magnetic resonance imaging findings and management. J Laryngol Otol. 2000;114(8):584-8.
6. Sadowski EA, Bennett LK, Chan MR, Wentland AL, Garrett AL, Garrett RW, et al. Nephrogenic systemic fibrosis: risk factors and incidence estimation. Radiology. 2007;243(1):148-57. doi: 10.1148/radiol.2431062144.
7. Costello JR, Kalb B, Martin DR. Incidence and Risk Factors for Gadolinium-Based Contrast Agent Immediate Reactions. TMRI. 2016;25(6):257-63. doi: 10.1097/RMR.0000000000000109.
8. Kanda T, Fukusato T, Matsuda M, Toyoda K, Oba H, Kotoku J, et al. Gadolinium-based Contrast Agent Accumulates in the Brain Even in Subjects without Severe Renal Dysfunction: Evaluation of Autopsy Brain Specimens with Inductively Coupled Plasma Mass Spectroscopy. Radiology. 2015;276(1):228-32. doi: 10.1148/radiol.2015142690.
9. Annesley-Williams DJ, Laitt RD, Jenkins JP, Ramsden RT, Gillespie JE. Magnetic resonance imaging in the investigation of sensorineural hearing loss: is contrast enhancement still necessary? J Laryngol Otol. 2001;115(1):14-21.
10. Zealley IA, Cooper RC, Clifford KM, Campbell RS, Potterton AJ, Zammit-Maempel I, et al. MRI screening for acoustic neuroma: a comparison of fast spin echo and contrast enhanced imaging in 1233 patients. BJR. 2000;73(867):242-7. doi: 10.1259/bjr.73.867.10817038.
11. Abele TA, Besachio DA, Quigley EP, Gurgel RK, Shelton C, Harnsberger HR, et al. Diagnostic accuracy of screening MR imaging using unenhanced axial CISS and coronal T2WI for detection of small internal auditory canal lesions. AJNR Am J Neuroradiol. 2014;35(12):2366-70. doi: 10.3174/ajnr.A4041.
12. Maslan JT, Lack CM, Zapadka M, Gasser TG, Oliver E. High resolution T2 MRI in the diagnosis of cerebellopontine angle and internal auditory canal lesions. Clin Imaging. 2017;45:8-11. doi: 10.1016/j.clinimag.2017.05.009.
13. Soulie D, Cordoliani YS, Vignaud J, Cosnard G. MR imaging of acoustic neuroma with high resolution fast spin echo T2-weighted sequence. EJR. 1997;24(1):61-5.
14. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. BMC. 2013;13:61. doi: 10.1186/1471-2288-13-61.

15. Ten Cate DF, Luime JJ, Hazes JM, Jacobs JW, Landewe R. Does the intraclass correlation coefficient always reliably express reliability? Comment on the article by Cheung et al. Arthritis Care Res. 2010;62(9):1357-8. doi: 10.1002/acr.20255.
16. Fortnum H, O'Neill C, Taylor R, Lenthall R, Nikolopoulos T, Lightfoot G, et al. The role of magnetic resonance imaging in the identification of suspected acoustic neuroma: a systematic review of clinical and cost effectiveness and natural history. Health technol Assess. 2009;13(18):iii-iv, ix-xi, 1-154. doi: 10.3310/hta13180.
17. Marx SV, Langman AW, Crane RC. Accuracy of fast spin echo magnetic resonance imaging in the diagnosis of vestibular schwannoma. Am J of Otolaryngol. 1999;20(4):211-6.
18. Stuckey SL, Harris AJ, Mannolini SM. Detection of acoustic schwannoma: use of constructive interference in the steady state three-dimensional MR. AJNR Am J Neuroradiol. 1996;17(7):1219-25.
19. Hermans R, Van der Goten A, De Foer B, Baert AL. MRI screening for acoustic neuroma without gadolinium: value of 3DFT-CISS sequence. Neuroradiology. 1997;39(8):593-8.
20. Allen RW, Harnsberger HR, Shelton C, King B, Bell DA, Miller R, et al. Low-cost high-resolution fast spin-echo MR of acoustic schwannoma: an alternative to enhanced conventional spin-echo MR? AJNR Am J Neuroradiol. 1996;17(7):1205-10.
21. Altman DG, Bland JM. Diagnostic tests. 1: Sensitivity and specificity. BMJ. 1994;308(6943):1552.
22. Lin D, Hegarty JL, Fischbein NJ, Jackler RK. The prevalence of "incidental" acoustic neuroma. Arch Otolaryngol Head Neck Surg. 2005;131(3):241-4. doi: 10.1001/archotol.131.3.241.
23. Friedmann DR, Grobelny B, Golfinos JG, Roland JT, Jr. Nonschwannoma tumors of the cerebellopontine angle. Otolaryngol Clinics North Am. 2015;48(3):461-75. doi: 10.1016/j.otc.2015.02.006.
24. Hajioff D, Raut VV, Walsh RM, Bath AP, Bance ML, Guha A, et al. Conservative management of vestibular schwannomas: third review of a 10-year prospective study. Clinical otolaryngol. 2008;33(3):255-9. doi: 10.1111/j.1749-4486.2008.01705.x.
25. Yoshimoto Y. Systematic review of the natural history of vestibular schwannoma. J Neurosurg. 2005;103(1):59-63. doi: 10.3171/jns.2005.103.1.0059.
26. Hunter JB, Francis DO, O'Connell BP, Kabagambe EK, Bennett ML, Wanna GB, et al. Single Institutional Experience With Observing 564 Vestibular Schwannomas: Factors Associated With Tumor Growth. Otol Neurotol. 2016;37(10):1630-6. doi: 10.1097/MAO.0000000000001219.
27. Kirchmann M, Karnov K, Hansen S, Dethloff T, Stangerup SE, Caye-Thomasen P. Ten-Year Follow-up on Tumor Growth and Hearing in Patients Observed With an Intracanalicular Vestibular Schwannoma. Neurosurgery. 2017;80(1):49-56. doi: 10.1227/NEU.0000000000001414.
28. van Linge A, Borsboom GJ, Wieringa MH, Goedegebure A. Hearing Loss Progresses Faster in Patients With Growing Intracanalicular Vestibular Schwannomas. Otol Neurotol. 2016;37(9):1442-8.
29. Wolbers JG, Dallenga AH, van Linge A, Te West M, Kummer EE, Mendez Romero A, et al. Identifying at diagnosis the vestibular schwannomas at low risk of growth in a long-term retrospective cohort. Clinical otolaryngol. 2016;41(6):788-92. doi: 10.1111/coa.12661.
30. Coelho DH, Tang Y, Suddarth B, Mamdani M. MRI surveillance of vestibular schwannomas without contrast enhancement: Clinical and economic evaluation. Laryngoscope. 2018;128(1):202-209. doi: 10.1002/lary.26589.

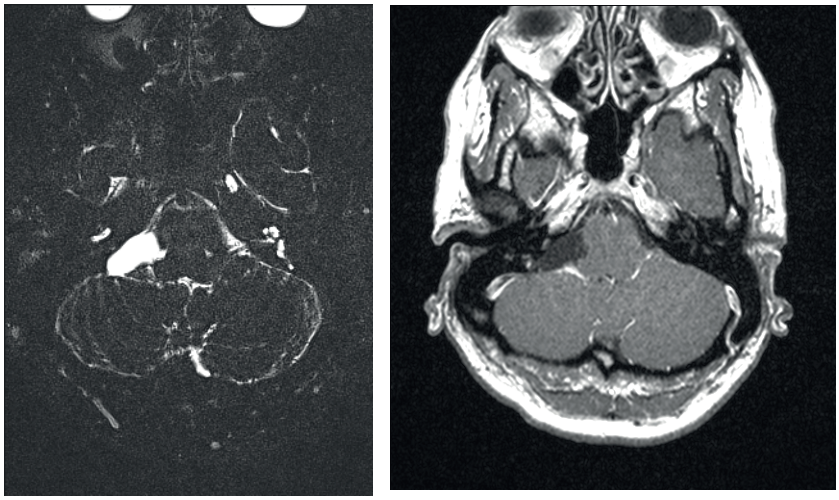
31. Verbist BM. Imaging of sensorineural hearing loss: a pattern-based approach to diseases of the inner ear and cerebellopontine angle. Insights Imaging. 2012;3(2):139-53. doi: 10.1007/s13244-011-0134-z.

32. Espiney Amaro C, Montalvao P, Huins C, Saraiva J. Lyme disease: sudden hearing loss as the sole presentation. J Laryngol Otol. 2015;129(2):183-6.

SUPPLEMENTAL CONTENT



Supplemental figure 5.1 Example of a lesion in the left CPA diagnosed as VS based on T2w (left), while GdT1w revealed features of a calcified meningioma (right).



Supplemental figure 5.2 Arachnoid cyst that was detected on T2w (left) but missed on GdT1w alone (right) by both observers.

CHAPTER 6

Development of a diagnostic model to identify patients at high risk for cerebellopontine angle lesions

Hentschel M.
Hannink G.
Steens S.
Rovers M.
Kunst H.

Submitted.

ABSTRACT

Objectives

We aimed to develop a diagnostic model to identify patients at high risk of a cerebellopontine angle (CPA) lesion.

Design

Cohort study.

Setting

Secondary care. University hospital in the Netherlands.

Participants

A consecutive cohort of patients with asymmetrical audiovestibular dysfunction (AAD) referred by a general practitioner, that underwent their first MRI examination of the CPA between 2005 and 2015.

Primary outcome measure

The presence/absence of a CPA lesion within a patient was used as outcome.

Results

The final model contained eleven variables, namely gender (OR 1.055 (95% CI 0.885-1.905)), sudden onset of hearing loss (OR 0.768 (95% CI 0.318-0.992)), gradual onset of hearing loss (OR 1.069 (95% CI 0.500-1.450)), unilateral tinnitus (OR 0.682 (95% CI 0.374-0.999)), complaints of unilateral aural fullness (OR 1.006 (95% CI 0.783-2.155)), instability (OR 1.006 (95% CI 0.580-2.121)), headache (OR 0.959 (95% CI 0.059-1.090)), facial numbness (OR 2.746 (95% CI 0.548-11.085)), facial nerve dysfunction during physical examination (OR 1.024 (95% CI 0.280-3.702)), and asymmetry in BC at 1 kHz (OR 1.013 (95% CI 1.000-1.027)) and 4 kHz (OR 1.008 (95% CI 1.000-1.026)).

Conclusions

The proposed diagnostic model is a first step in selecting patients with a high risk of a CPA lesion among those with AAD. It needs to be externally validated prior to its implementation in clinical practice. Clinicians may, however, use the model in the future to differentiate between subjects with a high and low risk of CPA lesions, and use it in shared decision making and to explain future strategies.

BACKGROUND

Currently, patients with asymmetrical audiovestibular dysfunction (AAD, asymmetrical hearing loss, asymmetrical tinnitus, dizziness) undergo a magnetic resonance imaging (MRI) examination to screen for lesions in the cerebellopontine angle (CPA), the space located between inner ear canal and brain. These are detected in approximately 3% of the screening population.[1, 2] Vestibular schwannoma (VS), a benign tumour originating from the myelin sheath of the vestibulocochlear nerve, is most common, whereas other types of lesions, such as meningiomas or arachnoid cysts, occur less often.[2-4]

Because the incidence of CPA lesions in the screening population is low, the majority of MRIs is negative. Optimization of the diagnostic process prior to imaging would ideally reduce the number of MRIs and its costs without missing CPA lesions. The magnitude of these savings can be illustrated based on cost prices spent on CPA screening. An MRI of the CPA costs about €206 (approximately £191). If all MRIs without VS could be avoided, approximately €293 (or £271, including price of consultation following MRI) could be saved per patient with AAD.[5, 6] In the Netherlands (17.3 million inhabitants)[7], this could result in potential savings of up to €3.2 million per year.[8] It therefore seems worthwhile to investigate new diagnostic strategies that can be used to preselect patients with a high risk of a CPA lesion for MRI.

A recent diagnostic meta-analysis did not reveal accurate existing non-imaging screening methods for detection of VS and CPA lesions.[9] In practice, clinicians use information on history, physical examination, and additional tests to generate a differential diagnosis. So far, only one study combined demographics, symptoms and audiometry to create a diagnostic model to select patients with AAD for MRI.[10] However, this study used a case-control design, did not perform an MRI in all controls and it was unclear how cases were diagnosed.

We therefore aimed to develop a multivariable diagnostic model for patients with AAD that can be used to identify patients at high risk of a CPA lesion for MRI. Such a diagnostic model would be a great asset in identifying patients at high risk of a CPA lesion, and could be used to guide doctors and patients in shared decision making regarding diagnostics and expectations.

METHODS

Population

The model was developed using a cohort of patients aged 16 years and older that visited the otolaryngology department of a university hospital (Radboud University Medical Center, Nijmegen, the Netherlands) with AAD complaints between January 2005 and February 2015. All consecutive patients referred by a general practitioner and undergoing their first MRI examination of the CPA in our hospital

were retrieved from the Radiology database (IMPAX 6.6.1.1527, version 6.6.1.0 2015, AGFA Healthcare N.V., Mortsel, Belgium). Patients referred by an otolaryngologist from another hospital were excluded, because they are usually diagnosed with a CPA lesion prior to referral. Inclusion would result in a higher incidence of CPA lesions compared to the regular screening population. Patients whose complete medical records were missing and patients with incomplete MRI images (e.g. due to claustrophobia) were also excluded.

The study was performed and reported according to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement.[11] The need for informed consent was waived by the local Medical Ethics Committee (CMO Radboudumc), because of the study's retrospective nature and size.

Outcome

The presence/absence of a CPA lesion within a patient was used as outcome, irrespective of its side and uni- or bilateral presence. CPA lesion diagnosis was based on the original neuroradiologist's report. The outcome was considered present whenever any type of CPA lesion was suspected based on MRI images as assessed by a neuroradiologist from our institution. Lesions of unknown origin were also considered CPA lesions, because these are usually considered abnormal and included in a (temporary) follow-up policy.

Size of CPA lesions

Information was gathered about the size of all VSs and meningiomas at time of diagnosis. In case MRI images were available we determined largest diameter in two directions on axial images: parallel to the internal auditory canal (split in an intra- and extrameatal portion delineated by the petrous bone) and largest diameter parallel to the petrous bone.

Potential predictors

Information on demographics, symptoms, physical examination, and pure-tone audiometry (PTA) results were collected from the patients' otolaryngology records using a pre-specified case report form. Potential predictors were selected based on previous studies and expert opinion.[9]

Demographics

Gender and age were included as potential predictors.

Symptoms

Hearing loss was scored asymmetrical whenever a subjective difference was reported between ears (including unilateral complaints). Moreover, we collected data about the onset of hearing loss. Patients were scored to have a sudden and/or gradual onset of hearing loss, when it was described as such in at least one ear. Duration of tinnitus was scored as either more or less than two months, the latter

also being applied for patients without tinnitus. Complaints of facial numbness and weakness, vertigo, instability and headache were scored as either absent or present. Symptoms were considered absent in case they were not mentioned in the patient record.

Physical examination

Facial nerve dysfunction was positive in case any abnormality was described on at least one side.

Pure-tone audiometry

PTA examinations performed within six months prior to MRI were included. We ensured blinding (of patients and examiners) by solely including PTAs performed prior to MRI. Data were retrieved from the clinical audiology database system.

We collected hearing thresholds in decibels hearing level (dBHL) of octave frequencies 0.5, 1, 2, 4, and 8 kHz for bone conduction (BC) and air conduction (AC). Absolute asymmetry of BC hearing thresholds between ears was calculated for each frequency. We calculated absolute asymmetry of the low and high Fletcher Index between ears using hearing thresholds in dBHL of octave frequencies 0.5, 1 and 2, and 1, 2 and 4 kHz AC, respectively. All were modeled as continuous variables.

Data analysis

For 13 of the 20 variables, data were missing ranging from 1.7% to 50.9% (Table 6.1). These missing data were imputed using multiple imputation by chained equations procedure using predictive mean matching. The R package mice was used to perform multiple imputation.[12] Missing data were assumed to be missing at random (MAR).

The MAR assumption appeared to be valid by visual exploration of missingness.[13, 14]

In the imputation model, we included all variables that appear in the complete data model, i.e. the model that was applied to the data after multiple imputation, including the outcome variable. In addition, we included baseline variables that are known to have influenced the occurrence of missing data.[13]

To determine the number of imputed datasets, an iterative multiple imputation approach, implemented in the R package 'imi', was used.[15] In short, the approach involves to keep adding imputed datasets until the estimates converge, or change very little as new imputed datasets are added. The two main parameters that are needed to be specified before applying our procedure on an incomplete dataset are 1) the stopping rule for the distance between two steps (ϵ), and 2) the number of steps (k_0) this criterion should be successively validated. As suggested by Nassiri et al. we used $\epsilon = 0.01$ as a conservative choice for the stopping rule and $k_0 = 3$ successive validation steps.[15] Based on this approach 90 imputed datasets were needed.

Model selection after multiple imputation was performed using a penalized logistic regression using least absolute shrinkage and selection operator (LASSO) taking into account the 90 multiple imputed datasets as implemented in the R package 'MAMI'. The LASSO implementation used in 'MAMI' is based on 'cv.glmnet' from the R package 'glmnet'. 'glmnet' is a package that fits a generalized linear model via penalized maximum likelihood.[16] The regularization path was computed for the LASSO penalty at a grid of values (n = 100) for the regularization parameter lambda (λ) using the package's default settings. 10-fold cross-validation was used to look for the optimal tuning parameter λ for the LASSO (the λ value that leads to the smallest mean squared error).

A variable will be formally 'selected' if it is selected in at least one imputed set of data, but its overall impact will depend on how often it is chosen.[17] This may lead to different results than pragmatic approaches which may select predictors only if they are contained in most imputed sets of data; or select variables based on a stacked dataset of all imputations and apply weights to this dataset[18]; or select variables based on averaged model selection criteria (e.g. AIC)[19]. Post model selection estimates obtained from the imputed datasets were pooled by means of Rubin's rules.

It is known from the literature that estimators post model selection not necessarily have a normal, or even symmetric distribution.[17, 20, 21] Moreover, bootstrapped confidence intervals can help to improve the coverage of model selection estimators.

Bootstrap model selection confidence intervals based on 200 bootstrap replicates, after imputation, were generated as described in detail elsewhere.[17, 22] In short, the following steps were taken; 1) 200 bootstrap samples of the original data (including missing observations) were created, 2) 90 imputed sets of data were generated for each bootstrap sample, 3) a model averaging (or selection) estimator was calculated for each imputed set of data in each bootstrap sample, 4) a model averaging (or selection) estimator was created after imputation for each bootstrap sample, 5) the average of the 200 estimates calculated in step 4 as the final point estimate, and 6) 95% CIs were constructed based on the percentiles of the empirical distribution produced by the 200 estimates of step 4.

Model performance measures, i.e. calibration intercept, calibration slope, and c-index, were estimated in each imputed dataset, and subsequently pooled using multiple imputation rules (so-called pooled performance strategy).[23]

The model was internally validated using bootstrap resampling for internal validation and estimation of the expected optimism was performed based on Musoro et al.[24] In short, the following steps were taken; 1) a bootstrap sample was taken from the incomplete data set and then multiple imputation was performed 90 times, 2) every model building step in the original model construction (see above) was repeated. The performance of the selected model on each imputed dataset was evaluated and averaged to obtain the apparent performance (pooled performance strategy), 3) the selected model was applied

to the original samples to determine the averaged test performance, 4) optimism, defined as optimism = apparent performance – test performance, was calculated, 5) steps 1 to 4 were repeated 100 times to obtain a stable estimate of the optimism, and finally 6) the optimism corrected performance was calculated.

The amount of miscalibration was quantified via the calibration slope β_{LP}. Correction was achieved by re-estimating the intercept and multiplying each estimated effect with a shrinkage factor *s* that was determined as follows. In every bootstrap run, model construction per imputed dataset was carried out as with the original sample and values of the linear predictors (LP) were calculated on the original samples. The intercept (α_{LP}) and slope (β_{LP}) of LP were estimated by regressing the outcome in the original sample on the LP. This process was repeated 100 times and *s* was calculated as the mean of the 100 estimates of β_{LP}. The re-calibrated model was $\bar{\alpha}_{LP} + \bar{\beta}_{LP}*(\alpha+\beta X)$.

To easily calculate an individual's risk of having a VS using the model, a dynamic nomogram was created. The nomogram is available via https://vs-model.shinyapps.io/predictCPA, where more data can be entered and corresponding predictions can be calculated.

TRIPOD recommends to evaluate a prediction model's net benefit.[11] Decision curve analysis (DCA) may help to summarize clinical usefulness of prediction models and support clinicians in decision making.[25, 26] DCA was performed after pooling the imputation specific predictions using Rubin's rules (so-called pooled prediction strategy).[23]
DCA is a plot of net benefit (NB) against threshold probability.
NB gives the proportion of "net" true positives in the dataset: the observed number of true positives is corrected for the observed proportion of false positives weighted by the odds of the risk threshold, and the result is divided by the sample size. This "net" proportion is equivalent to the proportion of true positives in the absence of false positives (i.e. perfect specificity).[25]

NB is calculated as follows:[27]
$net\ benefit = true\ positives/N - false\ positives/N * p_t/(1-p_t)$

Threshold probability (p_t) is defined as the minimum predicted risk of having a CPA lesion at which an otolaryngologist or patient would want an MRI. To represent a variety of preferences, a range of values is displayed.[27, 28] This range of values should be established prior to model reporting. Interviews with otolaryngologists from our centre working in the field of CPA lesions revealed that threshold values from 0% (MRI for all patients to find all CPA lesions, regardless of unnecessary MRIs) to 5% (indicating that one accepts 19 unnecessary MRIs to find 1 CPA lesion in a group of 20 patients) were regarded relevant by them. Net benefit represents the proportion of true positives (diagnosed CPA lesions) in absence of any false positives (specificity of 100%). To obtain standardized net benefit, the incidence of disease (intercept with y-axis) is set to 100%. Using DCA, one can compare the model to a 'scan all' (i.e. the current) or 'scan

none' strategy. Furthermore, we plotted the number of MRIs avoided per 1000 patients at risk against the threshold probability, which can be used to assess potential savings (in terms of MRIs) of the model. Additionally, the true and false positive rates were plotted against the threshold probability.

R statistical software version 3.6.0 (The R Foundation for Statistical Computing, Vienna, Austria) with packages 'imi', 'mice' and 'MAMI' were used for data analysis.[12, 15, 22]

Patient and public involvement statement

Representatives of the patient society for CPA lesions supported the study protocol. Meetings were organized with representatives to update them on study findings and exchange ideas and comments.

RESULTS

Population

We retrieved data of 2,725 adult patients with AAD who had visited our department and had undergone an MRI to screen for CPA lesions. We excluded 511 patients: 103 children, 6 missing patient records, 5 incomplete MRI examinations in which the CPA could not be properly assessed (3 claustrophobia, 2 movement/metal artefacts), and 397 patients were referred by an otolaryngologist from another hospital, resulting in 2,214 inclusions.

Diagnostic variables

Mean age at time of consultation was 56 (range: 16-93) years and 1149 (51.9%) patients were men. Table 6.1 displays an overview of patient characteristics and missings per potential diagnostic variable.

Outcome

Outcome data were available for all patients. Seventy-three CPA lesions were present in 69 (3.1%) subjects (i.e. 4 bilateral lesions), 42 (57.5%) were located on the right and 31 (42.5%) on the left side. Unilateral VS was found in 46 (2.1%) and bilateral VS in 2 patients (0.1%), 28 and 22 of VSs were found on the right and left side (56% and 44%), respectively. Another 19 (0.9%) patients had a unilateral meningioma (n=7; 0.3%), arachnoid cyst (n=5; 0.2%), lipoma (n=1; 0.05%), and lesion of unknown origin (n=6; 0.3%) in the CPA. One patient had bilateral metastases and one bilateral lesions of unknown origin.

Final diagnostic model

The final model consisted of eleven variables, namely gender (OR 1.055 (95% CI 0.885-1.905)), sudden onset of hearing loss (OR 0.768 (95% CI 0.318-0.992)), gradual onset of hearing loss (OR 1.069 (95% CI 0.500-1.450)), unilateral tinnitus (OR 0.682 (95% CI 0.374-0.999)), complaints of unilateral aural fullness (OR 1.006 (95% CI 0.783-2.155)), instability (OR 1.006 (95% CI 0.580-2.121)), headache (OR 0.959 (95% CI 0.059-1.090)), facial numbness (OR 2.746 (95% CI 0.548-11.085)), facial nerve dysfunction during physical

examination (OR 1.024 (95% CI 0.280-3.702)), and asymmetry in BC at 1 kHz (OR 1.013 (95% CI 1.000-1.027)) and 4 kHz (OR 1.008 (95% CI 1.000-1.026)). Table 6.2 displays coefficients and 95% CIs of variables obtained in the final model after internal validation.

Table 6.1 Overview of patient characteristics and missing data in 2,214 included patients.

			Descriptive n (%)*			
Variable	Total N=2214	Missing	CPA lesion n=69	Missing	No CPA lesion n=2145	Missing
Demographics						
- **Gender (male)**	1149 (51.9)	0	41 (59.4)	0	1108 (51.7)	0
- **Age†**	58 (16-93)	0	58 (16-86)	0	58 (16-93)	0
Hearing loss						
- **Asymmetrical**	1217 (55.0)	538 (24.3)	46 (66.7)	11 (15.9)	1171 (54.6)	527 (24.6)
- **Sudden onset ‡**	317 (14.3)	1108 (50.0)	8 (11.6)	31 (44.9)	309 (14.4)	1077 (50.2)
- **Gradual onset ‡**	397 (17.9)	1126 (50.9)	20 (29)	31 (44.9)	377 (17.6)	1095 (51)
Unilateral tinnitus	997 (45.0)	0	21 (30.4)	0	976 (45.5)	0
Unilateral aural fullness	277 (12.5)	0	9 (13)	0	268 (12.5)	0
Dizziness						
- **Vertigo**	347 (15.7)	38 (1.7)	9 (13)	0	338 (15.8)	38 (1.8)
- **Instability**	179 (8.1)	38 (1.7)	7 (10.1)	0	172 (8)	38 (1.8)
Headache	77 (3.5)	0	1 (1.4)	0	76 (3.5)	0
Facial complaints						
- **Facial numbness**	29 (1.3)	0	3 (4.3)	0	26 (1.2)	0
- **Facial weakness**	18 (0.8)	0	1 (1.4)	0	17 (0.8)	0
Physical examination						
- **Facial nerve dysfunction (HB≥2)**	20 (0.9)	0	1(1.4)	0	19(0.9)	0
PTA asymmetry§						
- **BC 0.5 kHz**	10 (0-65)	723 (32.7)	15 (0-55)	25 (36.2)	10 (0-65)	698 (32.5)
- **BC 1 kHz**	10 (0-75)	714 (32.2)	20 (0-70)	25 (36.2)	10 (0-75)	689 (32.1)
- **BC 2 kHz**	10 (0-75)	714 (32.2)	10 (0-65)	25 (36.2)	10 (0-75)	689 (32.1)
- **BC 4 kHz**	10 (0-85)	715 (32.3)	20 (0-75)	25 (36.2)	10 (0-85)	690 (32.2)
- **BC 8 kHz**	0 (0-55)	749 (33.8)	5 (0-40)	28 (40.6)	0 (0-55)	721 (33.6)
- **High FI**	13 (0-117)	521 (23.5)	25 (0-103)	18 (26.1)	13 (0-117)	503 (23.4)
- **Low FI**	13 (0-117)	521 (23.5)	17 (0-103)	18 (26.1)	13 (0-117)	503 (23.4)

PTA: pure-tone audiometry, BC: bone conduction, FI: Fletcher-index
* The number of patients and corresponding percentage is reported, unless stated otherwise in the first column
† Median years (range)
‡ In at least one ear
§ Median dB (range)

Table 6.2 Estimates of the final diagnostic model and 95% confidence intervals.

	Coefficient	OR	Lower 95% CI	Upper 95% CI
Intercept	-3.731			
Gender*	0.065	1.055	0.885	1.905
Sudden onset of hearing loss*	-0.325	0.768	0.318	0.992
Gradual onset of hearing loss*	0.082	1.069	0.500	1.450
Unilateral tinnitus*	-0.471	0.682	0.374	0.999
Unilateral aural fullness*	0.007	1.006	0.783	2.155
Instability*	0.007	1.006	0.580	2.121
Headache*	-0.052	0.959	0.059	1.090
Facial numbness*	1.242	2.746	0.548	11.085
Facial nerve dysfunction*	0.030	1.024	0.280	3.702
Asymmetry in BC at 1 kHz (dB)**	0.016	1.013	1.000	1.027
Asymmetry in BC at 4 kHz (dB)**	0.010	1.008	1.000	1.026

OR: Odds Ratio, CI: confidence interval, BC: bone conduction, kHz: kilohertz, dB: decibel.
The probability of having a CPA lesions can be calculated as follows using the regression coefficients presented above:
* Binary variables: 0 = absent, 1 = present
** Continuous variable
Probability (P) of having a CPA lesion = 1/ (1 + exp(-lp)), where lp = -3.73121 + (0.06519 x gender) + (-0.32472 x sudden onset of hearing loss) + (0.08241 × gradual onset of hearing loss) + (-0.47109 × unilateral tinnitus) + (0.00738 x unilateral aural fullness) + (0.00738 x instability) + (-0.05166 x headache) + (1.24230 × facial numbness) + (0.02952 x facial nerve dysfunction) + (0.01599 × asymmetry in BC at 1 kHz) + (0.00984 × asymmetry in BC at 4 kHz)
In the online dynamic nomogram data can easily be entered in the model. It can be found via https://vs-model.shinyapps.io/predictCPA.

The AUC of the model's ROC curve based on pooled predictions was 0.67 (95% CI 0.59-0.75), indicating acceptable discrimination. The calibration intercept was 0.00 (95% CI -0.24-0.24) and the calibration slope 1.15 (95% CI 0.64-1.66).

The net benefit curve in Figure 6.1A can be used to assess the standardized net benefit for different threshold probabilities at which one can use the model. The model's net benefit is higher than the current 'scan all' strategy for risk thresholds >1.8%. Figure 6.1B displays the number of MRIs avoided for different threshold probabilities. At a risk threshold of 1.8%, 2.5% and 5%, using the prediction model compared to the current strategy could avoid 1.1%, 22.4% and 44.5% of MRIs, respectively. Figure 6.1C shows the true and false positive rate for different threshold probabilities. As the risk threshold increases, the number of patients diagnosed with a CPA lesion decreases.

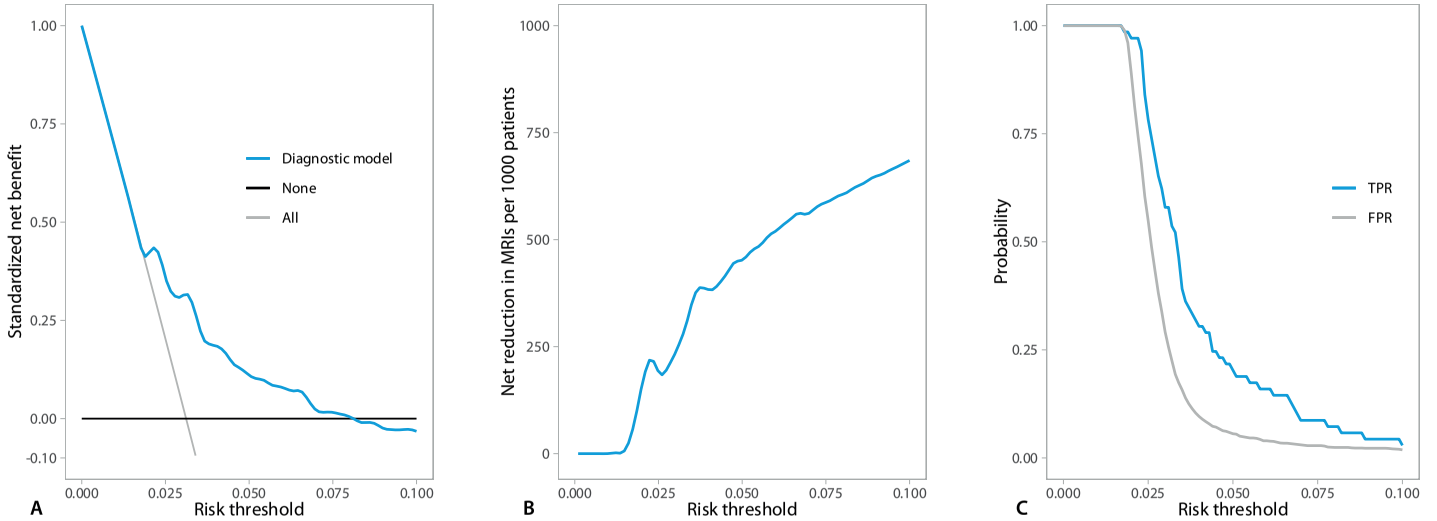


Figure 6.1 Decision curve analysis.

Figure 6.1A Standardized net benefit curve.

The x-axis represents the risk threshold, the y-axis represents the standardized net benefit. Net-benefit represents the proportion of true positives (detected CPA lesions) in absence of any false positives (i.e. specificity of 100%). The black line represents a strategy in which no MRIs are acquired, the net-benefit is 0. The grey line represents the current strategy, in which all patients have undergone an MRI. The blue line represents the prediction model.

Figure 6.1B MRIs avoided.

The x-axis represents the risk threshold, the y-axis the number of MRIs avoided per 1000 patients with AAD.

Figure 6.1C True and false positive rate.

The x-axis represents the risk threshold, the y-axis the probability of patients with true positive diagnoses (blue line) and false positive diagnoses (grey line). Decline of the blue line indicates that CPA lesions will be missed.

Size of CPA lesions

Of all patients with a VS, MRI images were not available for two patients. Of the patients with a unilateral VS, three had an intracochlear VS of which measurements were not included in our analyses. Of 45 remaining VSs (including two patients with bilateral VSs), 27 patients had a purely intrameatal localization of their VS. Mean size of the intrameatal portion parallel to the internal auditory canal was 6.1 mm (median 5 mm, range 0-14 mm). The mean extrameatal portion parallel to the internal auditory canal was 4.3 mm (median 0 mm, range 0-25 mm) and the mean size parallel to the petrous bone was 7.9 mm (median 5 mm, range 1-30 mm). Among the seven patients with a meningioma, mean intrameatal size was 2.1 mm (median 0, range 0-11 mm). The mean extrameatal portion parallel to the internal auditory canal was 12.3 mm (median 14 mm, range 4-21 mm) and parallel to the petrous bone mean size was 15.9 mm (median 13, range 5-37 mm).

DISCUSSION

We developed a clinical prediction model to identify those with high risk of a CPA lesion in AAD patients. This rule contains eleven variables, i.e. gender, sudden onset of hearing loss, gradual onset of hearing loss, unilateral tinnitus, unilateral aural fullness, instability, headache, facial numbness, facial nerve dysfunction and asymmetry in BC at 1 and 4 kHz. The presented decision curves can be used to compare the clinical value of the prediction model with the current 'scan all', and a 'scan none' strategy. The model can be used for the diagnostic workup of patients with AAD.

Comparison with other studies

The incidence of CPA lesions in our screening population was 3.1%, which is comparable with literature.[1, 2] The previously reported diagnostic model containing different PTA thresholds and presence of vertigo as diagnostic parameters achieved a sensitivity and specificity of 80% and 74%, respectively, for the threshold with the highest combination of sensitivity and specificity.[10] It was, however, developed using a case-control and split-sample design, which is considered less suitable for diagnostic modelling.[29] Moreover, only 55 of the 288 controls underwent MRI so that verification bias cannot be excluded. It is also unclear how CPA lesions were diagnosed in the 'cases'.[10] This hampers comparison with our study.

Strengths and limitations

To our knowledge, this is the first cohort study creating a prediction model for patients with AAD. We included patients who had not visited an otolaryngologist for their complaints prior to consultation in our hospital, to ensure our study population represented a standard screening population instead of that of a tertiary referral centre (where diagnoses are usually established beforehand). All patients underwent MRI, so outcome data were available for all patients. Moreover, our study comprised multiple types of variables (demographics, symptoms, physical examination and PTA) that can be used for screening, instead of studying only one type of screening test. Variables that serve as input

for the diagnostic rule can be acquired easily through history taking, physical examination and PTA. Both patients and examiners are ideally blinded for the outcome when determining values of potential diagnostic variables. We ensured blinding by solely including PTAs performed prior to MRI. Moreover, we used a timeframe of 6 months to ensure that PTAs were representative for hearing levels at the time of MRI.

Some potential limitations should also be discussed. First, our data were collected from patients' records, i.e. not all required information could be derived, resulting in missing values. We used multiple imputation, which is the recommended method compared to complete case analysis.[13, 30] Whenever complaints were not reported in the patient record, they were assumed to be absent. Possibly, some complaints were not reported by the otolaryngologist, or not mentioned by a patient. We assume however, that most complaints included in our analyses (e.g. hearing loss, tinnitus and dizziness) are usually registered, because they are common in patients with a suspected CPA lesion. Second, we excluded 11 patients, because their MRI examination was incomplete (e.g. claustrophobia) or their patient record was missing. We believe claustrophobia and missing records are not related to the outcome or other variables, therefore exclusion of these patients will not have influenced results. Third, criteria for MRI referral are known to vary. [31] Our local protocol prescribes an MRI for asymmetries of ≥10 dB on three consecutive frequencies, unilateral constant tinnitus ≥ 3 months, unilateral decreased or absent vestibular function. However, our national protocol on sensorineural hearing loss does not specify the amount or frequencies of asymmetry that are required for MRI referral. The group of included patients might therefore be heterogenous considering their hearing loss, but comprise patients that currently would undergo an MRI and are therefore representative for current practice.

Clinical implications

New screening methods to select patients with AAD for MRI will undoubtedly result in false negative results. Currently, it is a challenge to properly assess consequences of missed CPA lesions resulting from this diagnostic rule, for both patients (e.g. quality of life and functional outcomes) and society (i.e. costs). The majority of patients with a VS are obtained in a 'wait and scan' policy. A large proportion of VSs does not grow and thus remains untreated for years.[32-34] One could question the need of diagnosing CPA lesions for which a therapeutic consequence is lacking.

We believe that a next step in optimizing the diagnostic process of CPA lesions would be to focus on diagnosing those lesions requiring treatment (i.e. larger and/or growing lesions).

Tinnitus is often considered an indication for an MRI of the CPA.[35] The coefficient of 'unilateral tinnitus', however, turned out to be negative in our prediction model, which seems contradictory with clinical practice. Positive predictive value of unilateral tinnitus was previously shown to be low,[36, 37] which also results from our data: 17 and 21 out of 997 patients (1.7% and 2.1%) with unilateral tinnitus are diagnosed with a VS and CPA lesion, respectively. These numbers are lower than the incidence of CPA lesions in our

data, which explains why they are negative predictors. Moreover, prior to MRI, complaints of uniltaral tinnitus cannot yet be linked to the ear affected by a lesion.
Furthermore, asymmetry in hearinglevels at 1 and 4 kHz showed to be significant predictors for the presence of a CPA lesion, while the other tested frequencies did not. The presented combination of variables turned out to have best predictive abilities for CPA lesions. The contribution of other potential variables, including asymmetry at other frequencies, was too low to be included in the model. This does not mean, however, that they have absolutely no relation with the presence of CPA lesions, but asymmetry at 1 and 4 kHz had a higher contribution when combined with the other variables.

Although this diagnostic rule is a first step in selecting patients at high risk of a CPA lesion for MRI, its diagnostic accuracy would preferably be improved prior to its clinical use. Ideally, it would be possible to eliminate misdiagnoses (particularly false negatives), provide a safety net for false negatively diagnosed patients, or safely state that consequences of false negative diagnoses are negligible, before this diagnostic rule can be safely used. Moreover, the TRIPOD statement highly recommends external validation.[11] Preferably, this is done using prospectively collected data. DCA was used to compare the model's net benefit to a 'scan all' (i.e. the current) and 'scan none' strategy. DCA showed that the model has a higher net benefit than the current strategy when applying a threshold of >1.8%. Eventually, an impact study is needed to evaluate cost-effectiveness; costs and effects of applying the diagnostic rule in clinical practice need to be compared to the current diagnostic strategy in which an MRI is acquired in all patients with AAD.[38] This diagnostic rule can potentially help to focus the otolaryngologist's attention in history taking and requesting additional tests, but also may help a patient decide whether he/she wants to undergo an MRI.

CONCLUSION

The proposed diagnostic rule, containing eleven variables that can easily be assessed in every otolaryngology practice using history taking, physical examination and PTA, is a first step in identifying patients with a high risk of a CPA lesion among those with AAD. Following external validation, clinicians may use the model to differentiate between subjects with a low and high risk of a CPA lesion, and use it in shared decision making and to explain future strategies.

DECLARATIONS

Ethical approval and consent to participate

The need for informed consent was waived by the local Medical Ethics Committee (CMO Radboudumc), because of the study's retrospective nature and size.

Acknowledgements

Not applicable.

REFERENCES

1. Zealley I.A., Cooper R.C., Clifford K.M. et al. Mri Screening for Acoustic Neuroma: A Comparison of Fast Spin Echo and Contrast Enhanced Imaging in 1233 Patients. *Br J Radiol* 2000; 73(867): 242-247.
2. Kleijwegt M., Ho V., Visser O., Godefroy W., van der Mey A. Real Incidence of Vestibular Schwannoma? Estimations from a National Registry. *Otol Neurotol* 2016; 37(9): 1411-1417.
3. Kshettry V.R., Hsieh J.K., Ostrom Q.T., Kruchko C., Barnholtz-Sloan J.S. Incidence of Vestibular Schwannomas in the United States. *J Neurooncol* 2015; 124(2): 223-228.
4. Moffat D.A., Ballagh R.H. Rare Tumours of the Cerebellopontine Angle. *Clin Oncol (R Coll Radiol)* 1995; 7(1): 28-41.
5. Stangerup S.E., Tos M., Thomsen J., Caye-Thomasen P. True Incidence of Vestibular Schwannoma? *Neurosurgery* 2010; 67(5): 1335-1340.
6. Dutch National Health Care Institute. Guideline for Cost Research in Health Care. 2015 December 10. Available from: https://www.zorginstituutnederland.nl/binaries/content/documents/zinl-www/documenten/publicaties/overige-publicaties/1602-richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg/1602-richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg/Richtlijn+voor+het+uitvoeren+van+economische+evaluaties+in+de+gezondheidszorg.pdf.
7. Cbs Statline. Bevolking; Kerncijfers. 2019 October 21. Available from: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/37296NED/table?fromstatweb.
8. Scholte M., Hentschel M.A., Kunst H.P., Steens S.C.A., Rovers M.M., Grutters J.P.C. Potential Savings in the Diagnosis of Vestibular Schwannoma. *Clin Otolaryngol* 2018; 43(1): 285-290.
9. Hentschel M., Scholte M., Steens S., Kunst H., Rovers M. The Diagnostic Accuracy of Non-Imaging Screening Protocols for Vestibular Schwannoma in Patients with Asymmetrical Hearing Loss and/or Unilateral Audiovestibular Dysfunction: A Diagnostic Review and Meta-Analysis. *Clin Otolaryngol* 2016; 24(4): 815-823.
10. Chatrath P., Frosh A., Gore A., Nouraei R., Harcourt J. Identification of Predictors and Development of a Screening Protocol for Cerebello-Pontine Lesions in Patients Presenting with Audio-Vestibular Dysfunction. *Clin Otolaryngol* 2008; 33(2): 102-107.
11. Collins G.S., Reitsma J.B., Altman D.G., Moons K.G. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (Tripod): The Tripod Statement. *BMJ* 2015; 350: g7594.
12. van Buuren S., Groothuis-Oudshoorn, K. Mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw* 2011; 45(3): 1-67.
13. van Buuren S. Flexible Imputation of Missing Data. Chapman and Hall/CRC; 2012.
14. Cheng X.Y., Cook D., Hofmann H. Visually Exploring Missing Values in Multivariable Data Using a Graphical User Interface. *Journal of Statistical Software* 2015; 68(6): 1-23.
15. Nassiri V M.G., Verbeke G, Barbosa-Breda J. Iterative Multiple Imputation: A Framework to Determine the Number of Imputed Datasets. *The American Statistician* 2019: 1-17.
16. Friedman J., Hastie T., Tibshirani R. Regularization Paths for Generalized Linear Models Via Coordinate Descent. *J Stat Softw* 2010; 33(1): 1-22.
17. Schomaker M., Heumann C. Model Selection and Model Averaging after Multiple Imputation.

Computational Statistics & Data Analysis 2014; 71: 758-770.

18. Wood A.M., White I.R., Royston P. How Should Variable Selection Be Performed with Multiply Imputed Data? *Statistics in medicine* 2008; 27(17): 3227-3246.
19. White I.R., Royston P., Wood A.M. Multiple Imputation Using Chained Equations: Issues and Guidance for Practice. *Statistics in medicine* 2011; 30(4): 377-399.
20. Leeb H., Potscher B.M. Model Selection and Inference: Facts and Fiction. *Econometric Theory* 2005; 21(1): 21-59.
21. Wang H.Y., Zou G.H., Wan A.T.K. Model Averaging for Varying-Coefficient Partially Linear Measurement Error Models. *Electronic Journal of Statistics* 2012; 6: 1017-1039.
22. Schomaker M., Heumann C. Model Averaging and Model Selection after Multiple Imputation Using the R-Package Mami. 2017; http://mami.r-forge.r-project.org/
23. Wood A.M., Royston P., White I.R. The Estimation and Use of Predictions for the Assessment of Model Performance Using Large Samples with Multiply Imputed Data. *Biometrical journal Biometrische Zeitschrift* 2015; 57(4): 614-632.
24. Musoro J.Z., Zwinderman A.H., Puhan M.A., ter Riet G., Geskus R.B. Validation of Prediction Models Based on Lasso Regression with Multiply Imputed Data. *BMC Med Res Methodol* 2014; 14: 116.
25. Van Calster B., Wynants L., Verbeek J.F.M. et al. Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators. *Eur Urol* 2018; 74(6): 796–804.
26. Balachandran V.P., Gonen M., Smith J.J., DeMatteo R.P. Nomograms in Oncology: More Than Meets the Eye. *Lancet Oncol* 2015; 16(4): e173-180.
27. Vickers A.J., Van Calster B., Steyerberg E.W. Net Benefit Approaches to the Evaluation of Prediction Models, Molecular Markers, and Diagnostic Tests. *BMJ* 2016; 352: i6.
28. Capogrosso P., Vickers A.J. A Systematic Review of the Literature Demonstrates Some Errors in the Use of Decision Curve Analysis but Generally Correct Interpretation of Findings. *Med Decis Making* 2019: 272989X19832881.
29. Hendriksen J.M., Geersing G.J., Moons K.G., de Groot J.A. Diagnostic and Prognostic Prediction Models. *J Thromb Haemost* 2013; 11 Suppl 1: 129-141.
30. Altman D.G., Bland J.M. Missing Data. *BMJ* 2007; 334(7590): 424.
31. Hentschel M., Rovers M., Markodimitraki L., Steens S., Kunst H. An International Comparison of Diagnostic and Management Strategies for Vestibular Schwannoma. *Eur Arch Otorhinolaryngol* 2019; 276(1): 71-78.
32. Ferri G.G., Modugno G.C., Pirodda A., Fioravanti A., Calbucci F., Ceroni A.R. Conservative Management of Vestibular Schwannomas: An Effective Strategy. *Laryngoscope* 2008; 118(6): 951-957.
33. Hajioff D., Raut V.V., Walsh R.M. et al. Conservative Management of Vestibular Schwannomas: Third Review of a 10-Year Prospective Study. *Clin Otolaryngol* 2008; 33(3): 255-259.
34. Yoshimoto Y. Systematic Review of the Natural History of Vestibular Schwannoma. *J Neurosurg* 2005; 103(1): 59-63.
35. Nash R., Majithia A., Singh A. The Agreement between Protocols for the Investigation of Asymmetrical Audiovestibular Symptoms. *J Int Adv Otol* 2016; 12(1): 98-100.
36. Choi K.J., Sajisevi M.B., Kahmke R.R., Kaylie D.M. Incidence of Retrocochlear Pathology Found on Mri in Patients with Non-Pulsatile Tinnitus. *Otol Neurotol* 2015; 36(10): 1730-1734.
37. Vandervelde C., Connor S.E. Diagnostic Yield of Mri for Audiovestibular Dysfunction Using Contemporary Referral Criteria: Correlation with

Presenting Symptoms and Impact on Clinical Management. *Clin Radiol* 2009; 64(2): 156-163.

38. Moons K.G., Altman D.G., Vergouwe Y., Royston P. Prognosis and Prognostic Research: Application and Impact of Prognostic Models in Clinical Practice. *BMJ* 2009; 338: b606.

PART III

Management of vestibular schwannoma

CHAPTER 7

Quality of life in patients with a vestibular schwannoma: does size or treatment matter?

Hentschel M.
Wijn S.
Steens S.
Kunst H.
Rovers M.

ABSTRACT

Objective

To study health-related quality of life (HRQoL) of recently diagnosed patients with VS, stratified by size and management strategy.

Design

Cross-sectional survey.

Setting

University hospital.

Participants

A survey was conducted in 252 adult patients with a newly diagnosed unilateral VS between 2014 and 2017.

Main outcome measures

The survey included the disease-specific PANQOL, the SF-36 and EQ-5D-5L. Subgroups were defined according to VS size (Koos stage) and management strategy, i.e. treatment and wait and scan. Mean questionnaire and utility scores were calculated. A utility score is a numeric value ranging from death (0) to perfect health (1).

Results

The survey was completed by 174 patients (69%). Respondents had a mean age of 59 years (SD: 11.7, range: 22-89) and 50% was male. Mean utility scores were lowest for treated Koos 2 (0.673) and highest for wait and scan Koos 1 (0.829) patients. When comparing wait and scan patients, differences between Koos stages were small. Overall, treated patients had lower scores than wait and scan patients with the same Koos stage, particularly in Koos 2 (0.673 versus 0.820).

Conclusions

The lack of difference in HRQoL between Koos stages in patients in the wait and scan strategy suggests that small and medium-sized VSs might grow without having a large impact on the HRQoL. This should be further investigated using longitudinal data.

INTRODUCTION

Patients with a vestibular schwannoma (VS) usually have symptoms of asymmetrical hearing loss, tinnitus, dizziness and/or imbalance, but headaches and facial numbness and/or weakness may also occur.[1] Consequently, a VS can have a large impact on the health-related quality of life (HRQoL) of patients.[2, 3] Management strategies for VS usually comprise a wait and scan (W&S) strategy aiming to detect growth of the lesion by acquiring repeated magnetic resonance imaging (MRI) examinations, or active treatment, i.e. microsurgery or radiotherapy (including stereotactic radiosurgery).[4, 5] There is debate about the most suitable management strategy for VS.[6, 7] Currently, most patients undergo a W&S strategy, i.e. only patients with a large and/or growing VS (observed during W&S) receive active treatment.[8] The challenge is to decide when to switch from W&S to an active treatment strategy.
So far, several studies reported on the HRQoL in patients with VS, but detailed information about differences in HRQoL between patients according to size of their lesion and treatment strategies is lacking.[2, 9-17] Knowledge on this topic could help clinicians in patient counselling and decision making about whether or not to proceed to active treatment. Therefore, we studied the HRQoL of recently diagnosed VS patients, stratified by Koos stage and management strategy.

METHODS

Ethical considerations

We performed a survey among VS patients. The need for informed consent was waived by the Medical Ethics Committee of the Radboud University Medical Centre, Nijmegen, The Netherlands.

Patient enrolment

The survey was sent to adult patients that were newly diagnosed with a unilateral VS between January 1st 2014 and January 1st 2017 at the otolaryngology department of the Radboud University Medical Centre. Patients with severe comorbidity such as malignancies, dementia or genetic conditions (including Neurofibromatosis II) were excluded.

Information was retrieved from electronic patient records including age, gender, VS side and visit date(s) at the otolaryngology department. VS management at time of the survey was registered as W&S or active treatment (i.e. the patient had already received microsurgery or stereotactic radiosurgery). VS size was scored according to the Koos classification system, which defines 4 stages.[18] A Koos stage 1 lesion is limited to the internal auditory canal; stage 2 extends into the cerebellopontine angle without contacting the brainstem; stage 3 fills the entire cerebellopontine angle contacting the brainstem; and stage 4 displaces the brainstem and adjacent cranial nerves.[18] We determined the Koos stage for each patient by assessment of the most recent MRI examination (prior to treatment, if applicable). Patients without available MRI images were excluded since Koos stages could not be determined.

Eligible patients were sent an invitation to participate in the study. In case an e-mail address was available, patients received an electronic invitation. It contained a description of the study and provided the option to either refuse participation or continue to the survey. Reminders were sent after one and three weeks. A paper-based survey was sent to the remaining patients by post including an introductory letter, patient information flyer, informed consent and a postage-paid return envelope. No reminder was sent to the latter patients.

Questionnaires

We used three questionnaires, i.e. Penn Acoustic Neuroma Quality of Life (PANQOL) scale, 36-Item Short Form Survey (SF-36) and EuroQoL-5D five level questionnaire (EQ-5D). The first two measure HRQoL and the latter two measure utility scores, i.e. the SF-36 measuring both. Utility scores can be used to study cost-effectiveness (where they are used to quantify effectiveness), as well as to compare patients with different diseases. A utility score is a numeric value ranging from death (0) to perfect health (1). They can be derived using health states, which exist for every potential combination of answers to these questionnaires. The utility score of each specific health state has previously been determined through valuing of all possible health states by the general population. Utility scores are not available (yet) for the PANQOL.

The PANQOL questionnaire is a disease-specific quality of life instrument developed for patients with VS.[19] The questionnaire was recently translated and validated for the Dutch population.[20] The PANQOL consists of 26 items each having five optional answers on a Likert-scale ranging from 'strongly disagree' to 'strongly agree'. Questions can be grouped in the domains Anxiety, Balance, General, Hearing, Energy, Pain and Face. Scores were calculated on a scale from 0 to 100 (worst to best HRQoL, respectively) for each domain. The total instrument score was determined by calculating the mean of the 7 domain scores, constructing the Total domain score.[19]

The SF-36 is a multipurpose health survey containing 36 items and 8 domains, i.e. Physical Functioning (PF), Social Functioning (SF), Physical Role limitations (PR), Emotional Role limitations (ER), Mental Health (MH), Vitality (VT), Bodily pain (BP), and General Health perceptions (GH). For each domain, a score can be calculated on a scale of 0 to 100 with higher scores indicating better health and wellbeing. Mental and Physical Composite Scores were calculated as previously described.[21] Although it was not developed for this purpose, the SF-36 can also be used to calculate utility scores. Because the SF-36 is multidimensional and domains are not equally important, a transformation is needed to calculate a utility score, for which we used the SF-6D algorithm.[22] After transformation utility scores can range from 0.17 (worst health) to 1.00 (perfect health).[22-24]

The EQ-5D is a multi-attribute health state system developed by the EuroQol group.[25] It is the most frequently used instrument to measure utility scores and is recommended by several health care institutes.[26, 27] The questionnaire consists of five items to measure generic health status (i.e. mobility,

self-care, usual activities, pain/discomfort and anxiety/depression). Every item has five possible answers (ranging from 'no problems' to 'extreme problems (unable)'), providing a health state when combined. Preferences in ranking the importance of the different domains vary per population, so we used a specific rate for the Dutch population to calculate utility values.[28] The EQ-5D utility value can range from 0 (death) to 1 (perfect health). The EQ-5D additionally contains a visual analogue scale (VAS) looking like a thermometer, that patients can use to rate their overall HRQoL ranging from 0 (worst imaginable health) to 100 (best imaginable health). The VAS score can be used to compare utility scores with the self-reported HRQoL.

Analyses

All analyses were performed using RStudio, Version 1.1.383 (2009-2017 RStudio, Inc., Boston). The mean time between 'diagnosis and survey' and 'most recent MRI and survey' was calculated. In case a patient had received treatment, the MRI prior to treatment was used to determine the Koos stage. For the latter patients, mean time between 'diagnosis and treatment' and 'treatment and survey' was additionally calculated.

The 'mice' package was used for multiple imputation of missings.[29] In order to compare HRQoL, we calculated means and their corresponding 95% confidence intervals (CIs) for the PANQOL, SF-36, EQ-5D and VAS scores. Because treatment might influence the HRQoL, we stratified the scores according to management strategy (W&S versus active treatment) and Koos stages. Minimal clinically important differences, defining the minimal difference in scores that is found to be important for patients, have recently been determined for the PANQOL and SF-36 in VS patients.[30] Differences in scores between groups were compared to these minimal clinically important differences. Furthermore, a sensitivity analysis was performed by comparing the current results with 1) a complete case analysis and 2) an analysis in which only patients were included whose most recent MRI was within 12 months prior to the survey.

RESULTS

Population characteristics

Of the 252 patients who were invited to participate, 174 (69%) completed the questionnaire. The electronic questionnaire was completed by 144 out of 206 patients (70%), the paper questionnaire by 30 out of 46 patients (65%). Respondents had a mean age of 59 years (SD: 11.7, range: 22-89) and 50% was male. Most patients had a VS on the right side (55%). The prevalence of the different Koos stages was 36%, 25%, 16% and 24% for Koos 1-4, respectively, i.e. smaller VSs were more common. The population characteristics per Koos stage are presented in Table 7.1. At the time of completing the questionnaire most patients were followed in a W&S strategy. Active treatment (microsurgery or stereotactic radiosurgery) had most often been prescribed for patients with a Koos 4 VS and only once among Koos 1 patients. We chose not to display HRQoL for this latter patient, because of its restricted

usefulness. The average time between 'diagnosis and survey' was 630 days (SD: 290) and the average time between 'treatment and survey' for the 44 treated patients was 390 days (SD: 226). The most recent MRI at which the Koos stage had been determined was generally acquired within one year prior to the survey (mean: 336 days, SD: 229, median: 296 days).

Table 7.1 Patient characteristics of the study population according to their Koos stage (N=174).

	Koos stage			
Patient characteristics	**Koos 1 n=62**	**Koos 2 n=43**	**Koos 3 n=28**	**Koos 4 n=41**
Age (years (SD))	60.7 (11.0)	58.6 (11.2)	59.8 (10.2)	56.0 (13.8)
Gender (n (%))				
Male	31 (50.0)	27 (62.8)	11 (39.3)	18 (43.9)
Management (n (%))				
W&S	61 (98.4)	32 (74.4)	21 (75.0)	16 (39.0)
MS	0 (0)	2 (4.7)	0 (0.0)	13 (31.7)
SRS	1 (1.6)	9 (20.9)	7 (25)	12 (29.3)
Mean time between (days (SD)):				
Diagnosis and survey	571 (284)	652 (309)	618 (310)	704 (256)
Diagnosis and treatment	233 (NA)	583 (248)	585 (203)	265 (258)
Treatment and survey	757 (NA)	243 (144)	364 (218)	442 (228)
Most recent MRI and survey	309 (220)	323 (198)	282 (219)	429 (259)

n; number of patients; SD, standard deviation; W&S, wait and scan; MS, microsurgery; SRS, stereotactic radiosurgery; NA, not applicable since n=1.

Questionnaires

The PANQOL, SF-36, EQ-5D and VAS had a completion rate of 87%, 86%, 77% and 66%, respectively.

PANQOL

Figure 7.1 and Supplemental table S7.1 provide an overview of the results of the PANQOL according to Koos stage and management strategy. PANQOL scores were lowest for treated Koos 2 patients followed by Koos 4. Differences between Koos 1 and 2 stages for patients in a W&S strategy were limited and none exceeded the minimal clinically important difference (which are also displayed in supplemental table S7.1). Koos 2 and 3 patients in a W&S strategy had similar scores, except for the Hearing domain, which approached the minimal clinically important difference in favour of Koos 3. The difference between Koos 3 and 4 patients in a W&S strategy was only apparent on the Pain domain.

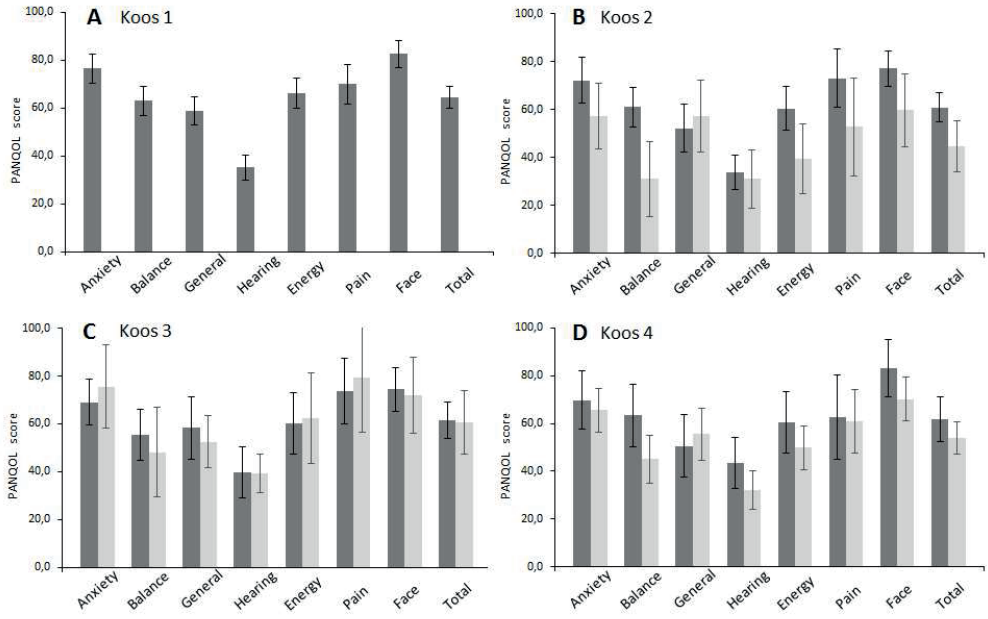


Figure 7.1 Penn Acoustic Neuroma Quality of Life (PANQOL) scores per domain for the different management strategies according to Koos stage: patients with (A) Koos 1, (B) Koos 2, (C) Koos 3, and (D) Koos 4. Scores can range from 0-100, higher scores denoting better health related quality of life. Values are presented as means, and error bars indicate 95% confidence intervals for the mean group score. Dark bars: wait and scan strategy. Light bars: treatment (i.e. microsurgery or stereotactic radiosurgery). NB: only one patient in the Koos 1 group had been treated, therefore results are not shown.

When comparing treated and W&S patients within a Koos stage, differences were largest for Koos 2 patients and in favour of the W&S group on all domains. A difference larger than the minimal clinically important difference was found for Anxiety, Balance, Hearing, Energy, Pain, and Total scores. Within Koos 3, treated and W&S patients had similar scores. For Koos 4, patients in the W&S strategy generally scored higher compared to treated patients, with Balance and Hearing domains exceeding the minimal clinically important difference.

When comparing treated patients across Koos stages, Koos 2 and 3 showed the largest differences in favour of Koos 3 with most scores exceeding the minimal clinically important difference. Treated Koos 3 patients scored slightly better than treated Koos 4 patients, with Pain and Hearing domains exceeding the minimal clinically important difference.

SF-36

SF-36 domain scores according to Koos stage and management strategy are displayed in figure 7.2 and Supplemental table S7.2. The lowest scores were found for treated Koos 2 patients on all domains. When comparing results of patients in a W&S strategy, results differed merely between Koos 1 and 2 except from Physical Role limitations. Koos 3 patients in a W&S strategy scored higher than Koos 2 on most domains, but the differences were most apparent in Physical Role limitations. Scores for the Koos 4 W&S group were slightly lower than Koos 3 with the largest differences in Physical Role limitations.

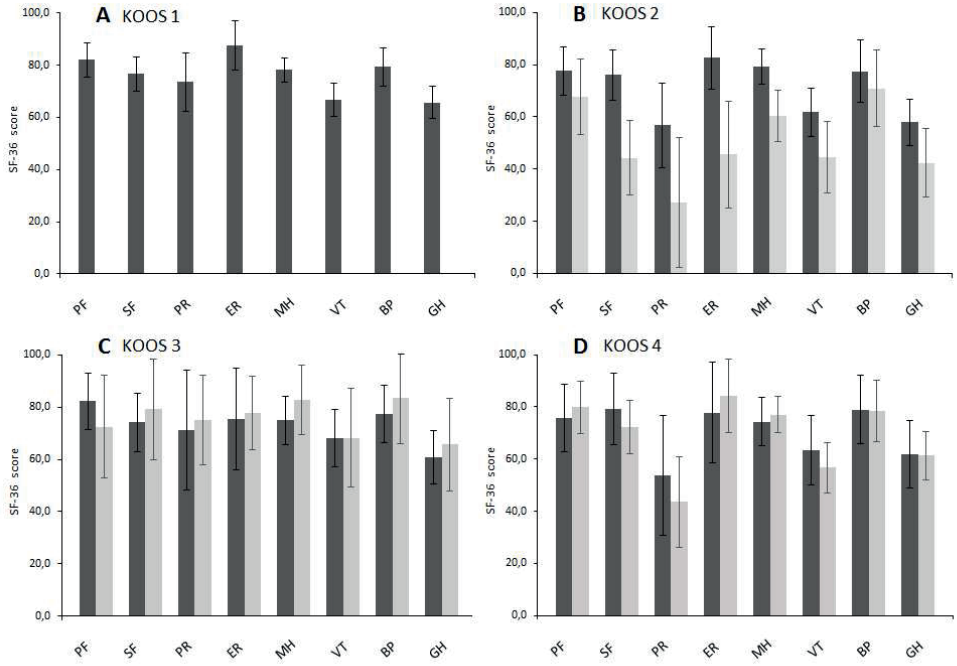


Figure 7.2 Self-administered Short-Form SF-36 health Survey (SF-36) scores per domain for the different Koos stages: patients with (A) Koos 1, (B) Koos 2, (C) Koos 3 and (D) Koos 4. Scores can range from 0-100, higher scores denoting better health related quality of life. Values are presented as means, and error bars indicate 95% confidence intervals for the mean group score. Dark bars: *wait and scan strategy*. Light bars: treatment (i.e. microsurgery or stereotactic radiosurgery). NB: only one patient in the Koos 1 group had been treated, therefore results are not shown. PF, Physical Functioning; SF, Social Functioning; PR, Physical Role limitations; ER, Emotional Role limitations; MH, Mental Health; VT, Vitality; BP, Bodily pain; GH, General health perceptions.

When comparing patients within the same Koos stage, differences were most apparent for Koos 2, with lower scores for treated patients. Within Koos 3 and 4 differences between treatment and W&S were relatively small.

Besides the low scores for treated Koos 2, comparison of treated patients across Koos stages showed small differences. Physical Role limitations scores were relatively low in treated Koos 4 patients.

Physical and Mental Composite scores irrespective of Koos stage or management strategy were 69.4 [95%CI: 65.9-72.9] and 72.3 [95%CI: 69.1-75.5], respectively. Across subgroups, Physical and Mental Composite Scores were lowest for treated Koos 2 patients, the difference with other subgroups exceeding the minimal clinically important differences for both scores. Additionally, difference in Physical Composite Scores between Koos 1 and 2 patients within W&S exceeded the minimal clinically important differences, in favour of Koos 1.

Utility and VAS scores

The mean utility and VAS scores according to Koos stage and management strategy are depicted in figure 7.3 and Supplemental table S7.3. EQ-5D derived utility scores were slightly higher than SF-36 utility scores for most subgroups and self-reported VAS scores were lower than EQ-5D derived utility scores.

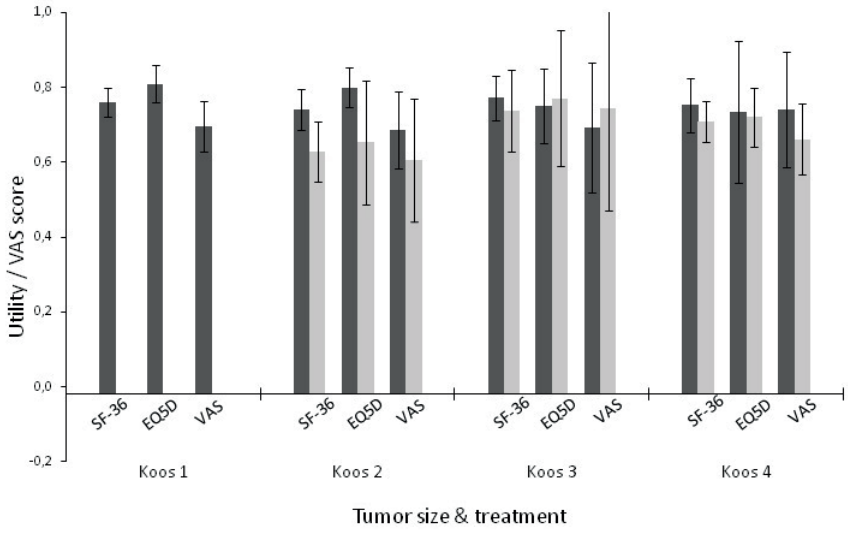


Figure 7.3 Utility (based on SF-36 and EQ-5D) and VAS scores for Koos stages 1 to 4.

Scores can range from 0-1, higher scores denoting better health related quality of life. Values are presented as means, and error bars indicate 95% confidence intervals for the mean group score. Dark bars: *wait and scan* strategy. Light bars: treatment (i.e. microsurgery or stereotactic radiosurgery). VAS, Visual Analogue Scale. NB: only one patient in the Koos 1 group had been treated, therefore results are not shown.

When comparing patients from different Koos and management strategies, Koos 1 patients in a W&S strategy had the highest combination of VAS and utility scores, whereas treated Koos 2 patients had lowest utility and VAS scores. Patients with a Koos 3 lesion that had been treated showed a higher VAS score compared to those in a W&S strategy, while the opposite was found for Koos 4.

Sensitivity analyses performed by only including patients with an MRI within one year prior to the survey and complete cases showed similar results for most subgroups. There were, however, some differences in scores for treated Koos 3 and 4 patients compared to the initial results, but this did not influence our conclusions.

DISCUSSION

We studied the HRQoL of patients with a VS, stratified according to Koos stage and management strategy. Mean utility scores ranged from 0.673 for treated Koos 2 patients, to 0.829 for Koos 1 patients in a W&S strategy. When comparing different Koos stages within a W&S strategy, differences were small. Treated Koos 2 and 4 patients, however, had slightly lower scores than those in a W&S strategy in the same Koos stage. In Koos 2, these differences were most apparent. Treated Koos 3 patients scored similar to those in the W&S strategy.

The results of our study seem in agreement with other studies that used the PANQOL and/or SF-36 in a population with VS, stratified according to management strategy.[9, 10, 13, 19] Only one study also stratified on VS size and management strategy for the PANQOL and found comparable results.[10] In contrast to other studies, we stratified patients according to Koos stage and management strategy, and did not compare the different treatment modalities (stereotactic radiosurgery and microsurgery). For these reasons, it is difficult to make exact comparisons.

Strengths and limitations

As far as we are aware, this is the first study to investigate HRQoL differences including utility scores for patients with VS, stratified according to Koos stage and management strategy. By including patients diagnosed between 2014 and 2017 we obtained a large representative sample of patients. Since we used PANQOL, SF-36, and EQ-5D questionnaires, the HRQoL was measured with both a disease specific questionnaire and general health utility scores, improving generalizability and comparability of our results. Utility scores can be used to compare HRQoL between patients with VS and other diseases. Moreover, utilities are needed for cost-effectiveness studies. Information on HRQoL in specific subgroups enables modelling of cost-effectiveness of different management strategies.

Some potential limitations should also be discussed. First, treatment was not equally distributed among Koos stages, and the number of treated patients per Koos stage was relatively small. To improve power,

we clustered treated patients irrespective of the treatment modality, thus ignoring possible differences in HRQoL following microsurgery and stereotactic radiosurgery. Second, due to the cross-sectional nature of our study, we had no information on the HRQoL of patients prior to treatment. We were therefore unable to determine absolute treatment effects. Moreover, the fact that treated patients generally scored lower could be the result of confounding by indication, since patients with a lower HRQoL might have been more eager to receive treatment compared to others, especially for the lower Koos stages for which a W&S strategy is more often applied. Third, we had to exclude patients without available MRI images since Koos stages had to be determined by these MRI images, which might have induced selection bias. We only had to exclude three patients for this reason and they all had a clear contra-indication, i.e. a pacemaker or metal implant, so that MRI imaging was impossible. We therefore believe that selection bias is not a real issue in our study. Fourth, there was a time interval between sending the EQ-5D and the other two questionnaires for most patients. Clinical parameters had, however, not changed for the latter patients. We therefore believe this most likely will not have influenced our results. Fifth, time between MRI and completion of the survey could be an interfering factor, because Koos stages were determined by assessment of the most recent MRI. Our sensitivity analysis, in which we excluded MRIs that were acquired more than a year prior to the survey, did not result in a different conclusion.

Clinical implications and future perspective

The small differences in HRQoL between Koos stages of patients in a W&S strategy suggest that VS size has less impact than expected. Even if the Koos stage increases, HRQoL seems unaffected, which suggests that the decision to change from a W&S strategy to active treatment could be delayed to a higher Koos stage with regard to HRQoL. It should, however, be kept in mind that VSs larger than 3 cm usually cannot be treated with stereotactic radiosurgery, thus extensive growth reduces treatment options to microsurgery only.[31] Furthermore, it should be noted that increased tumour volumes are associated with smaller success rates following stereotactic radiosurgery, especially for lesions ≥6 cm^3.[3, 32]

To improve HRQoL of patients with VS, patient counselling regarding expectations, coping strategies and other psychosocial support could be useful to manage symptoms.[5, 7, 12, 33, 34] It would be helpful to include HRQoL questionnaires in the W&S strategy to incorporate well-being of patients with VS next to clinical parameters such as VS growth. Patients with a severe reduction in HRQoL might be detected and helped with counselling and/or treatment. Prospective follow-up studies are needed to clarify the long-term impact of VS growth and (type of) treatment on the HRQoL.

In conclusion, our results show that HRQoL does not differ between VS patients in a W&S strategy with different Koos stages, and that treated patients generally score lower than wait and scan patients with the same Koos stage. Whether small and medium-sized VSs might be able to grow without having a large impact on the HRQoL should be further investigated using longitudinal data. Naturally, severe brainstem compression requires immediate treatment, and considerations regarding success rates of treatment in relation to VS size should always be taken into account.

REFERENCES

1. Broomfield SJ, O'Donoghue GM. Self-reported symptoms and patient experience: A British Acoustic Neuroma Association survey. 2016;30(3):294-301. doi: 10.3109/02688697.2015.1071323.
2. Brooker JE, Fletcher JM, Dally MJ, Briggs RJS, Cousins VC, Malham GM, et al. Factors associated with symptom-specific psychological and functional impact among acoustic neuroma patients. J Laryngol Otol. 2014;128:S16-S26. doi: 10.1017/S0022215113003216.
3. Hio S, Kitahara T, Uno A, Imai T, Horii A, Inohara H. Psychological condition in patients with an acoustic tumor. Acta Otolaryngol. 2012;133:42-6. doi: 10.3109/00016489.2012.709322. PubMed PMID: 23039793.
4. Stangerup SE, Tos M, Thomsen J, Caye-Thomasen P. True incidence of vestibular schwannoma? 2010;67:1335-40. doi: 10.1227/NEU.0b013e3181f22660.
5. Carlson ML, Habermann EB, Wagie AE, Driscoll CL, Van Gompel JJ, Jacob JT, et al. The Changing Landscape of Vestibular Schwannoma Management in the United States—A Shift Toward Conservatism. Otolaryngol Head Neck Surg. 2015;153:440-6. doi: 10.1177/0194599815590105.
6. Gait C, Frew EJ, Martin TPC, Jowett S, Irving R. Conservative management, surgery and radiosurgery for treatment of vestibular schwannomas: A model-based approach to cost-effectiveness. Clin Otolaryngol. 2014;39:22-31. doi: 10.1111/coa.12205.
7. Patnaik U, Prasad SC, Tutar H, Giannuzzi AL, Russo A, Sanna M. The Long-Term Outcomes of Wait-and-Scan and the Role of Radiotherapy in the Management of Vestibular Schwannomas. Otol Neurotol. 2015;36:638-46. doi: 10.1097/MAO.0000000000000657.
8. Kleijwegt M, Ho V, Visser O, Godefroy W, van der Mey A. Real Incidence of Vestibular Schwannoma? Estimations From a National Registry. Otol Neurotol. 2016;37:1411-7. doi: 10.1097/MAO.0000000000001169.
9. Carlson ML, Tveiten OV, Driscoll CL, Goplen FK, Neff BA, Pollock BE, et al. Long-term quality of life in patients with vestibular schwannoma: an international multicenter cross-sectional study comparing microsurgery, stereotactic radiosurgery, observation, and nontumor controls. Journal of neurosurgery. 2015;122:833-42. doi: 10.3171/2014.11.JNS14594.
10. Soulier G, van Leeuwen BM, Putter H, Jansen JC, Malessy MJA, van Benthem PPG, et al. Quality of Life in 807 Patients with Vestibular Schwannoma: Comparing Treatment Modalities. Otolaryngol Head Neck Surg. 2017:019459981769580. doi: 10.1177/0194599817695800.
11. Robinett ZN, Walz PC, Miles-Markley B, Moberly AC, Welling DB. Comparison of Long-term Quality-of-Life Outcomes in Vestibular Schwannoma Patients. Otolaryngol Head Neck Surg. 2014;150:1024-32. doi: 10.1177/0194599814524531.
12. Jufas N, Flanagan S, Biggs N, Chang P, Fagan P. Quality of Life in Vestibular Schwannoma Patients Managed by Surgical or Conservative Approaches. Otol Neurotol. 2015;36(7):1245-54.
13. McLaughlin EJ, Bigelow DC, Lee JYK, Ruckenstein MJ. Quality of Life in Acoustic Neuroma Patients. Otol Neurotol. 2015;36:653-6. doi: 10.1097/MAO.0000000000000674.
14. Wangerid T, Bartek J, Svensson M, Förander P. Long-term quality of life and tumour control following gamma knife radiosurgery for vestibular schwannoma. Acta Neurochir. 2014;156:389-96. doi: 10.1007/s00701-013-1924-5.

15. Cheng S, Naidoo Y, Da Cruz M, Dexter M. Quality of life in postoperative vestibular schwannoma patients. 2009;119:2252-7. doi: 10.1002/lary.20217. PubMed PMID: 19753619.
16. Wagner JN, Glaser M, Wowra B, Muacevic A, Goldbrunner R, Cnyrim C, et al. Vestibular function and quality of life in vestibular schwannoma: Does size matter? 2011;30(2):55. doi: 10.3389/fneur.2011.00055.
17. Tveiten OV, Carlson ML, Link MJ, Lund-Johansen M. Audiovestibular Handicap and Quality of Life in Patients With Vestibular Schwannoma and "Excellent" Hearing. Neurosurgery. 2017;80(3):386-392. doi: 10.1227/NEU.0000000000001238.
18. Koos WT, Day JD, Matula C, Levy DI. Neurotopographic considerations in the microsurgical treatment of small acoustic neurinomas. J Neurosurg. 1998;88(3):506-12. doi: 10.3171/jns.1998.88.3.0506.
19. Shaffer BT, Cohen MS, Bigelow DC, Ruckenstein MJ. Validation of a disease-specific quality-of-life instrument for acoustic neuroma: The penn acoustic neuroma quality-of-life scale. Laryngoscope. 2010;120(8):1646-54. doi: 10.1002/lary.20988.
20. van Leeuwen BM, Herruer JM, Putter H, Jansen JC, van der Mey AGL, Kaptein AA. Validating the Penn Acoustic Neuroma Quality of Life Scale in a Sample of Dutch Patients Recently Diagnosed With Vestibular Schwannoma. Otol Neurotol. 2013;34:952-7. doi: 10.1097/MAO.0b013e31828bb2bb.
21. Ware JE KM, Keller SK. SF-36® Physical and Mental Health Summary Scales: A User's Manual. Boston: The Health Institute; 1994.
22. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ. 2002;21:271-92. doi: 10.1016/S0167-6296(01)00130-8.
23. Ara R, Brazier J. Deriving an algorithm to convert the eight mean SF-36 dimension scores into a mean EQ-5D preference-based score from published studies (where patient level data are not available). Value Health. 2008;11(7):1131-43. doi: 10.1111/j.1524-4733.2008.00352.x.
24. McCabe C, Brazier J, Gilks P, Tsuchiya A, Roberts J, O'Hagan A, et al. Using rank data to estimate health state utility models. J Health Econ. 2006;25(3):418-31. doi: 10.1016/j.jhealeco.2005.07.008.
25. EuroQol-Group. Coastal dynamics. 1984;68:8-13. doi: 10.1093/infdis/jit783.
26. Guideline for economic evaluations in healthcare: Zorginstituut Nederland; 2016 [28th November, 2017]. 1-45]. Available from: https://www.zorginstituutnederland.nl/over-ons/publicaties/publicatie/2016/02/29/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg.
27. Position statement on use of the EQ-5D-5L valuation set: National Institute For Health And Care Excellence (NICE). Available from: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisal-guidance/eq5d5l_nice_position_statement.pdf.
28. Versteegh MM, Vermeulen KM, Evers SMAA, de Wit GA, Prenger R, Stolk EA. Dutch Tariff for the Five-Level Version of EQ-5D. Value Health. 2016;19(4):343-52. doi: 10.1016/j.jval.2016.01.003.
29. van Buuren S, Groothuis-Oudshoorn, K. mice: Multivariate Imputation by Chained Equations in R. J Stat Softw. 2011;45(3):1-67. doi: 10.18637/jss.v045.i03.
30. Carlson ML, Tveiten ØV, Yost KJ, Lohse CM, Lund-Johansen M, Link MJ. The Minimal Clinically Important Difference in Vestibular Schwannoma Quality-of-Life Assessment. Otolaryngol Head Neck Surg. 2015;153:202-8. doi: 10.1177/0194599815585508.

31. Pai I, Bowman J, Thomas N, Kitchen N, Strong A, Obholzer R, et al. Management of large and giant vestibular schwannomas. Skull Base. 2011;21(6):379-84 doi: 10.1055/s-0031-1287680.

32. Klijn S, Verheul JB, Beute GN, Leenstra S, Mulder JJ, Kunst HP, et al. Gamma Knife radiosurgery for vestibular schwannomas: evaluation of tumor control and its predictors in a large patient cohort in The Netherlands. J Neurosurg. 2016;124(6):1619-26. doi: 10.3171/2015.4.JNS142415.

33. van Leeuwen BM, Borst JM, Putter H, Jansen JC, van der Mey AGL, Kaptein AA. Emotional Intelligence in Association With Quality of Life in Patients Recently Diagnosed With Vestibular Schwannoma. Otol Neurotol. 2014;35:1650-7. doi: 10.1097/MAO.0000000000000423.

34. Van Leeuwen BM, Herruer JM, Putter H, Van Der Mey AGL, Kaptein AA. The art of perception: Patients drawing their vestibular schwannoma. Laryngoscope. 2015;125:2660-7. doi: 10.1002/lary.25386.

SUPPLEMENTAL CONTENT

Supplemental Table 7.1 Mean and 95% confidence intervals of PANQOL domains stratified by Koos stage and treatment. Penn Acoustic Neuroma Quality of Life (PANQOL) scores per domain for the different Koos stages. Scores can range from 0-100, higher scores denoting better HRQoL. W&S, wait and scan strategy; TR, treatment strategy (microsurgery or stereotactic radiosurgery); MCID, minimal clinically important difference. † Results for one treated patient in the Koos 1 group are not shown.

Domains	**Koos 1† (n=62)**	**Koos 2 (n=43)**		**Koos 3 (n=28)**		**Koos 4 (n=41)**		**MCID[30]**
	W&S (n=61)	W&S (n=32)	TR (n=11)	W&S (n=21)	TR (n=7)	W&S (n=14)	TR (n=27)	
Anxiety	76.5 [70.5-82.5]	72.3 [62.7-82.0]	57.3 [43.7-70.9]	69.0 [59.4-78.6]	75.6 [58.1-93.0]	69.8 [57.8-81.8]	65.6 [56.5-74.7]	10.8
Balance	63.0 [56.8-69.3]	61.1 [52.3-69.0]	31.0 [15.3-46.7]	55.3 [44.7-65.9]	48.2 [29.6-66.8]	63.3 [50.0-76.5]	45.2 [35.2-55.2]	16.4
General	58.9 [53.0-64.8]	52.3 [42.5-62.1]	50.2 [35.3-65.2]	58.3 [48.5-68.1]	52.5 [35.4-69.6]	50.5 [37.5-63.6]	55.5 [44.5-66.4]	15.4
Hearing	35.2 [29.9-40.5]	33.8 [26.7-40.9]	21.2 [8.9-33.6]	39.8 [30.8-48.8]	39.3 [24.1-54.5]	43.5 [32.7-54.2]	32.1 [24.1-40.1]	6.2
Energy	66.2 [59.9-72.5]	60.4 [51.2-69.5]	39.4 [24.8-54.0]	60.3 [50.0-70.6]	62.3 [42.8-80.7]	60.6 [47.7-73.4]	49.9 [40.7-59.2]	13.4
Pain	70.0 [61.8-78.2]	73.0 [60.8-85.2]	52.8 [32.4-73.1]	73.8 [60.1-87.4]	79.2 [54.2-100]	62.5 [44.9-80.1]	60.7 [47.4-74.0]	10.6
Face	82.6 [76.8-88.3]	77.2 [69.7-84.7]	59.8 [44.7-75.0]	74.4 [65.2-83.5]	71.9 [55.9-87.9]	83.2 [71.3-95.1]	70.1 [61.0-79.3]	20.2
Total	**64.5 [60.1-68.9]**	**61.0 [54.7-67.3]**	**44.6 [34.0-55.1]**	**61.5 [54.0-69.1]**	**60.8 [47.5-74.0]**	**61.7 [52.2-71.1]**	**53.8 [47.1-60.5]**	**11.2**

Supplemental Table 7.2 Mean and 95% confidence intervals of the domains of the SF-36 stratified by Koos stage and treatment status.

SF-36 domains	Koos 1[†] (n=62)	Koos 2 (n=43)		Koos 3 (n=28)		Koos 4 (n=41)		MCID[30]
	W&S (n=61)	W&S (n=32)	TR (n=11)	W&S (n=21)	TR (n=7)	W&S (n=14)	TR (n=27)	
PF	82.1 [75.6-88.6]	77.6 [68.3-86.9]	67.7 [53.1-82.3]	82.3 [71.4-93.1]	72.5 [52.8-92.2]	75.7 [62.8-88.6]	79.8 [69.7-89.9]	
SF	76.7 [70.1-83.3]	76.0 [66.4-85.7]	44.3 [30.0-58.6]	74.3 [63.1-85.5]	79.2 [59.8-98.5]	79.2 [65.5-92.8]	72.2 [62.1-82.3]	
PR	73.6 [62.4-84.9]	56.7 [40.5-72.9]	27.3 [2.4-52.2]	71.1 [52.1-90.0]	75.0 [41.3-100]	53.8 [30.9-76.8]	43.5 [26.2-60.7]	
ER	87.6 [78.2-97.0]	82.7 [69.2-93.1]	45.5 [25.2-65.8]	75.4 [60.0-90.9]	77.8 [50.3-100]	77.8 [58.3-97.2]	84.1 [70.0-98.1]	
MH	78.2 [73.6-82.8]	79.3 [72.7-86.0]	60.4 [50.6-70.2]	74.9 [67.2-82.6]	82.7 [69.4-95.9]	74.3 [64.9-83.7]	76.9 [70.0-83.8]	
VT	66.7 [60.2-73.1]	61.7 [52.3-71.0]	44.5 [30.7-58.4]	68.1 [57.2-78.9]	68.3 [49.6-87.1]	63.3 [50.1-76.6]	56.6 [46.8-66.4]	
BP	79.3 [72.0-86.5]	77.5 [65.6-89.4]	70.9 [56.3-85.4]	78.9 [67.8-90.1]	83.7 [64.7-100]	77.4 [64.3-90.6]	78.4 [66.8-90.0]	
GH	65.6 [59.4-71.8]	57.9 [49.0-66.8]	42.3 [29.1-55.4]	60.8 [50.6-71.1]	65.8 [48.1-83.6]	61.8 [48.7-74.9]	61.4 [52.1-70.6]	
PCS	74.0 [67.9-80.0]	65.7 [57.2-74.2]	52 [38.4-65.6]	72.4 [62.5-82.3]	73.2 [56.2-90.3]	67.1 [54.9-79.3]	65.3 [56.4-74.3]	7.8
MCS	75.3 [70.4-80.3]	72.5 [65.5-79.5]	48.7 [37.2-60.1]	72.5 [64.0-81.0]	74.9 [59.8-90.1]	73.0 [62.3-83.6]	71.5 [64.0-78.9]	6.7

Self-administered Short-Form SF-36 health Survey (SF-36) scores per domain for Koos stages 1 to 4. Scores can range from 0-100, higher scores denoting better HRQoL. W&S, wait and scan strategy; TR, treatment strategy (i.e. microsurgery or stereotactic radiosurgery); PF, Physical Functioning; SF, Social Functioning; PR, Physical Role limitations; ER, Emotional Role limitations; MH, Mental Health; VT, Vitality; BP, Bodily pain; GH, General health perceptions; PCS, Physical Composite Score; MCS, Mental Composite Score. [†] Results for one treated patient in the Koos 1 group are not shown.

Supplemental Table 7.3 Means and 95% confidence intervals of utility- and VAS scores stratified by Koos stage and treatment status.

Utility/VAS	Koos 1† (n=62)	Koos 2 (n=43)		Koos 3 (n=28)		Koos 4 (n=41)	
	W&S (n=61)	W&S (n=32)	TR (n=11)	W&S (n=21)	TR (n=7)	W&S (n=14)	TR (n=27)
SF-36	0.780 [0.741-0.819]	0.761 [0.707-0.815]	0.647 [0.567-0.728]	0.792 [0.732-0.852]	0.757 [0.647-0.866]	0.771 [0.699-0.843]	0.729 [0.674-0.784]
EQ-5D	0.829 [0.787-0.871]	0.820 [0.764-0.876]	0.673 [0.570-0.775]	0.770 [0.693-0.848]	0.791 [0.673-0.908]	0.755 [0.649-0.861]	0.740 [0.674-0.805]
VAS	74.0 [67.1-81.0]	69.2 [52.6-85.8]	62.0 [42.9-81.1]	72.3 [56.8-87.8]	77.4 [53.6-100]	75.0 [57.9-92.1]	70.5 [58.2-82.8]

Short-Form SF-36 health Survey (SF-36) and EQ-5D utility scores and VAS scores for Koos stages 1 to 4. Utility scores can range from 0.00-1.00, from worst to perfect health (a higher score denotes better HRQoL) and VAS scores can range from 0-100 (higher score means higher self-reported quality of life). SF-36, 36-item short-form health survey; EQ-5D, EuroQol five domains' questionnaire; VAS, visual analogue scale; W&S, wait and scan. † Only one patient in the Koos 1 group was treated, results not shown.

CHAPTER 8

In search of the most cost-effective monitoring strategy for vestibular schwannoma: a decision analytical modelling study

Scholte M.
Hentschel M.
Hannink G.
Kunst H.
Steens S.
Rovers M.
Grutters J.

Published as:
Clin Otolaryngol. 2019;44:525–533. doi: 10.1111/coa.13326.

ABSTRACT

Objectives

To assess the cost-effectiveness of frequently used monitoring strategies for vestibular schwannoma (VS).

Design

A state transition model was developed to compare six monitoring strategies for patients with VS: lifelong annual monitoring; annual monitoring for the first 10 years after diagnosis; scanning at 1-5, 7, 9, 12, 15 years after diagnosis and subsequently every 5 years; a personalized monitoring strategy for small and large tumours; scanning at 1, 2 and 5 years after diagnosis and no monitoring. Input data were derived from literature and expert opinion. Quality-adjusted life years (QALYs) and health care costs of each strategy were modelled over lifetime. Net monetary benefits (NMBs) were calculated to determine which strategy provided most value for money. Sensitivity analyses were performed to address uncertainty.

Results

Omitting monitoring is least effective with 18.23 (95%CI 16.84 – 19.37) QALYs per patient and lifelong annual monitoring is most effective with 18.66 (95%CI 17.42 – 19.65) QALYs. Corresponding costs were €6,526 (95%CI 5,923 – 7,058) and €9,429 (95%CI 9,197 – 9,643) per patient, respectively. Lifelong annual monitoring provided the best value with an NMB of €363,765 (339,040 – 383,697), but the overall probability of being most cost-effective compared to the other strategies was still only 23%. Sensitivity analysis shows that there is large uncertainty in the effectiveness of all strategies, with largely overlapping 95% confidence intervals for all strategies.

Conclusions

Due to the largely overlapping 95% confidence intervals of all monitoring strategies for VS, it is unclear which monitoring strategy provides most value for money at this moment.

INTRODUCTION

Vestibular schwannomas (VSs) are benign, slow growing tumours originating from Schwann cells of the vestibular part of the eighth cranial nerve.[1] They represent 6% of all intracranial tumours.[2] Patients with sporadic VS most commonly present between their 40s and 60s, some with small intracanalicular tumours and others with larger extrameatal tumours expanding into the cerebellopontine angle.[3]

The possibility to observe tumour development with magnetic resonance imaging (MRI) has led to the adoption of a "wait and scan" or "monitoring" policy in addition to treatment options; microsurgery and stereotactic radiosurgery (SRS).[3-5] Currently, treatment is mainly indicated for large and/or growing tumours. Due to monitoring it is known that approximately two thirds of VSs grow slowly or do not grow at all, which resulted in a decline of initial treatment and an increase in conservative management.[6, 7] At present, it is not possible to predict which VSs pose a threat and which can be safely left without intervention, therefore all patients undergo a monitoring strategy with extensive MRI scanning.

MRI scans are costly and, with a large proportion of patients in a monitoring strategy, contribute significantly to the high costs involved with VS.[8] Multiple monitoring protocols are used alongside each other, often lacking evidence of effectiveness.[9, 10] Therefore, a cost-effectiveness model was developed to determine the added value of monitoring strategies for VS.

METHODS

Ethical considerations

This modelling study was based on published literature and did not involve human subjects, therefore ethical approval or informed consent was not required.

Model development

To simulate the follow-up of patients in a monitoring strategy, we developed a state transition model in which we simulated costs and quality of life associated with multiple monitoring strategies for VS. The target population comprised VS patients who were initially assigned to the monitoring strategy, i.e. tumours smaller than Koos 4 at time of diagnosis or small Koos 4 tumours without symptoms of brainstem compression (hydrocephalus and symptoms caused by cranial nerve failure, e.g. swallowing problems).[11] The model starts at the age of 55, the mean age of diagnosis.[12] We assumed every patient was eligible for MRI and loss to follow-up did not occur. Based on clinical guidelines and expert interviews, the model was designed in a way that it resembles the clinical situation.

A state transition model describes the conditions that patients can be in (health states), how they can move among such states (transitions) and how likely such moves are (transition probabilities). Health

states in the model were "Koos 1", "Koos 2", "Koos 3", "Koos 4", "microsurgery", "post microsurgery", "SRS", "post SRS", and "dead" (Figure 8.1). Patients were assumed to enter the model via one of the Koos states. The Koos states represented different tumour sizes in the monitoring strategy, for patients who were not treated for their VS. We added treatment options to model the consequences of tumour growth. Tumour growth was defined as growth to the next Koos state, and in case of a Koos 4 tumour as growth to a Koos 4 tumour with brainstem compression. Growth within a Koos state and tumour shrinkage were not considered growth. Small and medium sized tumours (Koos 1 and 2) which showed growth continued to be monitored without treatment. Patients received SRS when growth from Koos 2 to 3 was detected and SRS or microsurgery when growth to Koos 4 was detected. We assumed that when a growing VS was not detected and treated in time, the patient would visit the hospital with symptoms of brainstem compression and would then receive microsurgical treatment. Quality of life was lower in the year prior to surgical treatment.

The model had a cycle length of 1 year with a lifelong time horizon. We applied discount rates to costs and effects, to adjust future costs and effects to present values. A discount rate of 4% was applied to costs and 1.5% to quality-adjusted life years (QALYs), according to Dutch guidelines.[13]

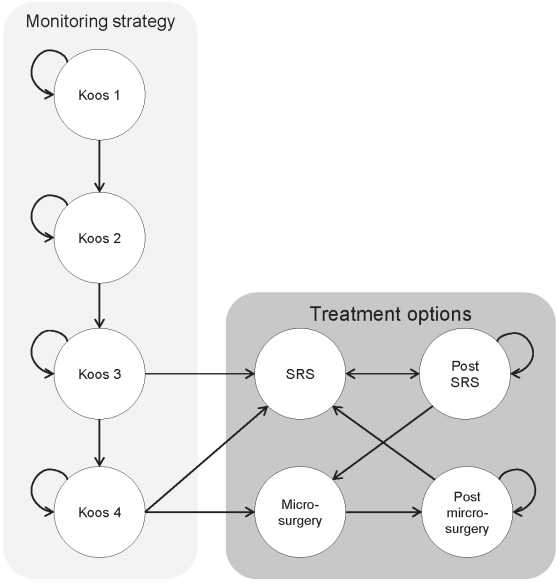


Figure 8.1 Influence diagram of the Markov model.

Patients could enter the model via one of the Koos states in the monitoring strategy. Koos 1 corresponds to an intracanalicular VS, Koos 2 to an extracanalicular VS without brainstem contact, Koos 3 to VS with brainstem contact and Koos 4 corresponds to VS that compresses the brainstem. When tumour growth was present, patients entered the next Koos state. In case of Koos state 3 and 4, patients exited the monitoring strategy when tumour growth was

detected on MRI. Leaving the monitoring strategy meant transition to one of the treatment options; stereotactic radiosurgery (SRS) or microsurgery. After treatment, patients were monitored for tumour growth. If tumour growth was detected after treatment, patients could receive additional treatment. The health state "dead" is not displayed, but could be entered from all health states.

Model validation

We verified the model's validity using the AdViSHE checklist.[14] This checklist covers five aspects of validation: conceptual model, input data, computerized model and operational validation and other validation techniques. The conceptual model was tested on its face validity (the model's appropriateness to represent the clinical process/disease) by consulting otolaryngology, radiology, SRS and neurosurgery experts in the Netherlands. The conceptual model was also cross-validated with other VS models in literature, however, no specific health-economic models for monitoring strategies were found. Face validity of the input data was tested by consulting the above-mentioned experts. The computerized model was validated by extreme value testing, to detect possible coding errors. Operational validity was tested by discussing the model outcomes with the above-mentioned experts. In addition, sensitivity analyses were performed to validate the outcomes with alternative input data. Last, the model was checked for inconsistencies by an independent expert.

Strategies

We modelled multiple monitoring strategies for the follow-up of VS: lifelong annual monitoring; annual monitoring for the first 10 years after diagnosis; scanning at 1-5, 7, 9, 12, 15 years after diagnosis and subsequently every 5 years; a personalized monitoring strategy for small and large tumours; scanning at 1, 2 and 5 years after diagnosis and no monitoring. In the personalized monitoring strategy, small tumours (Koos 1 & 2) are monitored 1-3, and 5 years following diagnosis and large tumours (Koos 3 & 4) are monitored 1-5, 8, 11 and 16 years following diagnosis. A strategy without monitoring was modelled to evaluate the consequences of omitting monitoring, since there is discussion about the added value of current monitoring strategies.[9] In the no monitoring strategy, we assumed that if symptoms of brainstem compression occurred, patients would visit the hospital and undergo microsurgery. We used conservative assumptions for this strategy: patients acquired brainstem compression when ≥2mm growth in Koos 4 occurred, quality of life was low for one year when brainstem compression occurred, and costs and consequences of microsurgery were adapted to large tumours by assuming a complication rate of 25% instead of 12.5%.[15]

Transition probabilities

Probabilities were derived from literature and expert opinion (Table 8.1). All expert based values were confirmed by at least two experts. Key inputs were the initial probabilities that divided patients over the Koos states, which were derived from Stangerup et al. 2010, i.e. 34.7%, 32.2%, 32.2%, and 0.9% for Koos 1, Koos 2, Koos 3, and Koos 4, respectively.[3] Transition among Koos states was defined by the probability of tumour growth to the next Koos state and the probability to have ≥2 mm growth in Koos 4. These probabilities were derived from a large (n=1217) retrospective study conducted in our hospital

(Supplemental Figure 8.1). For each Koos state, follow-up was at least 9 years. Thereafter, we assumed tumour growth not to occur.

Table 8.1 Model parameters

Parameter	Value†	Source
Probabilities		
Koos 1	0.347 (α 112, β 211)	Stangerup et al. 2010[3]
Koos 2	0.322 (α 104, β 219)	Stangerup et al. 2010[3]
Koos 3	0.322 (α 104, β 219)	Stangerup et al. 2010[3]
Koos 4	0.009 (α 3, β 320)	Stangerup et al. 2010[3]
Dead	Standard mortality rates	Statistics Netherlands[24]
Tumour growth to the next Koos state	Supplemental Figure 8.1	Patient cohort Radboudumc
SRS after growth in the Koos 3 state	1.00	Expert opinion
Microsurgery after growth in the Koos 4 state	0.900	Expert opinion
Microsurgery complications	0.125	Sughrue et al. 2011[15]
Death as a consequence of microsurgery	0.002	Sughrue et al. 2011[15]
Death as a consequence of SRS	0	Klijn et al. 2015[25]
Growth after microsurgery	0.003	Godefroy et al. 2009[26]
Growth after SRS	0.006	Klijn et al. 2015[25]
Microsurgery in case of growth in the post SRS state	0.400	Expert opinion
SRS in case of growth in the post microsurgery state	1.00	Expert opinion
Costs		
Consultation – tertiary hospital	€167	Dutch Guideline for costing research[13]
Consultation – general hospital	€82	Dutch Guideline for costing research[13]
MRI brain	€211	Dutch Guideline for costing research[13]
Microsurgery – uncomplicated	€10,406	Dutch health care administration
Microsurgery – complicated	€13,068	Dutch health care administration
SRS	€8,876	Dutch health care administration
Post microsurgery	€151 (90% of all patients are followed in a tertiary hospital after microsurgery)	Expert opinion
Post SRS	€153 (85% of all patients are followed in a general hospital	Expert opinion

Utilities		
Monitoring strategy	Year 1-3: 0.831 (SD 0.244) Year 4-6: 0.826 (SD 0.244) Year 7-9: 0.821 (SD 0.244) Year 10-12: 0.816 (SD 0.244) Year 13 and onwards: 0.811 (SD 0.244)	Gait et al. 2014, Godefroy et al. 2009[18, 19]
Symptoms of brainstem compression	0.537 (SD 0.283)	Turel et al. 2015[27]
First year after microsurgery	0.688	Gait et al. 2014, Sughrue et al. 2011[15, 19]
First year after SRS	0.789	Gait et al. 2014, Klijn et al. 2015[19, 25]
Post microsurgery	0.789	Godefroy et al. 2007[28]
Post SRS	0.811	Varughese et al. 2012[29]
Dead	0	

† β-distributions were assigned to some of the parameters for use in the probabilistic sensitivity analysis. The characteristics of the β-distribution are presented between brackets, either as an SD or as an α and β value (where α represents the number of events in a sample and β the number of non-events).
MRI, magnetic resonance imaging; SD, standard deviation; SRS, stereotactic radiosurgery.

Cost information

The cost analysis was performed from a healthcare perspective, meaning all healthcare costs were included. Costs were assessed in Euros (€) and based on the 2017 price level. When available, costs were derived from the Dutch guideline for costing research.[13] Otherwise, unit costs were obtained from hospital fees. Key costs were consultation costs, €167 for tertiary hospitals and €82 for general hospitals, and MRI scans of €211. Complication costs are included in the total costs of microsurgery and SRS. To determine annual costs after microsurgery or SRS, a scanning protocol with scans at 1-5, 7, 9, 12, 15 years after microsurgery or SRS and subsequently every 5 years was assumed as this is the current protocol in our hospital (Table 8.1).

Effects

Effectiveness was measured in QALYs, which is a combination of quality of life (utility) and survival. A utility reflects quality of life on a 0 to 1 scale, with 0 representing death and 1 representing full health. Most quality of life values for VS patients in literature are derived from the SF-36 questionnaire. We used an algorithm to construct a utility value from the domain scores of the SF-36 questionnaire (Table 8.1).[16, 17] Quality of life in the monitoring strategy was assumed to gradually decline, since symptoms of asymmetric sensorineural hearing loss, tinnitus and vertigo often progress over time.[18] In the year of treatment, a lower utility was assumed due to potential complications (by calculating the weighted mean of treatment associated complications and corresponding utilities).

Analysis

A hypothetical cohort of 1000 patients was sent through the model to determine mean expected costs and effects (QALYs) per patient for each strategy. We compared the monitoring strategies to each other by calculating the average costs per QALY. We also calculated the net monetary benefit (NMB), which represents the value of a strategy in monetary terms. The strategy with the highest NMB represents the most cost-effective strategy. The NMB is calculated by multiplying the gained QALYs by the threshold value minus costs of the monitoring strategy. We used a threshold value of €20,000 per QALY, as recommended by the Dutch guidelines.[13]

We performed a scenario analysis in which an alternative treatment scheme is used. In this scheme, growing Koos 2 tumours which were initially diagnosed as Koos 1 were treated with SRS when detected by MRI. We also performed a probabilistic sensitivity analysis with 10,000 simulations to investigate sampling uncertainty concerning the parameters in the model. We did this for important variables: initial probabilities, growth rates, utilities in the monitoring strategy, and the utility of brainstem compression (Table 8.1). The percentile method was used to calculate 95% confidence intervals (CIs) from simulations. Simulation results are presented in cost-effectiveness planes and cost-effectiveness acceptability curves (CEACs). All analyses were conducted using TreeAge Pro 2015 (TreeAge Software, Inc.) and percentiles were calculated in Excel 2007 (Microsoft).

RESULTS

We assessed the cost-effectiveness of multiple monitoring strategies for the follow-up of VS. Omitting monitoring is least effective with on average 18.23 (95%CI 16.84 – 19.37) QALYs while lifelong annual monitoring is most effective with 18.66 (95%CI 17.42 – 19.65) QALYs per patient. Overlapping 95%CIs were found regarding the effectiveness of all 6 monitoring strategies (Table 8.2). Lifelong annual monitoring was the most expensive strategy with average costs of €9,429 (95%CI 9,197 – 9,643) per patient. Omitting monitoring is the least expensive strategy with average costs of €6,526 (95%CI 5,923 – 7,058) per patient, which are mainly treatment costs.

Lifelong annual monitoring had the highest NMB, €363,765 (95%CI 339,040 – 383,697), and therefore provides most value for money (i.e. the strategy gained most QALYs at a price that we are willing to pay as society). This strategy was followed by annual monitoring for the first 10 years with an NMB of €362,174 (95%CI 336,438 – 382,311). The strategy with the lowest NMB, representing the least cost-effective strategy, was no monitoring with an NMB of €358,168 (95%CI 330,371 – 380,908). Although this strategy was least expensive, it also gains the least QALYs. The savings in this strategy do not weigh up against the QALYs lost, hence the lower NMB of this strategy. The 95%CIs for the NMBs were largely overlapping for all strategies (Table 8.2).

Table 8.2 Outcomes.

Strategy	Costs (€)	Effects (QALYs)	NMB (€)
1. Lifelong annual monitoring	9,429 (9,197 – 9,643)	18.66 (17.42 – 19.65)	363,765 (339,040 – 383,697)
2. Annual monitoring for the first 10 years after diagnosis	8,684 (8,297 – 9,033)	18.54 (17.26 – 19.55)	362,174 (336,438 – 382,311)
3. Scans at 1-5, 7, 9, 12, 15 after diagnosis and subsequently every 5 years	8,585 (8,232 – 8,911)	18.52 (17.27 – 19.54)	361,788 (336,809 – 382,335)
4. Personalised monitoring strategy for small and large tumours	8,149 (7,708 – 8,552)	18.46 (17.15 – 19.49)	360,986 (335,032 – 381,638)
5. Scans at 1, 2 and 5 years after diagnosis	8,032 (7,588 – 8,439)	18.44 (17.12 – 19.47)	360,774 (334,483 – 381,507)
6. No monitoring	6,526 (5,923 – 7,058)	18.23 (16.84 – 19.37)	358,168 (330,371 – 380,908)

NMB, net monetary benefit; QALY, quality adjusted life year.

Using an alternative treatment scheme resulted in additional costs, as more patients received treatment. Treatment outcomes for Koos 2 and Koos 3 tumours were the same, therefore no differences in quality of life were expected in case of annual monitoring. However, in other monitoring strategies, treating growing Koos 2 tumours resulted in higher quality of life as brainstem compression is prevented (Table 8.3). Therefore, alternative treatment is cost-effective compared to treating only Koos 3 and 4 in these monitoring strategies.

In Figure 8.2 the incremental results of the probabilistic sensitivity analysis are shown. There is uncertainty in the effectiveness of all strategies, resulting in largely overlapping 95%CIs for QALYs and NMBs. The CEAC shows that all strategies have a relatively low probability to be most cost-effective due to large uncertainty in the results (Figure 8.3). At a threshold of €20,000 per QALY, lifelong annual monitoring has a 23% probability to be the most cost-effective strategy, which is higher than annual monitoring for the first 10 years (18%), scans at 1-5, 7, 9, 12, 15 years after diagnosis and subsequently every 5 years (16%), personalized monitoring (16%), scans at 1, 2 and 5 years after diagnosis (15%) and no monitoring (11%).

Table 8.3 Additional costs and effects of using an alternative treatment scheme. In this strategy, growing Koos 2 tumours are treated with SRS when detected. We calculated the additional costs and effects for each monitoring strategy, compared to the same monitoring strategy in the base case analysis.

Strategy	Additional costs* (€)	Additional effects* (QALYs)	Incremental NMB** (€)
1. Lifelong annual monitoring	199	0.00	-199
2. Annual monitoring for the first 10 years after diagnosis	174	0.01	26
3. Scans at 1-5, 7, 9, 12, 15 after diagnosis and subsequently every 5 years	114	0.03	486
4. Personalised monitoring strategy for small and large tumours	82	0.03	518
5. Scans at 1, 2 and 5 years after diagnosis	88	0.03	512
6. No monitoring	0	0.00	0

NMB, net monetary benefit; QALY, quality adjusted life year.
* Outcomes of this sensitivity analysis were compared to the base case analysis, for each monitoring strategy.
** A positive incremental NMB indicates that the strategy is cost-effective compared to the base case analysis.

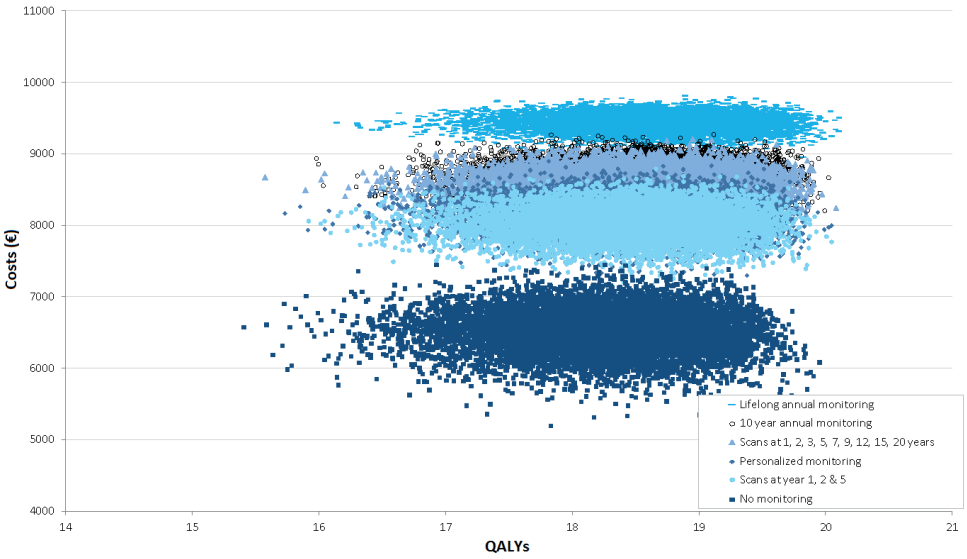


Figure 8.2 Outcomes of the probabilistic sensitivity analysis. This analysis quantifies the level of confidence of the model's conclusions. The four cost-effective monitoring strategies are displayed. Every dot represents the outcome of one analysis.

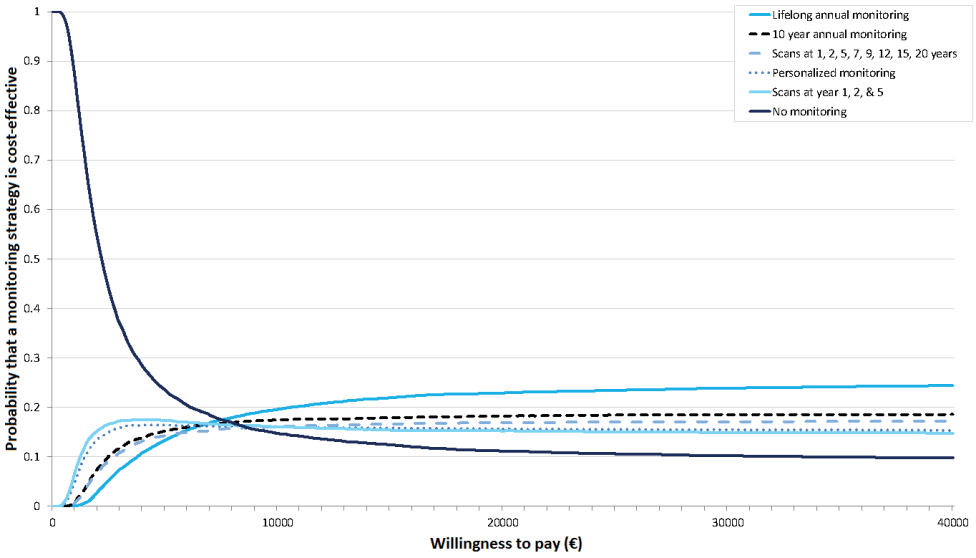


Figure 8.3 Cost-effectiveness acceptability curve of the four cost-effective strategies. This graph shows the probability that one of the strategies is most cost-effective for different willingness to pay values. The willingness to pay represents an estimate of what we might be prepared to pay for the health benefit.

DISCUSSION

Synopsis of key findings

We assessed the cost-effectiveness of multiple monitoring strategies for VS. Omitting monitoring is least effective with on average 18.23 (95%CI 16.84 – 19.37) QALYs while lifelong annual monitoring is most effective with 18.66 (95%CI 17.42 – 19.65) QALYs per patient. Corresponding costs were €6,526 (95%CI 5,923 – 7,058) and €9,429 (95%CI 9,197 – 9,643) per patient, respectively. Lifelong annual monitoring appeared to be most cost-effective with an NMB of €363,765 (95%CI 339,040 – 383,697). An alternative treatment scheme in which growing Koos 2 tumours are also treated was cost-effective when patients were not annually monitored. Sensitivity analysis shows that there is large uncertainty regarding the effectiveness of all strategies.

Strengths and limitations

To our knowledge, this is the first study to investigate cost-effectiveness of multiple monitoring strategies of VS. Others have studied cost-effectiveness of treatment strategies such as SRS or microsurgery.[19-21] However, the majority of patients with VS are nowadays observed through a monitoring strategy. In clinical practice, multiple monitoring strategies are used alongside each other, often lacking evidence of (cost-)effectiveness.[9] We therefore studied the cost-effectiveness of different monitoring strategies.

Some potential limitations should also be discussed. First, costs are based on Dutch healthcare prices, and may therefore slightly differ from other countries. The same applies to expert opinions, which can differ between hospitals and countries. We expect differences in exact costs and effects for other countries, but a similar trend. Given the detailed presentation of the model and its input parameters, those interested can assess the transferability of the results to their specific situation.

Second, we included VSs of all sizes into the monitoring strategy as this represents current practice in the Netherlands. Only 5% of tumours receive treatment directly following diagnosis. Monitoring for large tumours is more controversial, since the risk of brainstem compression is larger. When a less conservative management strategy is used, relatively smaller tumours will be included in the monitoring strategy with less severe consequences of undetected tumour growth. In this case, less intensive monitoring strategies would become more cost-effective.

Third, the construction of QALYs in this model required generic quality of life scores. We used the EQ-5D or SF-36 questionnaires, which are relatively insensitive for hearing problems compared to disease-specific questionnaires such as the PANQOL. However, there is currently no algorithm available that converts PANQOL outcomes to generic utility scores. Another generic questionnaire, the Health Utilities Index (HUI), does allow for the calculation of utility scores. Because it is more sensitive for hearing problems, the HUI seems more suitable to measure generic quality of life in patients with VS.[22] Unfortunately, we were unable to find utility scores measured by HUI for use in our model.

Last, transition from one health state to another in the monitoring strategy was based on the probability for a tumour to grow to the next Koos state. We chose these Koos states since they report clearly defined cut-off points, take tumour size and localization in relation to other structures into account, and have clearly defined consequences (i.e. recommended treatment).[23] We acknowledge that by using the Koos states as cut-off points, we were not able to detect growth within a Koos state. However, treatment options only change in case of progression to a next Koos state, therefore missing growth within a Koos state does not have consequences for treatment.

Implications for clinical practice

Currently, large differences in the management of VS are present. Multiple monitoring strategies are used alongside each other, without clear evidence of effectiveness.[9] In this analysis, we assessed the cost-effectiveness of several monitoring strategies. Looking at point estimates, lifelong annual monitoring seems most cost-effective. VSs are treated in time in this strategy, preventing serious consequences of brainstem compression.

However, the 95%CIs are largely overlapping with all other strategies. Based on the currently available evidence, the probability that lifelong annual monitoring is cost-effective is only 23%. This implies that if lifelong annual monitoring is implemented, the probability that this is the wrong decision is 77%. As

there is considerable uncertainty surrounding this decision, it might be better to wait for more evidence before we spend money on extensive monitoring strategies.
As shown in Figures 8.2 and 8.3, cost-effectiveness outcomes are very uncertain with probabilities for a strategy to be most cost-effective ranging from 11 to 23%. The uncertainty is mainly caused by uncertain effectiveness outcomes, due to the use of suboptimal effectiveness measures in literature and small sample sizes of study populations. Larger, high quality studies that investigate quality of life in VS patients assigned to a monitoring strategy using the HUI questionnaire are needed to achieve reliable effectiveness estimates. When research is initiated on this topic, a no monitoring strategy should be included. We used conservative assumptions for the no monitoring strategy, therefore we might be underestimating the cost-effectiveness of this strategy in this paper. Also, many patients remain in a monitoring strategy for life without needing treatment, therefore a no monitoring strategy could considerably lower the costs of monitoring.

In conclusion, due to the largely overlapping 95% CIs of all monitoring strategies for VS, it is unclear which monitoring strategy provides most value for money at this moment.

ACKNOWLEDGEMENTS

The authors thank Jeroen Verheul, Mark ter Laan and Jef Mulder for their valuable input and expert opinion.

CONFLICTS OF INTEREST

There are no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

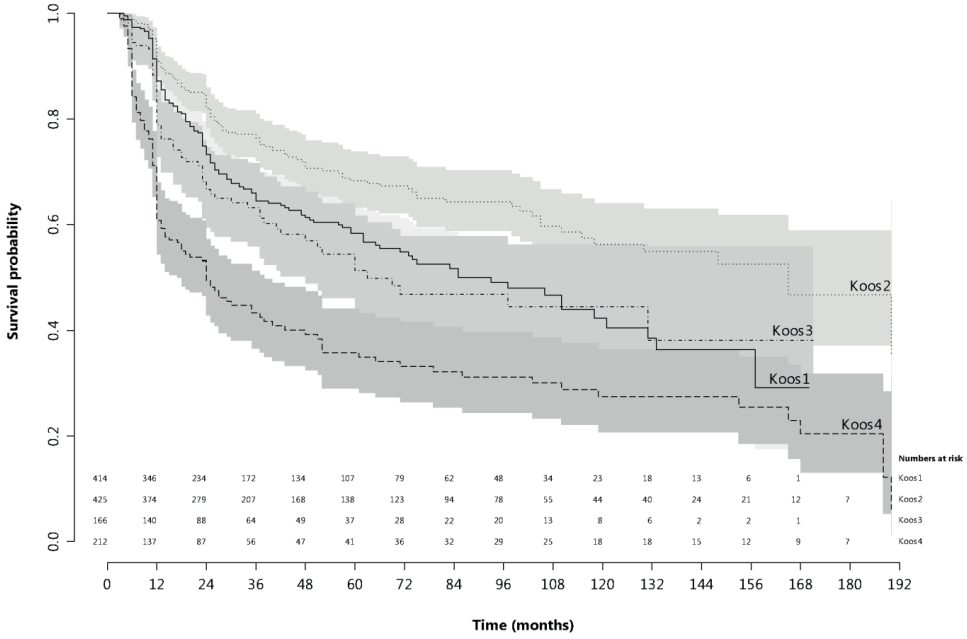
REFERENCES

1. Fortnum H, O'Neill C, Taylor R, Lenthall R, Nikolopoulos T, Lightfoot G, et al. The role of magnetic resonance imaging in the identification of suspected acoustic neuroma: a systematic review of clinical and cost effectiveness and natural history. Health Technol Assess. 2009;13(18):iii-iv, ix-xi, 1-154. Epub 2009/04/11. doi: 10.3310/hta13180.
2. Lin D, Hegarty JL, Fischbein NJ, Jackler RK. The prevalence of "incidental" acoustic neuroma. Arch Otolaryngol Head Neck Surg 2005;131(3):241-4. Epub 2005/03/23. doi: 10.1001/archotol.131.3.241.
3. Stangerup SE, Tos M, Thomsen J, Caye-Thomasen P. True incidence of vestibular schwannoma? Neurosurgery. 2010;67(5):1335-40. doi: 10.1227/NEU.0b013e3181f22660.
4. Carlson ML, Link MJ, Wanna GB, Driscoll CL. Management of sporadic vestibular schwannoma. Otolaryngol Clinics North Am. 2015;48(3):407-22. doi: 10.1016/j.otc.2015.02.003.
5. Conley GS, Hirsch BE. Stereotactic radiation treatment of vestibular schwannoma: indications, limitations, and outcomes. Curr Opin Otolaryngol Head Neck Surg. 2010;18(5):351-6. doi: 10.1097/MOO.0b013e32833c71a2.
6. Ferri GG, Modugno GC, Pirodda A, Fioravanti A, Calbucci F, Ceroni AR. Conservative management of vestibular schwannomas: an effective strategy. Laryngoscope. 2008;118(6):951-7. doi: 10.1097/MLG.0b013e31816a8955.
7. Hajioff D, Raut VV, Walsh RM, Bath AP, Bance ML, Guha A, et al. Conservative management of vestibular schwannomas: third review of a 10-year prospective study. Clin Otolaryngol. 2008;33(3):255-9. doi: 10.1111/j.1749-4486.2008.01705.x.
8. Pan P, Huang J, Morioka C, Hathout G, El-Saden SM. Cost analysis of vestibular schwannoma screening with contrast-enhanced magnetic resonance imaging in patients with asymmetrical hearing loss. J Laryngol Otol. 2016;130(1):21-4. doi: 10.1017/s0022215115002431.
9. Wolbers JG, Dallenga AH, van Linge A, Te West M, Kummer EE, Mendez Romero A, et al. Identifying at Diagnosis the Vestibular Schwannomas at Low Risk of Growth in a Long-term Retrospective Cohort. Clin Otolaryngol. 2016. doi: 10.1111/coa.12661.
10. Hentschel M, Rovers M, Markodimitraki L, Steens S, Kunst H. An international comparison of diagnostic and management strategies for vestibular schwannoma. Eur Arch Otorhinolaryngol. 2019;276(1):71-8. doi: 10.1007/s00405-018-5199-6.
11. Mindermann T, Schlegel I. Grading of vestibular schwannomas and corresponding tumor volumes: ramifications for radiosurgery. Acta Neurochir. 2013;155(1):71-4. doi: 10.1007/s00701-012-1553-4.
12. Stepanidis K, Kessel M, Caye-Thomasen P, Stangerup SE. Socio-demographic distribution of vestibular schwannomas in Denmark. Acta Otolaryngol. 2014;134(6):551-6. doi: 10.3109/00016489.2014.890293.
13. Zorginstituut Nederland. Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg. Accessed: 17-11-2015. Available from: https://www.zorginstituutnederland.nl/publicaties/publicatie/2016/02/29/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg.
14. Vemer P, Corro Ramos I, van Voorn GA, Al MJ, Feenstra TL. AdViSHE: A Validation-Assessment Tool of Health-Economic Models for Decision Makers and Model Users. Pharmacoeconomics. 2016;34(4):349-61. doi: 10.1007/s40273-015-0327-2.
15. Sughrue ME, Yang I, Aranda D, Rutkowski MJ, Fang S, Cheung SW, et al. Beyond audiofacial morbidity

after vestibular schwannoma surgery. J Neurosurg. 2011;114(2):367-74. doi: 10.3171/2009.10.jns091203.

16. Ara R, Brazier J. Deriving an algorithm to convert the eight mean SF-36 dimension scores into a mean EQ-5D preference-based score from published studies (where patient level data are not available). Value Health. 2008;11(7):1131-43. doi: 10.1111/j.1524-4733.2008.00352.x.
17. Ware JE, Jr. SF-36 health survey update. Spine. 2000;25(24):3130-9.
18. Godefroy WP, Kaptein AA, Vogel JJ, van der Mey AG. Conservative treatment of vestibular schwannoma: a follow-up study on clinical and quality-of-life outcome. Otol Neurotol. 2009;30(7):968-74. doi: 10.1097/MAO.0b013e3181b4e3c9.
19. Gait C, Frew EJ, Martin TP, Jowett S, Irving R. Conservative management, surgery and radiosurgery for treatment of vestibular schwannomas: a model-based approach to cost-effectiveness. Clin Otolaryngol. 2014;39(1):22-31. doi: 10.1111/coa.12205.
20. Morrison D. Management of patients with acoustic neuromas: a Markov decision analysis. Laryngoscope. 2010;120(4):783-90. doi: 10.1002/lary.20819.
21. Whitmore RG, Urban C, Church E, Ruckenstein M, Stein SC, Lee JY. Decision analysis of treatment options for vestibular schwannoma. J Neurosurg. 2011;114(2):400-13. doi: 10.3171/2010.3.jns091802.
22. Grutters JP, Joore MA, van der Horst F, Verschuure H, Dreschler WA, Anteunis LJ. Choosing between measures: comparison of EQ-5D, HUI2 and HUI3 in persons with hearing complaints. Qual Life Res. 2007;16(8):1439-49. doi: 10.1007/s11136-007-9237-x.
23. Erickson NJ, Schmalz PGR, Agee BS, Fort M, Walters BC, McGrew BM, et al. Koos Classification of Vestibular Schwannomas: A Reliability Study. Neurosurgery. 2018. doi: 10.1093/neuros/nyy409.
24. Centraal Bureau voor de Statistiek. Levensverwachting; geslacht en leeftijd, vanaf 1950. 2015 [cited 2016 26-1]. Available from: http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=37360ned&D1=0&D2=a&D3=1,11,21,31,41,51,55-99&D4=l&HDR=G1,T&STB=G2,G3&VW=T.
25. Klijn S, Verheul JB, Beute GN, Leenstra S, Mulder JJ, Kunst HP, et al. Gamma Knife radiosurgery for vestibular schwannomas: evaluation of tumor control and its predictors in a large patient cohort in The Netherlands. J Neurosurg. 2016;124(6):1619-26.
26. Godefroy WP, van der Mey AG, de Bruine FT, Hoekstra ER, Malessy MJ. Surgery for large vestibular schwannoma: residual tumor and outcome. Otol Neurotol. 2009;30(5):629-34. doi: 10.1097/MAO.0b013e3181a8651f.
27. Turel MK, Thakar S, Rajshekhar V. Quality of life following surgery for large and giant vestibular schwannomas: a prospective study. J Neurosurg. 2015;122(2):303-11. doi: 10.3171/2014.10.jns14534.
28. Godefroy WP, Hastan D, van der Mey AG. Translabyrinthine surgery for disabling vertigo in vestibular schwannoma patients. Clin Otolaryngol. 2007;32(3):167-72. doi: 10.1111/j.1365-2273.2007.01427.x.
29. Varughese JK, Wentzel-Larsen T, Pedersen PH, Mahesparan R, Lund-Johansen M. Gamma knife treatment of growing vestibular schwannoma in Norway: a prospective study. Int J Rad Oncol Biol Phys. 2012;84(2):e161-6. doi: 10.1016/j.ijrobp.2012.03.047.

SUPPLEMENTAL CONTENT



Supplemental Figure 8.1 Kaplan-Meier estimates of VS growth rates to the next Koos state, displayed per Koos state at diagnosis. We used a retrospective cohort (n=1217) of patients diagnosed with VS in the Radboudumc between 1990 and 2016. Patients with a unilateral VS, assigned to the monitoring strategy that had at least one follow-up MRI available were included. Transition among Koos states was defined by the probability of tumour growth to the next Koos state and the probability to have ≥2mm growth in Koos 4. Missing data were assumed to be missing at random. Missing data were imputed by multiple imputation using chained equations to create 25 imputed data sets. One randomly selected set was included in our model. We included the probability of tumour growth when sample sizes were ≥10 patients, thereafter growth was assumed not to occur.

CHAPTER 9

Development of a model to predict vestibular schwannoma growth: an opportunity to introduce new wait and scan strategies

Hentschel M.
Hannink G.
Steens S.
Mulder J.
Rovers M.
Kunst H.

Submitted.

ABSTRACT

Background

A large proportion of vestibular schwannomas (VSs) within a wait and scan (W&S) strategy remains stable in size and thus untreated. Follow-up, comprising repeated visits and imaging, contributes to high costs and burdening of patients. Preferably, we would identify patients with a high risk of future VS growth to tailor follow-up regimes. Therefore, we aimed to develop a prediction model to predict VS growth.

Methods

We used a cohort of patients diagnosed with VS in a university hospital between 1990 and 2016. Patients with a unilateral VS, admitted in a W&S strategy and at least one MRI available (following baseline MRI) were included. Twenty-two potential predictors were selected based on literature and interviews with experts. Data on demographics, symptoms, audiometry and MRI characteristics at time of diagnosis were collected from medical records. We used multiple imputation for missing data. A multivariable Cox regression model was used to select the variables using backward selection. VS growth of ≥2 mm was used as outcome and time to VS growth was expressed in months. Decision curve analyses (DCA) were performed to compare the model to the current strategy.

Findings

Of the 1217 analysed VS patients, 653 (53.7%) showed growth during follow-up. Median time to event (VS growth) was 13 months (range: 3-167) and median censoring time 44 months (range: 2-243). Balance complaints (HR 1.57 (95%CI: 1.31-1.88)) and tinnitus complaints in the affected ear (HR 1.36 (95%CI: 1.15-1.61)), Koos grade (Koos 1 is reference, Koos 2 HR 1.03 (95%CI: 0.80-1.31), Koos 3 HR 1.55 (95%CI: 1.16-2.06), Koos 4 HR 2.18 (95%CI: 1.60-2.96)), time since onset of symptoms (IQR HR 0.83 (95% CI: 0.77-0.88) and intrameatal diameter on MRI (IQR HR 1.67 (95%CI: 1.42-1.96)) were selected as significant predictors. Discrimination of the model (Harrell's C) was 0.69 (95%CI: 0.67-0.71) and calibration was good. DCA showed that the model has a higher net-benefit than the current strategy for calculated probabilities of VS growth of >12%, 15% and 21% for the first, second and third year, respectively.

Interpretation

Patients with balance and tinnitus complaints, a higher Koos grade, short duration of symptoms and a larger intrameatal diameter at time of diagnosis have a higher probability of future VS growth.

Funding

Health Care Efficiency Research grant by The Netherlands Organization for Health Research and development (ZonMw).

INTRODUCTION

Over the past years conservative management of unilateral vestibular schwannoma (VS) has gained popularity.[1] Currently, a 'wait and scan' (W&S) strategy is preferred in the majority of patients with a newly diagnosed VS.[2] The aim of a W&S strategy is to detect VS growth by means of repeated magnetic resonance imaging (MRI) examinations. In case of a large VS compressing surrounding tissues and/or detected growth during W&S, patients are usually referred for treatment, consisting of radiation therapy (i.e. stereotactic radiosurgery [e.g. Gamma Knife] or fractionated radiotherapy), or microsurgery. A large proportion of VSs observed within a W&S strategy remains stable in size and thus remains untreated during life.[3, 4] VSs are usually diagnosed in the sixth decade of life.[2, 5] W&S strategies are known to vary. A survey among otolaryngologists revealed several strategies, consisting of MRIs every 1-5 years, either continued until a specific age (75 or 80), for a specific period (4-21 years) or lifelong.[6] Thus, patients undergo a large number of MRIs during a lifetime. This contributes to the high costs associated with VS care and burdening of hospital visits for patients.[7] Preferably, we would select patients that need to be monitored carefully, because their VS has a high risk of future growth (and thus treatment), while others can be monitored less strictly or may even be omitted from further controls. This might improve (cost-) effectiveness of the W&S strategy, contribute to individualized patient care, and result in better informed patients with regard to the prognosis of their disease due to improved patient counselling. Therefore, the purpose of this study was to develop a clinical prediction model that can be used to predict VS growth for newly diagnosed patients assigned to a W&S strategy.

MATERIALS AND METHODS

We developed a multivariable prediction model to predict VS growth. Information on potential predictors and the outcome was retrospectively collected from patient records. The study protocol was published online (in Dutch, summary in English: https://www.zonmw.nl/nl/onderzoek-resultaten/doelmatigheidsonderzoek/programmas/project-detail/doelmatigheidsonderzoek/cost-effective-diagnostic-strategies-in-patients-with-asymmetrical-hearing-impairment-or-unilateral/verslagen/). The study was reported following the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement.[8]

Study population

In the Netherlands most patients with a newly diagnosed VS are referred to a specialized tertiary centre to determine further management. We consulted medical records of all patients that got assigned the diagnostic code 'cerebellopontine angle (CPA) lesion' and/or had undergone an MRI of the CPA in a tertiary hospital in the Netherlands (Radboud university medical center, Nijmegen) between 1990 and July 2016. We identified patients with a unilateral VS diagnosed by means of MRI. All patients initially assigned to a W&S strategy were included. The local W&S strategy prescribes MRIs at 1, 2, 3, 5, 7, 9, 12, and

15 years following diagnosis, then continuing every 5 years during the remaining lifetime of a patient. The W&S strategy could either be carried out in our own institution or in the referring clinic. To be able to study VS growth at least one follow-up MRI (either images or a report) had to be available. Thus, patients diagnosed with other modalities than MRI, those with bilateral VSs (i.e. neurofibromatosis), VSs that immediately had been treated, or without available follow-up, or CPA lesions other than VS were excluded.

Outcome

VS growth

Time-to-VS growth was defined as the number of months between the baseline MRI and the one on which VS growth was detected.

MRI examinations were assessed by one of the authors [MH] to determine whether growth had occurred. Each MRI was compared to the baseline MRI. Largest VS diameter was measured in two directions on axial images, i.e. parallel to the internal auditory canal (split in an intra- and extrameatal portion delineated by the petrous bone)[9] and largest extrameatal diameter parallel to the petrous bone. All measurements were rounded off to millimetres. Contrast enhanced T1-weighted images were preferably used to assess lesions. In case these were unavailable, T2-weighted images were used.

For intrameatal VSs, an increase in tumour diameter ≥2 mm parallel to the internal auditory canal was considered growth. For extrameatal VSs, growth was considered an increase ≥2 mm of the extrameatal portion in either direction.[9]

Whenever the W&S strategy was performed in another hospital and baseline or follow-up MRI images were unavailable, we evaluated growth based on the radiologists' reports. When the report stated that growth had occurred, we assumed this to be true.

Potential predictors

Twenty-two potential predictors were selected based on literature and interviews with three experts (otolaryngologists from our centre, working in the field of VS). Demographics (sex [male/female] and age), symptoms, pure-tone audiometry (PTA) and MRI findings at time of diagnosis were collected from the patients' medical records. Presence of complaints of hearing loss, tinnitus and aural pressure on the affected side were collected [yes/no]. The onset of hearing loss was classified [sudden/gradual]. Also, the presence of vertigo or balance complaints was collected [present/absent]. The time since onset of symptoms up to diagnosis was expressed in months [continuous].

Pure-tone audiometry

PTA data were retrieved from the clinical audiology database system AudiologicX (version 1.0.6, MarYor, the Netherlands). In our centre PTA is performed in a soundproof room according to standard

audiometric protocols. We collected hearing thresholds in dB hearing loss of octave frequencies 0.5, 1, 2, 4, and 8 kHz for air conduction (AC). Measurements on the affected side were used. Results of PTA performed within a range of 6 months prior and after diagnosis were included. In case a patient had multiple PTA examinations available, the one most proximate to the diagnostic MRI was selected.

Baseline MRI

Baseline MRI images were assessed to determine VS size as previously described [continuous], aspect [homogeneous/inhomogeneous], presence of cysts [yes/no] and Koos grading scale [grade 1-4], representing the lesion's size in relation to surrounding structures.[10]

Data analysis

Descriptive statistics were used to summarize the data. For 15 of the 22 potential predictors data were missing, ranging between 2.2% and 63.8% (Table 9.1). We assumed missing data to be missing at random (MAR). Imputation of missing values was performed using multiple imputation by chained equations, creating 25 imputation sets.[11]

Potential predictors were entered into a Cox regression model, taking into account the multiple imputed datasets. Akaike's information criterion was used as a selection criterion.[12] The probability of VS growth at a certain time point can be calculated by using the following formula: 1 - S(t), where $S(t) = S_0(t)^{\wedge exp(\beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n)}$.

In this formula, S(t) is the 'survival' of VS, i.e. the probability of no VS growth. $S_0(t)$ represents the baseline survival at time t and β_1, β_2 and β_n are the regression coefficients of the predictors x_1, x_2, and x_n, respectively, after having been pooled. Baseline survival is defined as the survival for the mean of all covariates in the model and can be transformed into a probability of future growth at the different time points for an individual patient.

For newly diagnosed VS patients assigned to a W&S strategy, predictions within the first five years following diagnosis are of interest to determine timing of the first follow-up MRI. Predictions at ten years are relevant for a patient's prognosis. Model performance was assessed on calibration using calibration plots for predictions at 1-5 and 10 years. The model's ability to discriminate between patients with successful or unsuccessful outcomes was estimated using Harrell's C.[13] Prediction models derived with multivariable regression analyses are known for overfitting. This results in too extreme predictions when the model is applied in new cases. Therefore, it was validated internally using bootstrapping techniques. Five hundred samples were drawn with replacement from the development sample. Bootstrapping techniques provide information on the performance of the model in comparable datasets and generate a shrinkage factor to adjust regression coefficients.[14] Thereafter, model performance was re-evaluated.

For development of multivariable prediction models, sample size is often based on the number of events per parameter estimated (EPP). This can be calculated by dividing the number of individuals with or without the outcome (whichever is lower) by the number of parameters to be estimated. We used 22 potential predictors that make up 24 parameters to be estimated (including multiple categories of the variable 'Koos grade'), amounting to an EPP of 23 (EPP = 564 'events [no VS growth]' divided by 24 parameters to be estimated). An EPP above 20 is considered to eliminate the estimated bias in regression coefficients and achieve reliable results.[15, 16]

A dynamic nomogram was created to easily calculate an individual's risk of VS growth. The nomogram is available via https://vs-model.shinyapps.io/predictVSgrowth, where more data can be entered and corresponding predictions on VS growth can be calculated.

TRIPOD recommends to evaluate net-benefit of prediction models.[16] Decision curve analysis (DCA) can help to summarize clinical usefulness of prediction models and support in decision making.[17, 18] In a DCA, net-benefit is plotted against threshold probability. In this study, net-benefit represents the proportion of true positives (detected VS growth) in absence of any false positives (i.e. specificity of 100%).[18] Threshold probability is defined as the minimum predicted risk of VS growth at which an otolaryngologist or patient would want the first follow-up MRI. A range of values for the threshold probability is displayed in order to represent a variation in preferences.[19, 20] Interviews with experts in the field of VS revealed a relevant range of risk threshold values of 10% (MRI in 10 patients to detect one case of VS growth and accept 9 false positives, i.e. unnecessary MRIs) to 30% (MRI in 10 patients to detect 3 cases of growth and accept 7 false positives). DCA was performed for the different time points (1-5 and 10 years). These can be used to compare the model to a 'scan all' (i.e. the current), and 'scan none' strategy and enable one to determine the threshold probability to initiate follow-up. Furthermore, we calculated the number of MRIs avoided for different threshold probabilities for each of the first five years. Data analysis was performed in R version 3.5.1 (R Foundation for Statistical Computing, Vienna, Austria) using packages 'rms' and 'rmda'.[21, 22]

Ethics statement

This study was performed with consent of the local medical ethics committee. The need for informed consent was waived, because of the retrospective nature and size of the study.

Role of the funding source

The funding source did not have any role in study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

RESULTS

Study population

We identified 1602 patients with an MRI-diagnosed unilateral VS. Three hundred and fourteen patients were excluded, because treatment was initiated at time of diagnosis, 239 (14.2%) and 75 (4.7%) were treated with microsurgery and radiation therapy, respectively. Another 14 (0.9%) were discharged from further controls due to patient preference or severe comorbidity. For the remaining 1274 (79.5%) patients a W&S strategy was initiated. Of these, 1217 had at least one follow-up MRI available and thus could be included for further study (Figure 9.1). Of the included VSs, 603 and 614 were located on the right and left side, respectively.

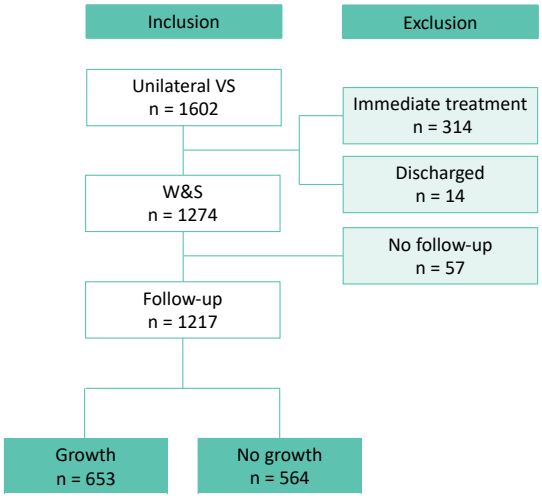


Figure 9.1 Flowchart displaying patient in- and exclusion.

Outcome

VS growth

MRI images were available for review for the majority of patients. Radiologists' reports were used to determine VS growth in 37.9% of the total number of examinations (n=3474). In 653 patients (53.7%) VS growth was detected at some point during W&S. Of these, 442 patients (67.7% of patients with a growing VS, or 36.3% of the total study population) also received treatment. Median time to VS growth was 13 months (range: 3-167), median censoring time (time to final follow-up for patients without VS growth) was 44 months (range: 2-243).

Predictors

Table 9.1 displays patient and VS characteristics at time of diagnosis. Median age of the included patients was 58.5 years (range 16.5-88.0) and 51.9% were male. Hearing loss was the most common complaint on the affected ear, followed by tinnitus. Baseline MRI images were available for review for 935 patients (76.8%), for the remaining patients we based VS presence on the radiologist's report. Other baseline MRI characteristics were registered as missing for the latter patients. Most patients presented with a Koos grade 1 (27.2%) or 2 (27.4%) VS. PTA results were available for 630 patients (51.8%).

Multivariable model

After backward selection, the following variables remained in the multivariable model: balance complaints and tinnitus complaints in the affected ear, Koos grading scale, duration of symptoms and the intrameatal diameter (Table 9.2). After multiplying the regression coefficients with the shrinkage factor (0.97) and updating the intercept, the model's performance was re-evaluated. The final model's discrimination yielded a Harrell's C of 0.69 (95% confidence interval (CI) 0.67 - 0.71), indicating good discrimination.[23] The model's calibration at the time points of interest was visualized with calibration plots and considered good for predictions at all time points (Supplemental Figure 9.1).

Patients with balance complaints (hazard ratio (HR) 1.57 (95% CI 1.31-1.88)) and tinnitus complaints in the affected ear (HR 1.36 (95% CI 1.15-1.61)), a higher Koos grade at time of diagnosis (HRs of 1.03 (95% CI 0.80-1.31), 1.55 (95% CI 1.16-2.06) and 2.18 (95% CI 1.60-2.96) for Koos grade 2, 3 and 4, respectively), short duration of symptoms (interquartile range (IQR) HR 0.83 (95% CI 0.77-0.88)), and a larger intrameatal diameter at time of diagnosis (IQR HR 1.67 (95% CI 1.42-1.96)) have a higher probability of future VS growth (Table 9.2).

Example

Using the proposed multivariable model for a patient whose complaints started 12 months ago, who has tinnitus but no balance problems, whose VS has an intrameatal diameter of 8 mm, and is classified as Koos grade 1, would have a probability of VS growth of 38% (95% CI 32-43%) two years following diagnosis. For five years following diagnosis, this probability increases to 55% (95% CI 48-61%) (Figure 9.2). A patient with the same characteristics, with the exception of having a Koos grade 4 VS rather than a Koos 1 grade at time of diagnosis would have a probability of future VS growth of 64% (95% CI 56-70%) and 82% (95% CI 75-87%) at 2, and 5 years following diagnosis, respectively (Figure 9.2). More variations can be entered online to calculate predictions at different time points (https://vs-model.shinyapps.io/predictVSgrowth).

Table 9.1 Characteristics of included patients with a unilateral VS obtained in a W&S strategy.

	Descriptive n (%)*					
At time of diagnosis	Total N=1217	Missing	VS growth n=653	Missing	No VS growth n=564	Missing
Male gender	632 (51.9)	-	334 (51.1)	-	298 (52.8)	-
Age (years, median (range))	58.5 (16.5-88.0)	-	57.7 (16.5-88.0)	-	59.2 (23.9-87.9)	-
Symptoms						
Hearing loss**	1105 (90.8)	-	598 (91.6)	-	507 (89.9)	-
Onset of hearing loss		776 (63.8)		418 (64.0)		358 (63.5)
- Sudden	104 (8.5)		57 (8.7)		47 (8.3)	
- Progressive	225 (18.5)		123 (18.8)		102 (18.1)	
Tinnitus**	778 (63.9)	-	439 (67.2)	-	339 (60.1)	-
Aural pressure**	200 (16.4)	-	114 (17.5)	-	86 (15.2)	-
Dizziness	481 (39.5)	-	282 (43.2)	-	199 (35.3)	-
- Balance complaints	306 (25.1)	27 (2.2)	196 (30.0)	13 (2.0)	110 (19.5)	14 (2.5)
- Vertigo	120 (9.8)		58 (8.9)		62 (11.0)	
Duration of symptoms (months, median (range))±	12 (0-983)	2 (0.2)	12 (0-627)	2 (0.3)	13 (0-983)	-
Koos grade		282 (23.2)		167 (25.6)		115 (20.4)
1	331 (27.2)		129 (19.8)		202 (35.8)	
2	334 (27.4)		187 (28.6)		147 (26.1)	
3	121 (9.9)		73 (11.2)		48 (8.5)	
4	149 (12.2)		97 (14.9)		52 (9.2)	
Median diameter (mm, median (range))						
Intrameatal (parallel to internal auditory canal)	8 (0†-16)	302 (24.8)	8 (0†-15)	176 (27.0)	6 (0†-16)	126 (22.3)
Extrameatal						
- Perpendicular to petrous bone (parallel to internal auditory canal)	4 (0-24)	293 (24.1)	5 (0-22)	175 (26.8)	2 (0-24)	118 (20.9)
- Parallel to petrous bone	7 (0-44)	297 (24.4)	8 (1-34)	174 (26.6)	6 (0-44)	123 (21.8)

Table 9.1 Continued

Aspect on MRI						
Inhomogeneous	241 (19.8)	444 (36.5)	153 (23.4)	254 (38.9)	88 (15.6)	190 (33.7)
Cystic	136 (11.2)	463 (38.0)	85 (13.0)	268 (41.0)	51 (9.0)	195 (34.6)
PTA (dB, median (range))						
- 250 Hz AC	25 (-5-110)	587 (48.2)	25 (-5-110)	335 (51.3)	20 (0-110)	252 (44.7)
- 500 Hz AC	25 (0-120)	587 (48.2)	30 (0-120)	335 (51.3)	25 (0-120)	252 (44.7)
- 1000 Hz AC	40 (0-120)	587 (48.2)	35 (0-120)	335 (51.3)	45 (0-120)	252 (44.7)
- 2000 Hz AC	55 (-10-120)	587 (48.2)	55 (0-120)	335 (51.3)	55 (-10-120)	252 (44.7)
- 4000 Hz AC	65 (0-120)	587 (48.2)	63 (0-120)	335 (51.3)	65 (5-120)	252 (44.7)
- 8000 Hz AC	75 (0-110)	588 (48.3)	70 (0-110)	335 (51.3)	75 (0-110)	253 (44.9)

VS, vestibular schwannoma; PTA, pure-tone audiometry; AC, air conduction;
* The number of patients and corresponding percentage is reported, unless stated otherwise in the first column
** Ipsilateral of VS
± Duration was set at 0 in case the complaints were absent
† Intrameatal size can be 0 for intracochlear and exclusively extrameatal VSs

Table 9.2 Predictors for VS growth. Using the baseline risk and regression coefficients, a patient's probability of VS growth can be calculated.

	Multivariable analysis	
Predictors	**Regression coefficient after shrinkage**	**HR (95% CI)**
Balance complaints	0.4360	1.57 (1.31-1.88)
Tinnitus	0.2998	1.36 (1.15-1.61)
Koos grade 1	reference	reference
Koos grade 2	0.0250	1.03 (0.80-1.31)
Koos grade 3	0.4240	1.55 (1.16-2.06)
Koos grade 4	0.7542	2.18 (1.60-2.96)
Time since onset of symptoms (months)	-0.0046	0.83 (0.77-0.88)*
Intrameatal diameter (mm)	0.1237	1.67 (1.42-1.95)*

* Interquartile range hazard ratio (interquartile range).
HR: hazard ratio.

Baseline survival is defined as the survival for the mean of all covariates in the model. Growth probability for a new patient can be calculated using the formula: 1 - S(t), where $S(t) = S_{base}$^exp(lp), and S_{base} is the baseline survival at the time point of interest, and lp is the centred linear predictor. The baseline survival for timepoints 1-5 and 10 years are: $S_{base12} = 0.7707905$, $S_{base24} = 0.6238292$, $S_{base36} = 0.5362379$, $S_{base48} = 0.4889781$, $S_{base60} = 0.4496488$, $S_{base120} = 0.3244630$] The linear predictor can be manually calculated as: *lp* = 0.4360386 * (Balance complaints - 0.2670501) + 0.024917 * (Koos grade 2 - 0.3393591) + 0.423974 * (Koos grade 3 - 0.1322925) + 0.7541666 * (Koos grade 4 - 0.1930978) + -0.0045785 * (Time since onset - 43.5937962) + 0.1237266 * (Intrameatal diameter - 7.3360723) + 0.2998087 * (Tinnitus - 0.6392769).

Decision curves

Figure 9.3 displays the net-benefit curves of the model for predictions at the different time points. The strategy with the highest net-benefit regarding the detection of VS growth at a specific threshold probability is clinically most useful. At risk thresholds > 12%, 15%, 21%, 23%, 25% and 35% for years 1-5, and 10 years, respectively, the developed model has a higher net-benefit compared to scanning all patients (Figure 9.3). Figure 9.4 displays the percentage of MRIs avoided for risk thresholds of 10%, 20% and 30% for each of the first five years.

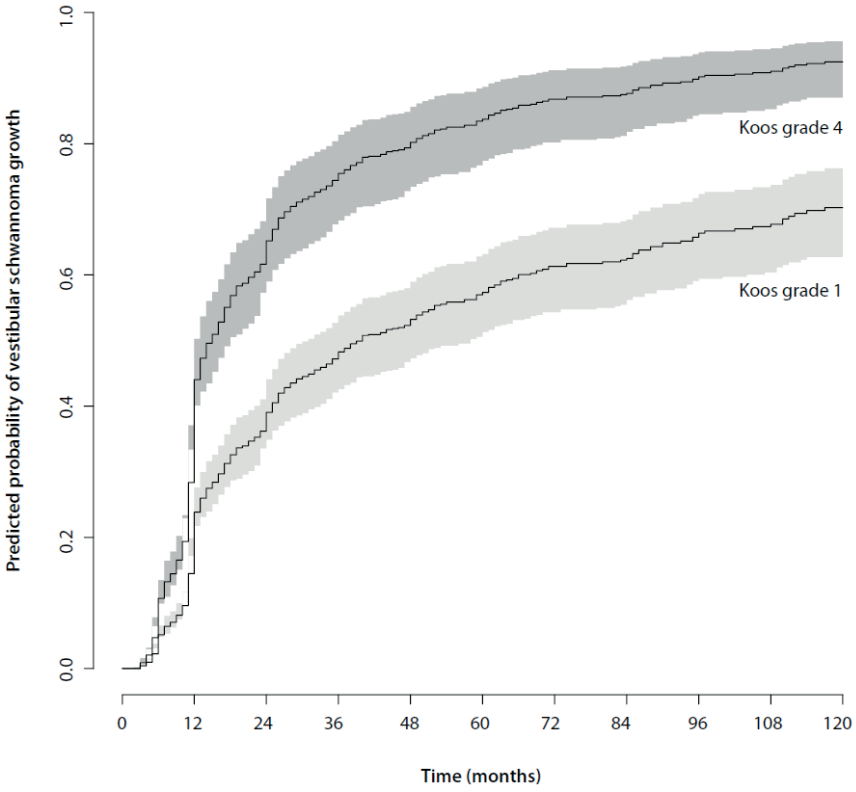


Figure 9.2 Predicted probabilities for a patient whose complaints started 12 months ago, who has tinnitus but no balance problems, whose VS has an intrameatal diameter of 8 mm, and is classified as Koos grade 1 (light grey) or Koos grade 4 (dark grey) at diagnosis.

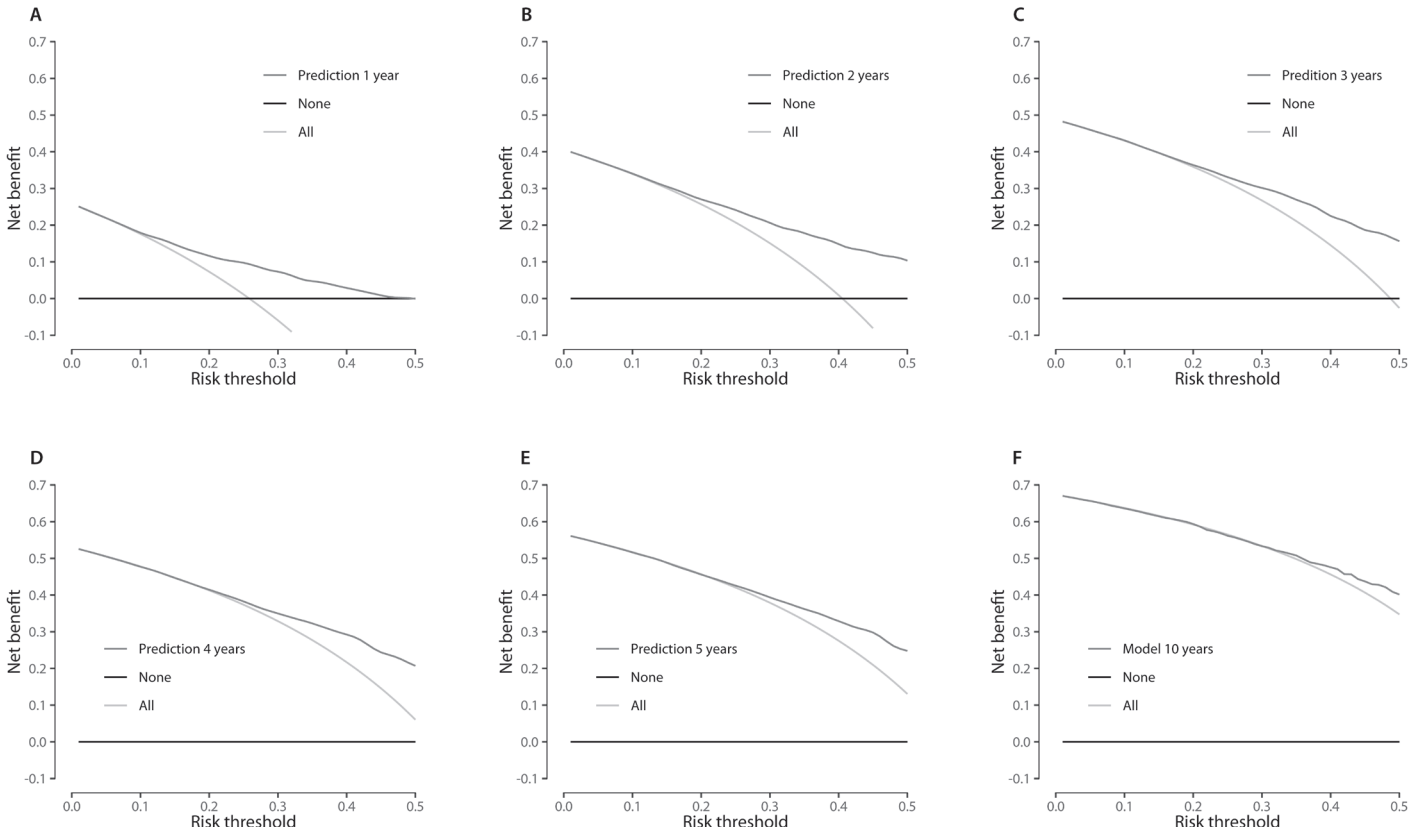


Figure 9.3 Net-benefit curves for time points 1-5 and 10 years.

The x-axis represents the risk threshold and the y-axis the net benefit.[18] Net-benefit represents the proportion of true positives (detected VS growth) in absence of any false positives (i.e. specificity of 100%).[18] The black line represents a strategy in which no MRIs are acquired, the net-benefit is 0. The grey line represents the current strategy, in which all patients have undergone an MRI. The coloured line represents the prediction model.

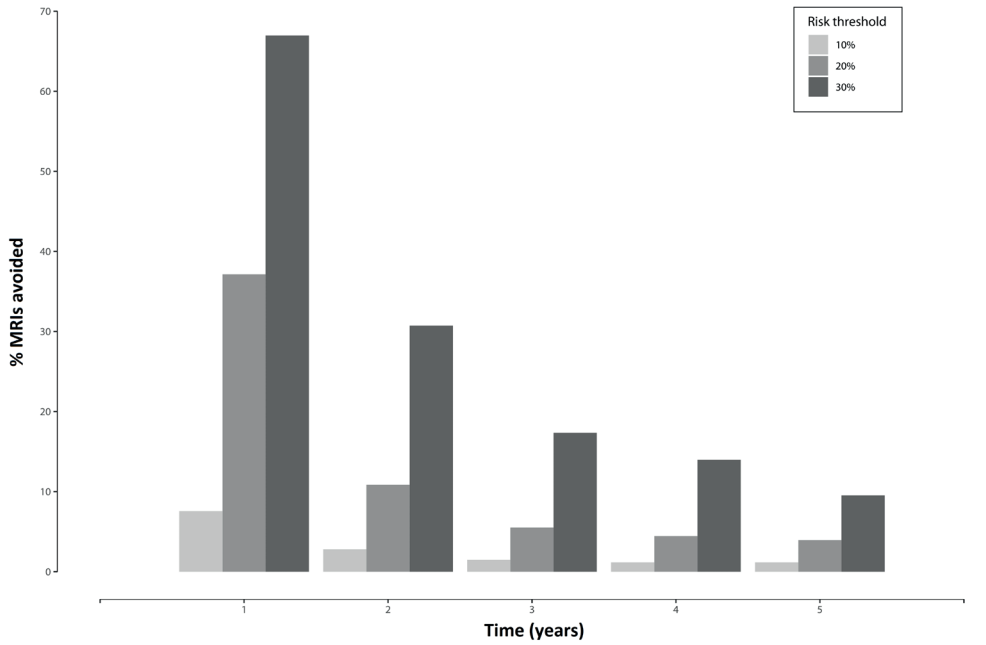


Figure 9.4 MRIs avoided for different risk thresholds and time points.

The x-axis displays different time points. The y-axis displays the proportion of MRIs avoided. Blue = risk threshold of 10%. Orange = risk threshold of 20%. Red = risk threshold of 30%.

DISCUSSION

We developed a multivariable time-to-event model predicting VS growth in newly diagnosed patients assigned to a W&S strategy. Our results show that patients with balance complaints and tinnitus complaints in the affected ear, a higher Koos grade, short duration of symptoms and a larger intrameatal diameter at time of diagnosis have a higher probability of VS growth. This prediction model may be helpful in the development of new W&S strategies and contributes to individualized care for patients diagnosed with VS. Individual patient data can be entered online (https://vs-model.shinyapps.io/predictVSgrowth) to calculate a patient's probability of VS growth.

Several authors have tried to identify features of VS that might predict future growth, growth rate and/or treatment. Most, however, looked at VS subgroups (e.g. VSs limited to the internal auditory canal[24, 25], or non-cystic VSs[26, 27]) used small sample sizes[3, 24, 28-33], used volume measurements to assess growth[30] or included growth rate in the first year of follow-up in their analyses[32]. Based on findings of studies on the topic, including the current study, we might state that age and sex are no strong predictors for VS

growth.[4, 24, 29-35] Multiple authors did find an association between VS growth or treatment and VS size at time of diagnosis[4, 24, 26, 36, 37], balance complaints[4, 26, 35], or extension in the CPA[26, 33, 38, 39].

The largest comparable study comprising 564 patients was performed by Hunter et al.[4] They evaluated risk factors for VS growth and found similar results, with larger initial VS diameter and disequilibrium complaints being identified as significant predictors (with increased HRs for both).[4] Age, sex, asymmetrical hearing loss and vertigo were not identified as significant predictors.[4] Tinnitus, however, was not selected in their model, whereas it was in the current study. This difference in findings could be explained by the fact that we, unlike Hunter et al., linked presence of symptoms to the affected ear.[4]

We have collected data of a large cohort of patients with unilateral VS. To our knowledge this is the largest study reporting on a complete cohort of VS patients in a W&S strategy with a relatively long duration of follow-up (mean 41 months). Predictors included in the model consist of presenting symptoms, baseline MRI parameters and PTA results, which can easily be obtained in every otolaryngology practice. By comparing each MRI to the baseline MRI (instead of the previous MRI), we were able to identify slow growing lesions.

Some potential limitations should also be discussed. First, PTA results were available for a slight majority of patients. From 2003 onwards, PTA results were digitally available. Thus, data from earlier days are missing at random (MAR). Second, MRI images were assessed by one person [MH]. Although inter- and intra-observer reliability is high for VS measurements, it is not 100%.[40] We, however, decided not to add a second reader as in daily clinical practice the outcome will also be assessed by a single person.

Third, we had to rely on radiologists' reports rather than MRI images in a large minority of cases (37.9%). Although assessment by radiologists from another institution might have been slightly different from our measurement method, we assumed that the presence of growth (yes/no) was properly assessed in these cases. In case of suspected growth, patients were usually referred to our clinic and MRI images could be assessed. For examinations of which we had both a report and measurements available, we used our measurements for analyses. We were able to compare our findings to the radiologist's report in these cases, and agreement was reached in 79%.

Regional otolaryngologists might have reported less often about stable VSs compared to growing VSs, since the latter are referred to our centre for further management. This might have resulted in an overestimation of VS growth. The proportion of patients with VS growth varies in literature, which is partially explained by abovementioned differences in study methods and follow-up. The proportion of VS growth found in our study (53.7%) was comparable to a study by Artz et al.[26] Kirchmann et al., who studied intracanalicular VSs (Koos grade 1) observed growth in 37% of patients, which is similar to our Koos 1 patients (38%).[25] In the study by Hunter et al. growth was detected in 40.8% of patients.[4] However, the fact that 36.3% of our patients was eventually treated is comparable to their data.[4]

Fourth, VS growth was defined as a ≥2 mm increase in diameter, while slice thickness was larger in the MRI examinations from the earliest study period. This might have resulted in an underreporting of VS growth in these earlier MRIs. Finally, we measured VS size in two directions, which might have led to missed growth in another direction.

The model was developed and internally validated in a Dutch population. External validation is necessary prior to its clinical use. Subsequently, the proposed multivariable model can be used in the consulting room to assess an individual patient's probability of having future VS growth. These findings can, next to patient counselling, also be used to establish a more individualized W&S strategy for patients. Increasing the interval between subsequent MRIs is relatively safe in selected patients, because potential growth can still be identified at a later time.

The data from this study enable further study of new W&S strategies. As mentioned previously, the range of threshold probabilities deemed relevant by specialists was 10-30%. When making predictions for 1-5 years following diagnosis, the model performs better than the current strategy with threshold probabilities within range preferred by the experts, i.e. the thresholds were > 12%, 15%, 21%, 23%, 25%, respectively. After 10 years the model performs similar to a 'scan all' strategy for the preferred threshold range of 10-30%, for a threshold of >35% the model has a higher net benefit than the current strategy.

Of all VSs that grew, less than 5% and 1% did so after 7 and 10 years, respectively. Given these data, we would at least suggest termination of follow-up after 10 years for non-growing VSs.

Future studies might reveal which changes in symptoms should prompt patients to visit their clinician. In case the model's performance could be further improved, it might even be possible to safely omit selected patients from further controls. It is difficult to assess the impact of missed VS growth and subsequent delayed treatment on clinical outcomes, especially since our data show that growth ≥2 mm does not necessarily lead to treatment. Stereotactic radiosurgery is usually not performed in VSs exceeding 3 cm, so delayed detection of growth beyond this size will result in a more invasive treatment, i.e. microsurgery.[41] So far, long term quality of life seems comparable for both treatment strategies, although results according to VS size are unknown.[42]

Abandoning all monitoring will initially result in the greatest cost reduction. However, based on current knowledge, long-term cost-effectiveness (including quality of life) and functional outcomes of the latter strategy are difficult to assess.

Further exploration of new W&S strategies, including their cost-effectiveness, is needed to reach an optimal W&S schedule.

CONCLUSION

Patients with balance and tinnitus complaints, a higher Koos grade, short duration of symptoms and a larger intrameatal diameter at time of diagnosis have a higher probability of future VS growth following diagnosis. Clinicians may use these variables to determine which recently diagnosed patients in a W&S strategy should be monitored more carefully.

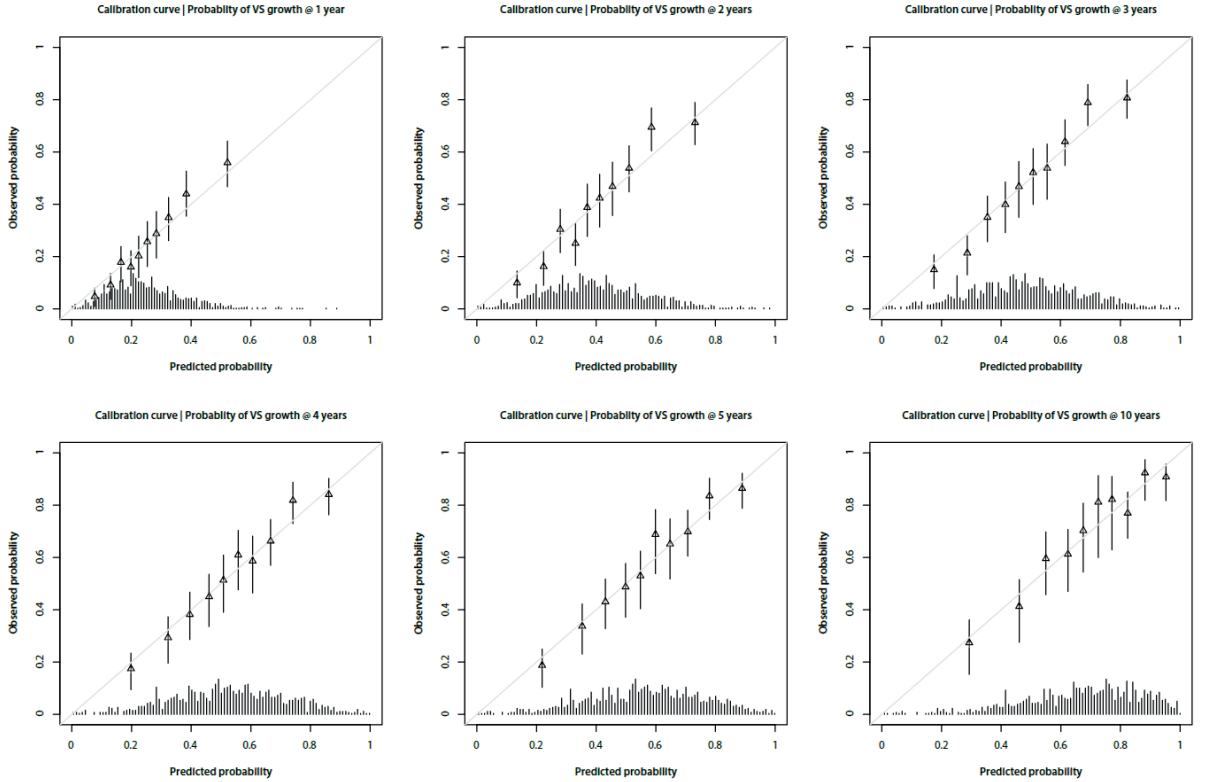
FUNDING

This study was funded by a Health Care Efficiency Research grant by The Netherlands Organization for Health Research and development (ZonMw).

ACKNOWLEDGEMENTS

Conflicts of interest: none. The authors want to thank Stan Wijn and Beau de Korte for their valuable input.

SUPPLEMENTARY RESOURCES



Supplemental figure 9.1 Calibration curves for vestibular schwannoma growth at 1-5 and 10 years.

The triangles represent the observed versus the predicted probabilities of VS growth. The vertical lines represent the 95% CI of the observed probabilities. The broom plot at the bottom shows the distribution of predicted probabilities.
VS = vestibular schwannoma, grey line = ideal.

REFERENCES

1. Carlson ML, Habermann EB, Wagie AE, Driscoll CL, Van Gompel JJ, Jacob JT, et al. The Changing Landscape of Vestibular Schwannoma Management in the United States--A Shift Toward Conservatism. Otolaryngol Head Neck Surg. 2015;153(3):440–6. doi: 10.1177/0194599815590105.
2. Kleijwegt M, Ho V, Visser O, Godefroy W, van der Mey A. Real Incidence of Vestibular Schwannoma? Estimations From a National Registry. Otol Neurotol. 2016;37(9):1411–7. doi: 10.1097/MAO.0000000000001169.
3. Hajioff D, Raut VV, Walsh RM, Bath AP, Bance ML, Guha A, et al. Conservative management of vestibular schwannomas: third review of a 10-year prospective study. Clin Otolaryngol. 2008;33(3):255-9. doi: 10.1111/j.1749-4486.2008.01705.x.
4. Hunter JB, Francis DO, O'Connell BP, Kabagambe EK, Bennett ML, Wanna GB, et al. Single Institutional Experience With Observing 564 Vestibular Schwannomas: Factors Associated With Tumor Growth. Otol Neurotol. 2016;37(10):1630–6. doi: 10.1097/MAO.0000000000001219.
5. Kshettry VR, Hsieh JK, Ostrom QT, Kruchko C, Barnholtz-Sloan JS. Incidence of vestibular schwannomas in the United States. J Neurooncol. 2015;124(2):223-8. doi: 10.1007/s11060-015-1827-9.
6. Hentschel M, Rovers M, Markodimitraki L, Steens S, Kunst H. An international comparison of diagnostic and management strategies for vestibular schwannoma. Eur Arch Otorhinolaryngol. 2018;276(1):71–8.
7. Scholte M, Hentschel MA, Hannink G, Kunst HPM, Steens SC, Rovers MM, et al. In search of the most cost-effective monitoring strategy for vestibular schwannoma: a decision analytical modelling study. Clin Otolaryngol. 2019;44(4):525–33.
8. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. J Clin Epidemiol. 2015;68(2):134-43. doi: 10.1016/j.jclinepi.2014.11.010.
9. Kanzaki J, Tos M, Sanna M, Moffat DA, Monsell EM, Berliner KI. New and modified reporting systems from the consensus meeting on systems for reporting results in vestibular schwannoma. Otol Neurotol. 2003;24(4):642–8. Epub 2003/07/10. PubMed PMID: 12851559.
10. Koos WT, Day JD, Matula C, Levy DI. Neurotopographic considerations in the microsurgical treatment of small acoustic neurinomas. J Neurosurg. 1998;88(3):506–12. doi: 10.3171/jns.1998.88.3.0506.
11. van Buuren S, Groothuis-Oudshoorn, K. mice: Multivariate Imputation by Chained Equations in R. J Stat Softw. 2011;45(3):1-67. doi: 10.18637/jss.v045.i03.
12. Wood AM, White IR, Royston P. How should variable selection be performed with multiply imputed data? Stat Med. 2008;27(17):3227–46. doi: 10.1002/sim.3177.
13. Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. Stat Med. 2004;23(13):2109–23. doi: 10.1002/sim.1802.
14. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J. 2014;35(29):1925–31. doi: 10.1093/eurheartj/ehu207.
15. Ogundimu EO, Altman DG, Collins GS. Adequate sample size for developing prediction models is

not simply related to events per variable. J Clin Epidemiol. 2016;76:175–82.

16. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015;162(1):W1–73. Epub 2015/01/07. doi: 10.7326/M14-0698.
17. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. Lancet Oncol. 2015;16(4):e173–80.
18. Van Calster B, Wynants L, Verbeek JFM, Verbakel JY, Christodoulou E, Vickers AJ, et al. Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators. Eur Urol. 2018;74(6):796–804. Epub 2018/09/23. doi: 10.1016/j.eururo.2018.08.038.
19. Capogrosso P, Vickers AJ. A Systematic Review of the Literature Demonstrates Some Errors in the Use of Decision Curve Analysis but Generally Correct Interpretation of Findings. Med Decis Making. 2019:272989X19832881.
20. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. BMJ. 2016;352:i6.
21. Harrell, F. E. rms: Regression Modeling Strategies. R package version 5.1-2. 2018. Available from: https://CRAN.R-project.org/package=rms.
22. Brown, M. rmda: Risk Model Decision Analysis. R package version 1.6. 2018. Available from: https://CRAN.R-project.org/package=rmda.
23. Steyerberg EW. Clinical prediction models. A practical approach to development, validation, and updating. New York: Springer Science+Business Media; 2009.
24. Younes E, Montava M, Bachelard-Serra M, Jaloux L, Salburgo F, Lavieille JP. Intracanalicular Vestibular Schwannomas: Initial Clinical Manifestation, Imaging Classification, and Risk Stratification for Management Proposal. Otol Neurotol. 2017;38(9):1345–50. doi: 10.1097/MAO.0000000000001538.
25. Kirchmann M, Karnov K, Hansen S, Dethloff T, Stangerup SE, Caye-Thomasen P. Ten-Year Follow-up on Tumor Growth and Hearing in Patients Observed With an Intracanalicular Vestibular Schwannoma. 2017;80(1):49–56. Epub 2016/08/30. doi: 10.1227/NEU.0000000000001414. PubMed PMID: 27571523.
26. Artz JC, Timmer FC, Mulder JJ, Cremers CW, Graamans K. Predictors of future growth of sporadic vestibular schwannomas obtained by history and radiologic assessment of the tumor. Eur Arch Otorhinolaryngol. 2009;266(5):641–6. doi: 10.1007/s00405-008-0791-9.
27. Timmer FC, Artz JC, Beynon AJ, Donders RT, Mulder JJ, Cremers CW, et al. Prediction of vestibular schwannoma growth: a novel rule based on clinical symptomatology. Ann Otol Rhinol Laryngol. 2011;120(12):807–13. doi: 10.1177/000348941112001206.
28. Godefroy WP, Kaptein AA, Vogel JJ, van der Mey AG. Conservative treatment of vestibular schwannoma: a follow-up study on clinical and quality-of-life outcome. Otol Neurotol. 2009;30(7):968–74. doi: 10.1097/MAO.0b013e3181b4e3c9.
29. Jethanamest D, Rivera AM, Ji H, Chokkalingam V, Telischi FF, Angeli SI. Conservative management of vestibular schwannoma: Predictors of growth and hearing. The Laryngoscope. 2015;125(9):2163–8. doi: 10.1002/lary.25159.
30. van de Langenberg R, de Bondt BJ, Nelemans PJ, Dohmen AJ, Baumert BG, Stokroos RJ. Predictors of volumetric growth and auditory deterioration in vestibular schwannomas followed in a wait and scan policy. Otol Neurotol. 2011;32(2):338–44. doi: 10.1097/MAO.0b013e3182040d9f.

31. Nutik SL, Babb MJ. Determinants of tumor size and growth in vestibular schwannomas. Journal of neurosurgery. 2001;94(6):922–6. doi: 10.3171/jns.2001.94.6.0922. PubMed PMID: 11409521.
32. Tschudi DC, Linder TE, Fisch U. Conservative management of unilateral acoustic neuromas. Am J Otol. 2000;21(5):722–8.
33. Hughes M, Skilbeck C, Saeed S, Bradford R. Expectant management of vestibular schwannoma: a retrospective multivariate analysis of tumor growth and outcome. Skull Base. 2011;21(5):295–302. doi: 10.1055/s-0031-1284219.
34. Stangerup SE, Caye-Thomasen P, Tos M, Thomsen J. The natural history of vestibular schwannoma. Otol Neurotol. 2006;27(4):547-52. doi: 10.1097/01.mao.0000217356.73463.e7.
35. Malhotra PS, Sharma P, Fishman MA, Grumbine FL, Tholey R, Dam VQ, et al. Clinical, radiographic, and audiometric predictors in conservative management of vestibular schwannoma. Otol Neurotol. 2009;30(4):507-14. doi: 10.1097/MAO.0b013e31819d3465.
36. Suryanarayanan R, Ramsden RT, Saeed SR, Aggarwal R, King AT, Rutherford SA, et al. Vestibular schwannoma: role of conservative management. J Laryngol Otol. 2010;124(3):251-7. doi: 10.1017/S0022215109992362.
37. Solares CA, Panizza B. Vestibular schwannoma: an understanding of growth should influence management decisions. Otol Neurotol. 2008;29(6):829-34. doi: 10.1097/MAO.0b013e318180a4c4.
38. Wolbers JG, Dallenga AH, van Linge A, Te West M, Kummer EE, Mendez Romero A, et al. Identifying at diagnosis the vestibular schwannomas at low risk of growth in a long-term retrospective cohort. Clin Otolaryngol. 2016;41(6):788–92. doi: 10.1111/coa.12661.
39. Ferri GG, Modugno GC, Pirodda A, Fioravanti A, Calbucci F, Ceroni AR. Conservative management of vestibular schwannomas: an effective strategy. The Laryngoscope. 2008;118(6):951-7. doi: 10.1097/MLG.0b013e31816a8955.
40. Coelho DH, Tang Y, Suddarth B, Mamdani M. MRI surveillance of vestibular schwannomas without contrast enhancement: Clinical and economic evaluation. 2018;128(1):202-9. doi: 10.1002/lary.26589.
41. Pai I, Bowman J, Thomas N, Kitchen N, Strong A, Obholzer R, et al. Management of large and giant vestibular schwannomas. Skull Base. 2011;21(6):379–84.
42. Papatsoutsos E, Spielmann PM. Self-Evaluated Quality of Life and Functional Outcomes After Microsurgery, Stereotactic Radiation or Observation-Only for Vestibular Schwannoma of the Adult Patient: A Systematic Review. Otol Neurotol. 2018;39(2):232–41. Epub 2018/01/10. doi: 10.1097/MAO.0000000000001664.

PART IV

General discussion and summaries

CHAPTER 10

General discussion

The overall aims of this thesis were to study 1) the (cost-)effectiveness of various diagnostic strategies in patients with asymmetrical audiovestibular dysfunction, i.e. suspected of having a cerebellopontine angle (CPA) lesion such as a vestibular schwannoma (VS), the most common lesion and 2) the (cost-) effectiveness of various 'wait and scan' (W&S) strategies in patients diagnosed with VS. In this discussion, I will summarise the main findings and discuss them from a more general perspective. Furthermore, recommendations for future guidelines for VS will be discussed.

DIAGNOSIS OF VESTIBULAR SCHWANNOMA

We modelled the potential savings in diagnosing VS by comparing the current diagnostic strategy to hypothetical new strategies. The results showed that the potential savings resulting from a perfect diagnostic strategy, i.e. a strategy where only patients with VS are referred for a magnetic resonance imaging (MRI) examination, amounted to 3.2 million euros per year in the Netherlands. Our international comparison of diagnostic strategies showed great variation, both within and between countries. This finding seems to be the result of a lack of evidence on reliable screening instruments for selecting which patients should undergo an MRI examination. At this time, no uniform guideline on diagnosing VS is available.

Based on our results and prior evidence, it seems possible to more selectively refer patients with asymmetrical audiovestibular dysfunction for MRI examinations. In our systematic review and meta-analysis, we pooled the results of five current screening strategies. The data were based on moderate to low quality studies. The highest diagnostic accuracy was shown by the pure-tone audiometry protocol that was proposed by the American Academy of Otolaryngology-Head and Neck Surgery (AAO-HNS), which recommends an MRI examination for patients with an average asymmetry of ≥15 dB at frequencies 0.5-3 kHz.[1, 2] Our own newly developed diagnostic model showed that gender, sudden onset of hearing loss, gradual onset of hearing loss, unilateral tinnitus, complaints of unilateral aural fullness, instability, headache, facial numbness, facial nerve dysfunction during physical examination, and asymmetry in bone conduction at 1 kHz and 4 kHz are important diagnostic features. Moreover, it seems possible to reduce MRI sequences in regard to diagnosing CPA lesions since high accuracy was found for a diagnostic imaging strategy with T2-weighted MRI only (compared to a strategy with contrast enhanced T1-weighted MRI). Box 10.1 provides an overview of our key findings regarding the (cost-)effectiveness of diagnosing VS.

Box 10.1 Key findings regarding the (cost-)effectiveness of diagnosing vestibular schwannoma.

- There is substantial room for improvement in diagnosing VS.
- Of the existing strategies that were tested in our study population, the diagnostic accuracy of the Seattle protocol was highest.
- Additionally, gender, sudden onset of hearing loss, gradual onset of hearing loss, unilateral tinnitus, complaints of unilateral aural fullness, instability, headache, facial numbness, facial nerve dysfunction during physical examination, and asymmetry in bone conduction at 1 kHz and 4 kHz seem important diagnostic features to identify CPA lesions.
- High diagnostic accuracy was found for T2-weighted MRI compared to contrast enhanced T1-weighted MRI for diagnosing CPA lesions.

Both current and new strategies will result in VSs being missed. For the AAO-HNS protocol, these misses would add up to approximately 30 undiagnosed VSs per year in the Netherlands compared to the current strategy (which leads to approximately 330 VS diagnoses each year).[3] The amount of VSs that stay undetected due to the diagnostic model depends on the applied threshold for MRI referral (at the threshold with the highest sensitivity, approximately 20 VSs would stay undiagnosed in the Netherlands per year compared to the current strategy). Also, lesions ≤2 mm could be missed on T2-weighted MRI. Thus, if we introduce a new diagnostic strategy in clinical practice, it will undoubtedly lead to undiagnosed patients with VS. This population includes patients with a large or growing VS that with the current scheme would receive treatment. The impact on functional outcomes, the quality of life and costs for these patients remain uncertain.

The Dutch Health Care Institute (Zorginstituut Nederland) uses thresholds for the willingness to pay per Quality Adjusted Life Year (QALY, an extra life year in perfect health) gained according to a disease's impact. A threshold of 50,000 euros per QALY is used for diseases with a medium burden.[4] It is difficult to state whether current and/or new strategies result in the necessary QALYs to justify their costs. There is a lack of longitudinal information on QALYs for the different Koos grades (a VS staging system based on the VS size and its relation to other structures)[5] and management strategies. According to our research, the current strategy does not seem to be the most cost-effective, as compared to some alternatives. If we want to try to improve cost-effectiveness, we must introduce new screening methods despite the lack of evidence on outcomes.

MANAGEMENT OF VESTIBULAR SCHWANNOMA

Our international comparison revealed great variation in management strategies and their indications. Again, this finding seems to be the result of a lack of evidence. There is no common guideline available for the management of VSs.

In the Netherlands, an increasing number of patients is assigned to a W&S strategy. In our study population (n=1602), 239 and 75 patients (19.6%) were treated with microsurgery and radiation therapy following diagnosis respectively. Fourteen patients (0.9%) were excluded from further controls due to severe comorbidity or patient preference. The majority of patients (n= 1274, 79.5%) was assigned to a W&S strategy. Currently a 'one-size-fits-all' approach is used in most clinics. Only a minority of patients assigned to a W&S strategy is eventually treated. During follow-up (median 34 months, range 2-242), growth was detected in 653 patients (53.7% of the 1217 patients) from our study population, 39.0% and 59.1% of which had intrameatal (Koos grade 1) and extrameatal (Koos grade 2 to 4) VSs at presentation respectively. Of all 653 patients with VS growth, 442 (67.7%) were treated. The estimated costs of the current W&S strategy in a patient's lifetime are about 8,500 euros per patient.[6] With increasing life expectancy, these costs will further increase.[7] A 'perfect W&S strategy' (in which an MRI examination is acquired at the time there is an indication for treatment, i.e. significant VS growth is observed) would result in cost savings of 1470 euros per patient, or approximately 0.5 million euros per year. We assessed the cost-effectiveness of multiple monitoring strategies for VS. Sensitivity analysis showed that there is large uncertainty regarding the effectiveness of all tested strategies, so based on current evidence it is difficult to state which strategy is preferred.

A multivariable model with clinical parameters that can be used at the time of diagnosis to predict VS growth showed fair predictive abilities. This model enables the development of tailored W&S strategies and can be used for patient counselling. We showed that the quality of life did not seem to differ between patients with a different Koos grade in a W&S strategy. Box 10.2 summarizes key findings regarding the management of VS. We propose to introduce a restricted W&S strategy in which patients with a low and high risk of future VS growth are categorized according to the margin for growth and the risk of growth (e.g. based on our prediction model). These risk categories can be used to personalize the W&S strategy or might in the future even be used as an indication for early treatment. New and less strict W&S strategies will, however, lead to patients with the growth of their VS going unnoticed. The consequences of missing a growing VS are uncertain. Thus, close monitoring of the effects on health outcomes by means of a national registry is warranted.

Box 10.2 Key findings on management of vestibular schwannoma.

- The quality of life did not seem to differ between patients with different Koos grades in a W&S strategy.
- A prediction model showed that tinnitus complaints in the affected ear, balance complaints, Koos grade, duration of symptoms and intrameatal diameter are significant predictors for future VS growth.

RECOMMENDATIONS FOR A GUIDELINE

A(n) (inter-)national guideline may reduce existing differences in clinical practice guidelines and costs.[8] Clinicians and policy makers may consider the available results, including the results of this thesis, and develop such an international guideline. Below, we will discuss our considerations regarding several potential recommendations.

Diagnosis

First, we recommend the introduction of a new diagnostic strategy in clinical practice followed by close monitoring of its effects on patient outcomes. Several options are available, and each has its own merits and drawbacks.

In our meta-analysis, the pure-tone audiometry protocol that was proposed by the AAO-HNS showed the highest diagnostic accuracy, with a pooled sensitivity of 90.9% (95% confidence interval (CI): 84.2-94.9) and a specificity of 57.5% (95% CI: 49.4-65.2). Sensitivity and specificity of our diagnostic model depend on the threshold applied. For the range of threshold values considered relevant (0.018-0.05), sensitivity and specificity vary from 22-99% and 4-94%, respectively. The highest combination of sensitivity and specificity was reached for a threshold of 0.038, sensitivity and specificity reached 50.7% and 80.6% respectively, which is inferior to diagnostic accuracy of the AAO-HNS criteria. However, testing the AAO-HNS criteria on our own screening population (averaging 2 and 4 kHz measurements to obtain 3 kHz thresholds) resulted in a lower sensitivity and specificity, i.e. 62.7% and 52.2% respectively. In our meta-analysis, the pooled sensitivity and specificity of the Seattle protocol, where an MRI examination is recommended for an average asymmetry of ≥15 dB at frequencies 1-8 kHz, were 89.2% (95% CI: 79.9-94.5) and 43.8% (95% CI: 25.8-63.6) respectively. Testing the Seattle protocol on our own screening population yielded higher results, with a sensitivity and specificity of 73.5% and 50.3% respectively. Based only on internal validation, the newly developed diagnostic model may achieve a higher diagnostic accuracy. For example, at a threshold of 0.03, sensitivity of the diagnostic model reaches 76.8%, thus correctly identifying the most VS patients. With a specificity of 49.3%, it would achieve a moderate reduction in MRI examinations of 51% when applying a threshold of 0.03. Potential savings could add up to about 1.4 million euros per year on MRIs only in the Netherlands. Table 10.1 provides an overview of the diagnostic strategies discussed above. Our diagnostic model has, however, not yet been externally validated. We recommend doing this prior to its implementation in clinical practice. For this future study, one of the study populations included in our meta-analysis could be used. The option of omitting all diagnostic MRI examinations for patients with asymmetrical audiovestibular dysfunction should also be discussed. It would initially result in the greatest cost reduction (4.5 million euros per year), but again the potential consequences of such a strategy in terms of the quality of life and functional outcomes remain unclear. Since more VSs remain undiagnosed, the uncertainty is even greater compared to the aforementioned options. All stakeholders, including patients, should be involved in the discussion about what strategy should be adopted in clinical practice.

Table 10.1 Overview of the diagnostic accuracy of non-imaging strategies to diagnose vestibular schwannoma.

		Diagnostic accuracy	External validation
Diagnostic method		Pooled	In our study population
AAO-HNS*	Sensitivity	90.9%‡	62.7%
	Specificity	57.5%‡	52.2%
Seattle†	Sensitivity	89.2%‡	73.5%
	Specificity	43.8%‡	50.3%
		Reported	In other study population
Diagnostic model	Sensitivity	76.8%§	?
	Specificity	49.3%§	?

AAO-HNS: American Academy of Otolaryngology-Head and Neck Surgery
* Average asymmetry of ≥15 dB at frequencies 0.5-3 kHz
† Average asymmetry of ≥15 dB at frequencies 1-8 kHz
‡ Pooled results (chapter 2)
§ Results after internal validation and applying a threshold of 0.03. Diagnostic accuracy varies depending on the applied threshold values (chapter 6)

Of the existing strategies, we prefer the Seattle protocol because it has the highest diagnostic accuracy in our population and has already been studied in several other cohorts.[1, 2, 9] However, the sensitivity level of the Seattle protocol decreased substantially after testing it on our population.

In order to externally validate our diagnostic model, we need data on the symptoms and pure-tone audiometry of a new study population comprising patients with asymmetrical audiovestibular dysfunction. Availability of such data might enable a retesting of the Seattle protocol since it is also based on pure-tone audiometry results. In the Netherlands, the introduction of either the Seattle protocol or the diagnostic model could save at least 1.4 million euros per year. We therefore propose the use of a study population to validate both the Seattle protocol and our diagnostic model. The study populations included in our meta-analysis could be used.

Management

Our developed model showed that tinnitus complaints in the affected ear, balance complaints, the Koos grade, the duration of symptoms and the intrameatal diameter are significant predictors of future VS growth. The model is not yet externally validated but seems promising in distinguishing between patients with a low and high risk of future growth. We propose the use of a multivariable model to determine the interval between diagnosis and the first follow-up. Since the model's discrimination is fair, we do not recommend to completely discharge low risk patients based on the model's predictions. The data from this study enable further study of new W&S strategies. Of all the VSs that grew, less than 5% and 1% did so after 7 and 10 years respectively. Given these data, we suggest the termination of follow-ups after 10 years of no growth. Moreover, the interval between subsequent MRI examinations could be altered using the probability of VS growth that is calculated by the model. Following external validation, we suggest to use calculated probabilities to determine further strategies. Since stereotactic radiosurgery can be used for VSs with a diameter of up to 3 cm,[10] there is more margin for growth for VSs that are small at time of diagnosis. Growth of these smaller VSs may stay undetected for a relatively longer period without excluding the latter treatment option.

For all potential new W&S strategies, an issue that must be taken into account is that the consequences of delayed and failed detection of VS growth, incidents which are inevitable, are difficult to assess. Abandoning all monitoring procedures will initially result in the greatest cost reduction. However, based on current knowledge, assessment of long-term cost-effectiveness (including the quality of life) and functional outcomes of the latter strategy is most challenging. Moreover, a safety net for both missed VS diagnoses and VS growth would ideally be introduced, which would still enable detection of VS (growth) at a later time. It is known that hearing loss progresses faster in the affected ear of patients with intrameatal VSs.[11] This is useful information, but patients will not be able to detect these subtle differences. The question is whether repeated pure-tone audiometry, costing about one fifth of an MRI examination,[12] can identify these differences and could serve as a screening tool for VS growth. The association between hearing deterioration and extrameatal VSs could not be shown,[11] while the importance of detecting VS growth is more relevant for this patient category because of the potential indication for treatment.

METHODOLOGICAL LIMITATIONS

This thesis contains two studies in which a prediction model is developed, one that can be used for diagnosing VS and one to predict VS growth for patients in a W&S strategy. Although both studies are unique due to their large sample sizes, some limitations should also be discussed. First, for most variables there was missing information, for which we used multiple imputation. This method is commonly used to deal with missing information and dependent on the available data.[13, 14] Although all data were retrospectively collected from patient records, we assume that the specialist's report was accurate.

Second, both models were only internally validated, i.e. shrinkage was performed within the available data of the tested population. External validation is warranted prior to their use in clinical practice because internally validated models tend to be overfitted.

Moreover, in the study that reported on the quality of life in VS patients, the subgroups were small. Confidence intervals were overlapping, which made it difficult to draw firm conclusions. These results were used for the cost-effectiveness analysis of W&S strategies, which contributed to the broad and largely overlapping CIs in this study as well. Among other reasons, it is why the effects of new strategies should be monitored in clinical practice.

These limitations underpin the need for a(n) (inter-)national registry, in which data are prospectively and uniformly collected. In this way, the available strategies can be evaluated. Moreover, a registry will contribute to future studies on VS, as a frequent limitation of VS studies is their limited power. An (inter-) national registry would contribute to studies with its wealth of patient data.

FUTURE PERSPECTIVES

Future studies

We recommend external validation of the Seattle protocol and our own diagnostic model. The final validated diagnostic strategy should subsequently be studied prospectively, e.g. by implementing it in clinical practice followed by the close monitoring of its effects. This analysis is needed in order to assess the consequences of a delayed detection of a VS resulting from a new strategy's implementation. It should reveal the number of clinically relevant delayed diagnoses and their impact. Data from our studies could serve as a comparison. In order to do so, outcomes of different management strategies (W&S, radiotherapy and microsurgery) should be studied. Only in the case where the chosen management strategy or its timing are affected by the delay in the diagnosis or detection of VS growth may a patient's outcome be affected. Hopefully, such a study would also clarify which patients benefit from the treatment. The ultimate goal of the W&S strategy should be to identify only those patients that benefit from (early) treatment. If treatment provides few advantageous effects, we might even refrain from any follow-ups for selected patients.

(Inter-)national collaboration

Given the low incidence of VSs and the current variety in clinical practice, it is a challenge to find suitable data to externally validate new strategies. Increased collaboration on a national and/or international level would enable further testing. Moreover, it would contribute to the exchange of evidence between centres and might reduce differences in clinical practice. Nationally, all VSs are managed in a limited number of centres. If these centres would collaborate in registering all new VS patients, we could overcome one of the main limitations in VS research, namely the limited power of studies. Furthermore,

it would be possible to compare results of the current diagnostic and management strategies of the different centres. Such a registry could contribute to the collection of longitudinal data on not only VS size but also parameters such as the quality of life and functional outcomes (e.g. hearing levels and facial nerve function). Ideally, the registry would be accessible to multiple disciplines since VSs are treated by different specialists (otolaryngology, neurosurgery, radiotherapy). Such registries already exist for other diseases. For example, the Dutch Institute for Clinical Auditing (DICA) provides insight into the quality of care by comparing and analysing outcomes of several diseases.[15] For this purpose, clinical data, treatment strategies, patient reported outcome and experience measurements (PROMs and PREMs) are collected.[15]

Diagnosing large vestibular schwannomas

Clinicians agree on the need for the treatment of VSs of extreme size in order to prevent complications of brainstem compression. In our data consisting of patients in a W&S strategy, the majority (55%) of VSs were classified as Koos grade 1 or 2 at the time of discovery, 22% were classified as Koos 3 or 4. For the remaining 23%, the Koos grade was unknown, but the latter patients were all assigned to a W&S strategy, suggesting that these VSs were not of extreme size. Detection of smaller VSs will initially not result in a health gain for patients. The diagnosis itself could even be regarded as a burden by some patients because once diagnosed, it is unclear what management strategy is best for some lesions. This obscurity may cause uncertainty and doubt. The ultimate aim should be to develop new methods to diagnose VSs and offer an indication for current or future treatment. In order to know which VSs warrant treatment, more evidence is needed on what management strategy is preferred according to the VS size and a patient's symptoms. Such methods might even be used to determine which patients benefit from early intervention instead of using a W&S strategy. Ideally, only the latter lesions would be diagnosed.

Stop screening in symptomatic patients?

What if we dare to completely quit screening in patients with asymmetrical audiovestibular dysfunction? Following the introduction of this strategy, we could expect several scenarios. First, it might lead to a decrease in the number of patients receiving treatment since a proportion of VSs would remain undetected. Second, it might cause an increase in invasive treatments, i.e. microsurgery. This rise in numbers would be the consequence of an increased number of VSs exceeding a size of approximately 3 cm at the time of diagnosis, thereby excluding radiosurgery as a treatment option.[10] Third, it might result in an increase in the number of patients receiving multiple treatments such as radiosurgery and microsurgery. It seems that there is a relation between VS size and tumour control rates following stereotactic radiosurgery.[16-18] If tumour volume at the time of radiosurgery negatively affects tumour control rates, an increase in VS volume at the time of treatment will lead to a decrease in tumour control. Additional treatment might then be necessary.

If this strategy would be implemented, its effects should be monitored closely. Patients will follow different pathways until the diagnosis and will be diagnosed by different specialists. Care should be taken so that the few patients that present with severe symptoms due to the progression of a large VS are registered. Some patients might present in the neurology or neurosurgery department due to signs of brainstem compression. Collaboration between specialists and centres is a prerequisite to identify all VS patients.

Insurance companies are increasingly involved in the development of strategies to improve the efficiency within health care systems. As the current diagnostic strategy costs about 4.5 million euro per year, health care insurers could be motivated to contribute to the introduction of new strategies. If a new strategy is implemented, the number of newly diagnosed VSs should be monitored during a specified period, including whether any adverse events occurred. In order to finance a national registry on VS, collaboration with insurance companies could be explored. It would be an investment of the insurance companies in potential health care cost savings.

Shared decision-making

For many specialists, the idea to stop all screening procedures is too extreme. An alternative could be found in shared decision-making, which is increasingly being applied in the consultation room. For example, consultation cards have been developed by the Otorhinolaryngology Society to help patients decide about common surgeries.[19] The question is what would happen if patients would take the lead. Given the current variety of strategies, there are multiple options available to each patient depending on the clinic they visit. A patient could be given the available information on the probability of having a VS as well as having a VS that needs treatment someday and thereafter decide whether he/she wants a diagnostic MRI examination. Some might consider it important to know what is causing their symptoms, while others might not find it worth the effort to undergo an MRI examination. In practice, patient age and vitality seem to play a role in this kind of decision. It would be interesting to investigate how shared decision-making would affect the number of acquired MRI examinations.

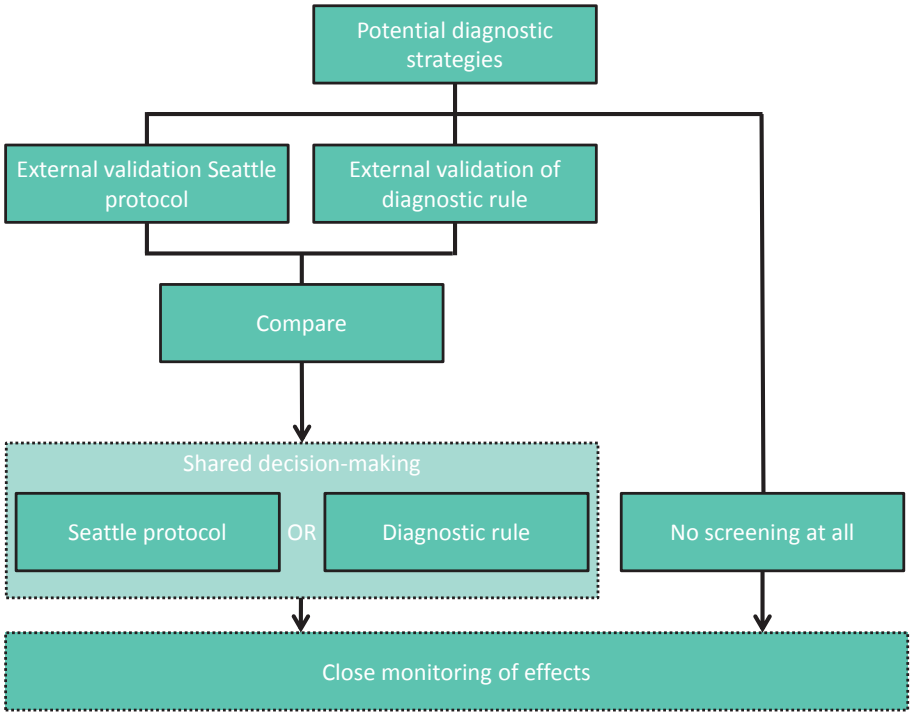


Figure 10.1 Potential future diagnostic strategies.

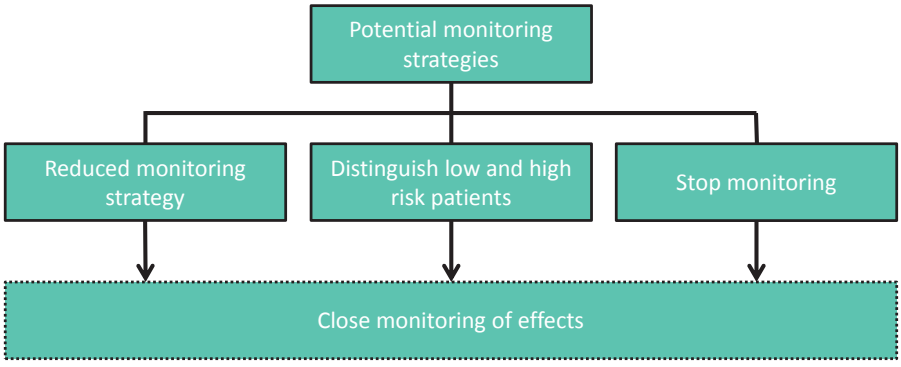


Figure 10.2 Potential future monitoring (W&S) strategies.

CONCLUDING REMARKS

There is substantial room for improvement in the diagnosis and management of VS. There is a lack of high-quality evidence, which may have caused the variation present in clinical practice. In this thesis, we propose several strategies that may lead to more (cost-)effective scenarios, which are summarized in Figures 10.1 and 10.2. None of the described strategies are perfect. The effects that undetected VSs and VS growth have on patient outcomes and long-term cost-effectiveness remain uncertain. We therefore recommend to closely monitor the introduction of any new strategy.

REFERENCES

1. Saliba I, Bergeron M, Martineau G, Chagnon M. Rule 3,000: a more reliable precursor to perceive vestibular schwannoma on MRI in screened asymmetric sensorineural hearing loss. 2011;268(2):207-12. doi: 10.1007/s00405-010-1378-9. PubMed PMID: 20835831.
2. Obholzer RJ, Rea PA, Harcourt JP. Magnetic resonance imaging screening for vestibular schwannoma: analysis of published protocols. 2004;118(5):329-32. doi: 10.1258/002221504323086480. PubMed PMID: 15165304.
3. Dawes PJ, Mehta D, Arullendran P. Screening for vestibular schwannoma: magnetic resonance imaging findings and management. J Laryngol Otol. 2000;114(8):584-8. PubMed PMID: Medline:11027046.
4. Zorginstituut Nederland. Ziektelast in de praktijk. May 2018. Available from: https://www.zorginstituutnederland.nl/publicaties/rapport/2018/05/07/ziektelast-in-de-praktijk.
5. Koos WT, Day JD, Matula C, Levy DI. Neurotopographic considerations in the microsurgical treatment of small acoustic neurinomas. J Neurosurg. 1998;88(3):506-12.
6. Zorginstituut Nederland. Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg. February 2016. Available from: https://www.zorginstituutnederland.nl/publicaties/publicatie/2016/02/29/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg.
7. CBS Statline. Levensverwachting; geslacht, leeftijd (per jaar en periode van vijf jaren). May 2018. Available from: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/37360ned/table?ts=1551767226769.
8. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. BMJ. 1999;318(7182):527-30.
9. Gimsing S. Vestibular schwannoma: when to look for it? J Laryngol Otol. 2010;124(3):258-64.
10. Pai I, Bowman J, Thomas N, Kitchen N, Strong A, Obholzer R, et al. Management of large and giant vestibular schwannomas. Skull Base. 2011;21(6):379-84.
11. van Linge A, Borsboom GJ, Wieringa MH, Goedegebure A. Hearing Loss Progresses Faster in Patients With Growing Intracanalicular Vestibular Schwannomas. Otol Neurotol. 2016;37(9):1442-8.
12. Radboudumc. Passantentarieven. August 2015. Available from: https://www.radboudumc.nl/getmedia/ba7d64d9-ae01-41d5-958f-fbaf961d19f6/Passantentarieven_Radboudumc_2015.aspx.
13. Altman DG, Bland JM. Missing data. BMJ. 2007;334(7590):424.
14. van Buuren S. Flexible Imputation of Missing Data: Chapman and Hall/CRC; 2012.
15. Dutch Institute for Clinical Auditing (DICA) [Internet]. DICA; [cited March 17, 2019]. Available from: https://dica.nl/.
16. Frischer JM, Gruber E, Schoffmann V, Ertl A, Hoftberger R, Mallouhi A, et al. Long-term outcome after Gamma Knife radiosurgery for acoustic neuroma of all Koos grades: a single-center study. J Neurosurg. 2018:1-10.
17. Bailo M, Boari N, Franzin A, Gagliardi F, Spina A, Del Vecchio A, et al. Gamma Knife Radiosurgery as Primary Treatment for Large Vestibular Schwannomas: Clinical Results at Long-Term Follow-Up in a Series of 59 Patients. World Neurosurg. 2016;95:487-501.

18. Klijn S, Verheul JB, Beute GN, Leenstra S, Mulder JJS, Kunst HPM, et al. Gamma Knife radiosurgery for vestibular schwannomas: evaluation of tumor control and its predictors in a large patient cohort in The Netherlands. J Neurosurg. 2016;124(6):1619-26.
19. Consultkaart. Voorlichting op het gebied van Keel-Neus-Oorheelkunde en Heelkunde van het Hoofd-Halsgebied [Internet]. Nederlandse Vereniging voor Keel-Neus-Oorheelkunde en Heelkunde van het Hoofd-Halsgebied; 2017 [cited March 7, 2019]. Available from: https://www.kno.nl/consultkaart/.

CHAPTER 11

Summary

PART I: CURRENT PRACTICE AND EVIDENCE

Vestibular schwannomas (VSs) are benign lesions originating from the nerve sheath of the vestibulocochlear nerve. They are the most common type of lesions occurring in the cerebellopontine angle (CPA). The vestibulocochlear nerve transports sensory information from the inner ear to the brain. VSs can therefore cause asymmetrical audiovestibular complaints (AAC, asymmetrical sensorineural hearing loss, unilateral tinnitus and/or dizziness). Currently, all patients with these symptoms undergo a Magnetic Resonance Imaging (MRI) examination. This leads to a substantial amount of MRIs with negative findings, because the incidence of VS in this screening population varies between 1 and 4.7% (i.e. more than 95% of MRIs are negative for VS).

Once diagnosed, a majority of VS patients is assigned to a 'wait and scan' (W&S) strategy, which consists of repeated MRIs over time, since VS growth may form an indication for treatment. Growth is detected in approximately 40-60% of VSs at some point during follow-up, the percentage that is eventually treated is slightly smaller. Again, a substantial amount of MRIs acquired during W&S is negative for growth and a substantial proportion of patients remains untreated.

The overall aims of this thesis were to optimize 1) the (cost-)effectiveness of various diagnostic strategies in patients with AAC, i.e. suspected of a CPA lesion, and 2) the (cost-)effectiveness of (alternative) W&S strategies in patients diagnosed with a VS.

In **chapter 2**, we therefore performed a systematic diagnostic literature review and meta-analysis in which we assessed the diagnostic accuracy of different non-imaging screening protocols that can be used prior to MRI to select patients at high risk of VS. We included studies that compared non-imaging screening protocols to MRI as gold reference standard. Five pure-tone audiometry protocols were studied by multiple authors; pooled estimates for sensitivity ranged from 88% (95% confidence interval (CI) 84-91) to 91% (95% CI: 52-99) and specificity from 31% (95% CI: 10-66) to 58% (95% CI: 49-65). Due to heterogeneity we were unable to pool the results of the other tests (auditory brainstem response, presenting symptoms, electronystagmography, caloric irrigation and hyperventilation test), but all reported low diagnostic accuracy. Based on the pooled results from this meta-analysis, the AAO-HNS protocol which prescribes an MRI for patients with an average asymmetry of ≥ 15 dB at frequencies 0.5–3 kHz has the highest combination of sensitivity (91% (95% CI: 84-95) and specificity (58% (95% CI: 49-65).

Chapter 3 compares international diagnostic and management strategies for VS, for which we sent a questionnaire to otolaryngologists. Guidelines for diagnosis and management of VS are used by 44% and 42% of respondents, respectively. In the diagnostic strategy for VS, different types and combinations of audiovestibular function tests are used when deciding whether a patient should undergo an MRI. Most respondents (84%) prefer a W&S strategy in case of a small intrameatal VS (Koos 1). Variety in

management strategies increases for patients with a medium to large sized VS (Koos 2, 3 and 4). The details of each management strategy (W&S, microsurgery, stereotactic radiosurgery and fractionated radiotherapy) also differ among respondents. More evidence and/or consensus seem warranted to reduce uncertainties for patients, and differences in outcome and costs that might result from the variety of currently applied strategies.

PART II: DIAGNOSIS OF VESTIBULAR SCHWANNOMA

Since MRIs are costly and have a low yield in case of VS screening, we have examined the potential savings of a hypothetical perfect diagnostic test in **chapter 4**. This test should be used prior to MRI, so only a subset of patients will be referred for MRI. Using a decision analytical model, we compared the current strategy to hypothetical new strategies, assigning MRI to the following: 1) all patients with pathology, 2) all patients with important pathology and 3) only patients with
VS. This resulted in potential cost savings for each strategy, ranging from €256 (95% CI: €250 - €262) per patient for strategy 1 to €293 (95% CI: €290 - €296) per patient for strategy 3. Future diagnostic strategies can cost up to these amounts per patient to be still cost saving. Annually, for the Netherlands, €2.8 to €3.2 million could be saved. The model shows that substantial savings could be generated if it is possible to further optimize the diagnosis of VS.

In **chapter 5** we evaluated the diagnostic accuracy of high resolution T2-weighted MRI (T2w) for detecting CPA lesions, including VS, compared to a combined protocol including gadolinium enhanced T1-weighted MRI (GdT1w). On T2w all CPA lesions >2 mm in size were identified (sensitivity: 90% (95% CI: 73.5-97.9)). Negative predictive value reached 99.5% (95% CI: 98.7-99.9). One missed lesion of 2 mm would have been detected in clinical practice, as this was one of 14 patients for which additional GdT1w would have been ordered based on T2w alone, increasing sensitivity to 93% (95% CI: 77.9-99.2) and negative predictive value to 99.7% (95% CI: 98.9-100). Inter-rater agreement between an experienced neuroradiologist and a less experienced observer for T2w was 98% (95% CI: 96.4-98.8). Thus, T2w has a very high diagnostic accuracy for the presence of CPA lesions in patients with AAC. However, in a screening protocol with T2w only, smallest vestibular schwannomas as well as rare differential diagnoses that probably only would be detected on GdT1w may remain unnoticed.

To preselect patients with high risk for a CPA lesion from those with AAD, we developed a prediction model in **chapter 6**. Several diagnostic variables (demographics, symptoms, physical examination and pure-tone audiometry) were studied. The final model contained eleven variables: gender, sudden onset of hearing loss, gradual onset of hearing loss, unilateral tinnitus, complaints of unilateral aural fullness, instability, headache, facial numbness, facial nerve dysfunction during physical examination, and asymmetry in audiometry at 1 kHz and 4 kHz. The model has a higher net benefit than screening all patients when applying a threshold of >1.8%. At the stated threshold 1.1% MRIs can be saved, but

this increases to 44.5% for a threshold of 5%. The proposed diagnostic rule is a first step in identifying patients with a high risk of a CPA lesion among those with AAD. Following external validation, clinicians may use the model to differentiate between subjects with a low and high risk of a CPA lesion, and use it in shared decision making and to explain future strategies.

PART III: MANAGEMENT OF VESTIBULAR SCHWANNOMA

To enable studies on (cost-)effectiveness of VS, we studied quality of life of recently diagnosed patients with VS, stratified by size (Koos stage) and management strategy (i.e. treatment and W&S) in **chapter 7**. A survey containing the disease-specific PANQOL, the SF-36 and EQ-5D-5L was conducted among adult patients with a newly diagnosed unilateral VS. A utility score is a numeric value ranging from death (0) to perfect health (1). Mean utility scores were lowest for treated Koos 2 (0.673) and highest for W&S Koos 1 (0.829) patients. When comparing W&S patients, differences between Koos stages were small. Overall, treated patients had lower scores than W&S patients with the same Koos stage, particularly in Koos 2 (0.673 versus 0.820). The lack of difference in quality of life between Koos stages in patients in the W&S strategy suggests that small and medium-sized VSs might grow without having a large impact on the quality of life. This should be further investigated using longitudinal data.

The latter data on quality of life were used in **chapter 8** to assess the cost-effectiveness of frequently used monitoring strategies for VS, by development of a state transition model. Six monitoring strategies for patients with VS were compared: lifelong annual monitoring; annual monitoring for the first 10 years after diagnosis; scanning at 1-5, 7, 9, 12, 15 years after diagnosis and subsequently every 5 years; a personalized monitoring strategy for small and large tumours; scanning at 1, 2 and 5 years after diagnosis; and no monitoring. Omitting monitoring is least effective (18.23 quality adjusted life years (QALYs) per patient (95% CI: 16.84-19.37)) and lifelong annual monitoring is most effective (18.66 QALYs (95% CI: 17.42-19.65)). Corresponding costs were €6,526 (95% CI: 5,923-7,058) and €9,429 (95% CI: 9,197-9,643) per patient, respectively. Lifelong annual monitoring provided the best value with a net monetary benefit of €363,765 (95% CI: 339,040-383,697), but the overall probability of being most cost-effective compared to the other strategies was still only 23%. There is large uncertainty in the effectiveness of all strategies, with largely overlapping 95% CIs for all strategies. It is therefore unclear which monitoring strategy provides most value for money at this moment.

A large proportion of VSs within a W&S strategy remains stable in size and thus untreated. This contributes to high costs and burdening of hospital visits. Preferably, we would identify patients with a high risk of future VS growth (and treatment). In **chapter 9** we used a cohort of patients diagnosed with a unilateral VS in the Radboudumc between 1990 and 2016. At least one follow-up MRI had to be available. Twenty-one potential predictors were selected based on literature and interviews with experts. Data on demographics, symptoms, audiometry and MRI characteristics at time of diagnosis were collected from medical records. We used multiple imputation for missing data. Cox proportional

hazards with backward selection was used to develop a model to predict VS growth ≥2 mm. Of the 1217 analysed VS patients, 653 (53.7%) showed growth during follow-up. Median time to event (VS growth) was 13 months and median censoring time 44 months. Balance complaints and tinnitus complaints in the affected ear, Koos grade, duration of complaints and intrameatal diameter on MRI were selected as significant predictors. Discrimination of the model (Harrell's C) was 0.69 and calibration was good. Using this model, it is possible to calculate the probability of VS growth at future time points. In this study we report predictions for 1-5 and 10 years following diagnosis and propose a new W&S strategy based on the model.

There is substantial room for improvement in the diagnosis and management of VS. There is a lack of high-quality evidence, which may have caused the variation present in clinical practice. In this thesis, we propose several strategies that may lead to more (cost-)effective scenarios (see figures 10.1 and 10.2 of general discussion). None of the described strategies are perfect. The effects that undetected VSs and VS growth have on patient outcomes and long-term cost-effectiveness remain uncertain. We therefore recommend to closely monitor the introduction of any new strategy.

CHAPTER 12

Summary in Dutch
Nederlandse samenvatting

DEEL I: HUIDIGE STRATEGIE EN WETENSCHAPPELIJK BEWIJS

Het vestibulair schwannoom, ook wel brughoektumor genoemd, is een goedaardige tumor, die uitgaat van de zenuwschede van de nervus vestibulocochlearis, de gehoor- en evenwichtszenuw. Het zijn de meest voorkomende laesies in de brughoek, de ruimte tussen de inwendige gehoorgang en de kleine hersenen. De nervus vestibulocochlearis, die door de inwendige gehoorgang en brughoek loopt, geeft informatie van het binnenoor door aan de hersenen. Brughoektumoren kunnen daarom asymmetrische gehoor- en evenwichtsklachten veroorzaken (asymmetrisch perceptief gehoorverlies, eenzijdige of asymmetrische tinnitus (oorsuizen) en/of duizeligheid). Momenteel ondergaan alle patiënten met deze symptomen een Magnetic Resonance Imaging (MRI) onderzoek. Dit resulteert in een groot aantal MRI's waarbij geen afwijkingen gevonden worden, omdat het voorkomen van brughoektumoren in deze screeningpopulatie slechts 1% tot 4,7% is. Oftewel, meer dan 95% van de MRI's toont geen brughoektumor.
Eenmaal gediagnosticeerd, komt het merendeel van de patiënten met een brughoektumor in een zogenaamd 'wait and scan' (W&S) traject. Dit is een afwachtend beleid, waarbij op afgesproken tijden een controle MRI gemaakt wordt. Het merendeel van de patiënten heeft nooit een behandeling nodig, maar groei van een brughoektumor kan een reden zijn om wel over te gaan tot behandeling. Groei wordt in ongeveer 40-60% van de brughoektumoren vastgesteld gedurende de follow-up, maar het percentage behandelde patiënten ligt lager, omdat bij kleinere brughoektumoren wat langer wordt afgewacht. Binnen het W&S traject is er dus een aanzienlijk deel van de MRI's waarop geen groei wordt gezien van de brughoektumor, en een nog groter deel van de patiënten wordt nooit behandeld.

Het doel van het onderzoek dat in dit proefschrift beschreven is, is om de kosteneffectiviteit te optimaliseren. Dit geldt zowel voor verschillende diagnostische strategieën voor patiënten die verdacht worden van een brughoektumor, als voor (alternatieve) W&S strategieën in patiënten die reeds gediagnosticeerd zijn met een brughoektumor.

In **hoofdstuk 2** beschrijven we de resultaten van een literatuurstudie en meta-analyse waarin we de diagnostische nauwkeurigheid van verschillende niet-beeldvormende screeningmethodes beoordelen. Deze screeningmethodes worden gebruikt om patiënten met een hoog risico op een brughoektumor te selecteren voor MRI onderzoek. We includeerden studies die niet-beeldvormende screeningmethodes vergeleken met MRI als referentiestandaard. De gecombineerde (gepoolde) resultaten van vijf protocollen die gebruik maakten van uitslagen van gehoortesten toonden een sensitiviteit (terecht positieve uitslagen onder patiënten met een brughoektumor) variërend van 88% (95% betrouwbaarheidsinterval (BI): 84-91%) tot 91% (95% BI: 52-99%) en een specificiteit (terecht negatieve uitslagen onder patiënten zonder brughoektumor) van 31% (95% BI: 10-66%) tot 58% (95% BI: 49-65%). Op basis van de gepoolde resultaten van de meta-analyse, behaalde het AAO-HNS protocol de hoogste combinatie van sensitiviteit (91% (95% BI: 84-95%)) en specificiteit (58% (95% BI: 49-65%)); dit protocol adviseert een MRI voor patiënten met een gemiddeld verschil in gehoorverlies tussen

beide oren van ≥15 decibel (dB) op de frequenties 0,5–3 kilohertz (kHz). Als gevolg van de verschillen in zowel de gebruikte methoden als de onderzochte populaties was het niet mogelijk om de resultaten van de andere niet-beeldvormende screeningmethoden (hersenstamaudiometrie, symptomen, elektronystagmografie, calorische testen en de hyperventilatietest) te poolen, maar voor al deze testen werd een lage diagnostische nauwkeurigheid gerapporteerd.

Hoofdstuk 3 geeft een overzicht van de diverse internationale diagnostische en management strategieën voor brughoektumoren. Hiervoor werd een vragenlijst verstuurd naar KNO-artsen in elf landen. Richtlijnen voor diagnostiek en management van brughoektumoren worden door respectievelijk 44% en 42% van de respondenten gebruikt. Er worden voor de diagnostiek van brughoektumoren verschillende types en combinaties van gehoor- en evenwichtstesten gebruikt om te bepalen of een patiënt een MRI moet ondergaan. De meeste respondenten (84%) hebben voorkeur voor een W&S beleid voor kleine, intrameataal (in de inwendige gehoorgang) gelegen brughoektumoren (gradering: Koos 1). De variatie in behandelstrategieën neemt toe voor de middelmatige tot grote brughoektumoren (Koos 2, 3 en 4). De details van elke behandelstrategie (W&S, microchirurgie, stereotactische radiochirurgie en gefractioneerde radiotherapie) verschillen ook tussen respondenten. Om zowel de onzekerheden voor patiënten, de verschillen in klinische uitkomsten en de kosten te reduceren die mogelijk het gevolg zijn van de variatie in behandelstrategieën, is het nodig om meer wetenschappelijk bewijs en/of consensus te verkrijgen.

DEEL II: DIAGNOSTIEK VAN BRUGHOEKTUMOREN

Omdat MRI's kostbaar zijn en er 21 tot 100 MRI's nodig zijn om één brughoektumor te vinden, hebben we in **hoofdstuk 4** de potentiële kostenbesparing onderzocht van een hypothetische perfecte diagnostische test. Deze test zou voorafgaand aan een MRI toegepast moeten worden, zodat alleen die patiënten die ook daadwerkelijk een brughoektumor hebben, verwezen worden voor een MRI. Door middel van een besliskundig model vergeleken we de huidige strategie met hypothetische nieuwe strategieën, waarbij de volgende patiënten een MRI ondergingen: 1) alle patiënten met een afwijking op de MRI scan, 2) alle patiënten met een klinisch relevante afwijking op de MRI scan en 3) alle patiënten met een brughoektumor. Dit resulteerde in potentiële kostenbesparingen per strategie, variërend van €256 (95% BI: €250 - €262) per patiënt voor strategie 1, tot €293 (95% BI: €290 - €296) per patiënt voor strategie 3. Op jaarbasis leidt dit tot potentiële besparingen van respectievelijk €2,8 tot €3,2 miljoen in Nederland. Indien de aanvullende kosten van toekomstige strategieën lager zijn dan bovengenoemde bedragen, dan zullen deze strategieën kostenbesparend zijn. Het model toont dus aan dat aanzienlijke besparingen mogelijk zijn, indien het lukt om de diagnostiek van brughoektumoren verder te optimaliseren.

In **hoofdstuk 5** bestudeerden we de diagnostische nauwkeurigheid van verschillende MRI-protocollen. Afhankelijk van het gekozen MRI protocol worden weefsels anders weergegeven. We vergeleken de nauwkeurigheid van hoge resolutie T2-gewogen MRI (T2w) in het detecteren van brughoektumoren, met een gecombineerd protocol waaraan een T1-gewogen MRI met contrast (GdT1w) was toegevoegd. Op T2w werden alle 27 brughoektumoren van >2 mm omvang (van de in totaal 30 brughoektumoren) gedetecteerd (sensitiviteit: 90% (95% BI: 73,5-97,9%)) in de 678 bekeken oren. De negatief voorspellende waarde (proportie patiënten met een negatieve MRI die daadwerkelijk geen brughoektumor heeft) was 99,5% (95% BI: 98,7-99,9%). Eén gemiste laesie van 2 mm zou in de klinische praktijk wel gedetecteerd zijn, omdat dit één van de 14 patiënten was waarbij er twijfel bestond op basis van T2w alleen. Voor deze groep zouden we in de klinische praktijk aanvullende GdT1w beelden vervaardigd hebben op basis van de bevindingen op T2w. Hiermee steeg de sensitiviteit naar 93% (95% BI: 77,9-99,2%) en de negatief voorspellende waarde naar 99,7% (95% BI: 98,9-100%). De overeenkomst in betrouwbaarheid tussen een ervaren neuroradioloog en een minder ervaren beoordelaar bedroeg 98% (95% BI: 96,4-98,8%) voor T2w. Kortom, T2w heeft een zeer hoge diagnostische betrouwbaarheid bij het detecteren van brughoektumoren bij patiënten met asymmetrische gehoor- en evenwichtsklachten. Echter, in een screeningprotocol waarbij alleen T2w vervaardigd wordt, kunnen de kleinste brughoektumoren en vermoedelijk ook zeldzame andere afwijkingen gemist worden, die wel gedetecteerd zouden worden middels GdT1w. Het is onduidelijk wat de klinische relevantie hiervan is.

Om een preselectie te maken van patiënten met een hoog risico op een brughoektumor binnen de groep patiënten met asymmetrische gehoor- en evenwichtsklachten, werd een diagnostisch model ontwikkeld in **hoofdstuk 6**. Diverse diagnostische variabelen (demografische gegevens, symptomen, lichamelijk onderzoek en toonaudiometrie) werden bestudeerd. Het uiteindelijke model bestond uit elf variabelen: geslacht, plotseling ontstaan van gehoorverlies, geleidelijk ontstaan van gehoorverlies, éénzijdig oorsuizen (tinnitus), een vol gevoel in één oor, instabiliteit, hoofdpijn, doof gevoel in het gelaat, afwijkende functie van de aangezichtszenuw bij lichamelijk onderzoek en de asymmetrie tussen beide oren bij toonaudiometrie op de 1 en 4 kHz frequentie. Op basis van het model kan per patiënt berekend worden wat het risico is dat hij/zij een brughoektumor heeft. Voor een dergelijk model kan een netto opbrengst berekend worden, wat een verhouding weergeeft tussen de correct positieven (mensen die terecht een MRI ondergaan, omdat er een brughoektumor ontdekt wordt) versus de vals positieven (mensen die onterecht een MRI ondergaan, omdat ze geen brughoektumor blijken te hebben). Het model heeft een hogere netto opbrengst dan het screenen van alle patiënten wanneer een drempelwaarde van >1,8% wordt toegepast als verwijscriterium voor MRI. Indien we dit model zouden gebruiken met een drempelwaarde van 1,8%, kan een afname in het aantal MRI's van 1,1% behaald worden. Echter, dit percentage neemt toe tot 44,5% wanneer een drempelwaarde van 5% wordt toegepast. Het voorgestelde model is een eerste stap in de selectie van patiënten met een hoog risico op een brughoektumor binnen de groep met asymmetrische gehoor- en evenwichtsklachten. De diagnostische betrouwbaarheid zou verder verbeterd, en het model extern gevalideerd moeten worden voordat het kan worden toegepast in de klinische praktijk. Artsen kunnen het model in de toekomst

gebruiken om te differentiëren tussen patiënten met een hoog en laag risico op een brughoektumor. Daarnaast kan het model gebruikt worden in gedeelde besluitvorming tussen patiënt en dokter ('shared decision making'), dat wil zeggen in het overleg waarin het beleid nader besproken wordt.

DEEL III: BEHANDELING VAN BRUGHOEKTUMOREN

Om studies naar (kosten)effectiviteit van brughoektumoren mogelijk te maken, hebben we in **hoofdstuk 7** de kwaliteit van leven van recent gediagnosticeerde brughoektumorpatiënten bestudeerd, ingedeeld naar tumorgrootte (Koos klasse) en behandelstrategie, te weten een ingreep (bestraling/operatie) of W&S. Drie vragenlijsten (de ziekte specifieke PANQOL, de SF-36 en EQ-5D-5L) werden verspreid onder volwassen patiënten met een recent gediagnosticeerde eenzijdige brughoektumor. Op basis van de vragenlijsten kon onder andere een utiliteitsscore berekend worden. Dit is een numerieke waarde die kan variëren van dood (0) tot perfecte gezondheid (1). Gemiddelde utiliteitsscores waren het laagste voor behandelde Koos 2 patiënten (0,673) en het hoogste voor Koos 1 patiënten in een W&S traject (0,829). De verschillen tussen alle W&S patiënten (Koos 1 tot 4) waren klein. Over het algemeen hadden behandelde patiënten lagere scores dan patiënten met een W&S beleid bij hetzelfde Koos stadium, met name bij Koos 2 was het verschil tussen patiënten die behandeld waren en patiënten met een W&S beleid relatief groot (0,673 versus 0,820). Het gebrek aan verschil in kwaliteit van leven tussen de verschillende Koos stadia voor patiënten in een W&S traject suggereert dat kleine en middelgrote brughoektumoren kunnen groeien zonder dat dit grote impact heeft op de kwaliteit van leven. Dit dient verder onderzocht te worden door de kwaliteit van leven van brughoektumorpatiënten te volgen over de tijd.

De verkregen data over kwaliteit van leven uit hoofdstuk 7 werden vervolgens gebruikt om de kosteneffectiviteit van veelgebruikte W&S strategieën te beoordelen (**hoofdstuk 8**)**.** Zes mogelijke alternatieve W&S strategieën voor patiënten met een brughoektumor werden vergeleken, waarbij op de volgende momenten een MRI werd vervaardigd: jaarlijks gedurende de rest van het leven; jaarlijks voor de eerste tien jaar na diagnose; 1-5, 7, 9, 12 en 15 jaar na diagnose en nadien elke 5 jaar; een gepersonaliseerde follow-up strategie voor kleine en grote tumoren; 1, 2 en 5 jaar na diagnose; en helemaal geen follow-up. De kosten werden uitgedrukt in euro's, de effectiviteit in 'quality adjusted life years' (QALYs), waarbij één QALY gelijk staat aan een jaar leven in perfecte gezondheid. Het staken van de follow-up bleek het minst effectief (18,23 QALYs per patiënt (95% BI: 16,84-19,37)), terwijl een jaarlijkse follow-up gedurende de rest van het leven het meest effectief bleek te zijn (18,66 QALYs (95% BI: 17,42-19,65)). Bijbehorende kosten bedroegen respectievelijk €6.526 (95% BI: 5.923-7.058) en €9.429 (95% BI: 9.197-9.643) per patiënt. 'Net-monetary benefit' is de waarde van een strategie uitgedrukt in geld. De strategie met de hoogste 'net-monetary benefit' is het meest kosteneffectief. Een jaarlijkse follow-up gedurende de rest van het leven leverde het meeste op qua gezondheidswinst, met een 'net monetary benefit' van €363.765 (95% BI: 339.040-383.697). Echter, de algehele kans dat dit de meest kosteneffectieve

strategie is, vergeleken met de andere strategieën was slechts 23%. Er is grote onzekerheid rondom de effectiviteit van alle strategieën, met grotendeels overlappende 95% BI's tussen de strategieën. Om deze reden is het momenteel onduidelijk welke follow-up strategie het meeste oplevert.

Een groot deel van de brughoektumoren in een W&S traject blijft stabiel in omvang en daarom onbehandeld. Dit draagt bij aan de hoge kosten en belasting voor de patiënt ten gevolge van de ziekenhuisbezoeken voor de controles en MRI scans. Bij voorkeur zouden we patiënten met een hoog risico op groei van hun brughoektumor (en behandeling) identificeren. In **hoofdstuk 9** onderzochten we een groep patiënten uit het Radboudumc, gediagnosticeerd met een brughoektumor tussen 1990 en 2016, waarvan ten minste één follow-up MRI beschikbaar was. Eenentwintig potentiële voorspellers voor brughoektumorgroei werden geselecteerd op basis van literatuur en interviews met experts. Data met betrekking tot demografie, symptomen, toonaudiometrie en MRI karakteristieken ten tijde van de diagnose werden verzameld uit het patiëntendossier. We gebruikten multipele imputatie voor missende data. Vervolgens ontwikkelden we een voorspellend ('Cox proportional hazards') model om brughoektumorgroei van ≥2 mm te voorspellen. Bij 653 (53,7%) van de 1217 geanalyseerde brughoektumorpatiënten groeide de brughoektumor tijdens de follow-up. De mediane tijd totdat er groei optrad was 13 maanden en de mediane follow-up van patiënten waarbij geen groei vastgesteld werd, bedroeg 44 maanden. Balansproblemen, oorsuizen in het aangedane oor, Koos klasse, duur van de klachten en de diameter van de brughoektumor in de inwendige gehoorgang op MRI werden geselecteerd als significante voorspellers van groei. De discriminatie, het vermogen van het model om het wel/niet groeien van een brughoektumor goed te voorspellen, was 0,69 (uitgedrukt als Harrell's C, waarbij 0,5 het slechtst en 1 het best haalbare is) en de kalibratie (overeenkomst van frequentie van geobserveerde groei en de voorspelde kans op groei) was goed. Dit model kan gebruikt worden om de kans op groei van een brughoektumor te berekenen 1, 2, 3, 4, 5 en 10 jaar na de diagnose. Via https://vs-model.shinyapps.io/predictVSgrowth kunnen verschillende variaties ingevoerd worden om het model voorspellingen te laten berekenen.

Er is aanzienlijke ruimte voor verbetering binnen de diagnostiek en de behandeling van brughoektumoren. Kwalitatief hoogstaand wetenschappelijk bewijs ontbreekt vooralsnog, wat de huidige praktijkvariatie kan verklaren. In dit proefschrift stellen we een aantal strategieën voor die mogelijk leiden tot meer (kosten)effectieve scenario's (zie figuren 10.1 en 10.2 van de 'general discussion'). Geen van de beschreven strategieën is perfect. Het effect op de uiteindelijke uitkomst en de lange termijn kosteneffectiviteit van een brughoektumor die gemist wordt tijdens de diagnose, dan wel het missen van groei van een brughoektumor blijft onduidelijk. Daarom is ons advies om de invoer van een nieuwe strategie nauwkeurig te monitoren.

PART V

Appendices

CHAPTER 13

List of publications

Wegner I, Bittermann AJN, Hentschel MA, van der Heijden GJM, Grolman W. Pure-tone audiometry in otosclerosis. Insufficient evidence for the diagnostic value of the Carhart notch. *Otolaryngol Head Neck Surg* 2013; 149(4): 528-32.

Hentschel MA, Huizinga P, van der Velden DL , Wegner I, Bittermann AJN, van der Heijden GJM, Grolman W. Limited evidence for the effect of sodium fluoride on progression of hearing loss in patients with otosclerosis. *Otol Neurotol* 2014; 35(6): 1052-7.

Hentschel MA, ten Koppel PGJ, Bonte-Mineur F. Een zeldzame oorzaak voor aften in de mond. *Nederlands Tijdschrift voor Keel-Neus-Oorheelkunde* 2015; 2: 75-8.

Hentschel MA, Haaksma ML, van de Belt TH. Wearable technology for the elderly: Underutilized solutions. *Eur Ger Med* 2016; 7: 399-401.

Hentschel MA, Scholte M, Steens SCA, Kunst HPM and Rovers MM. The diagnostic accuracy of non-imaging screening protocols for vestibular schwannoma in patients with asymmetrical hearing loss and/or unilateral audiovestibular dysfunction: a diagnostic review and meta-analysis. *Clin Otolaryngol* 2017; 42(4): 815-823.

Scholte M, Hentschel MA, Kunst HPM, Steens SCA and Rovers MM. Potential savings in diagnosing vestibular schwannoma. *Clin Otolaryngol* 2018; 43(1): 285-290.

Hentschel MA, Kunst HPM, Rovers MM and Steens SCA. Diagnostic accuracy of high resolution T2-weighted MR versus contrast enhanced T1-weighted MR to screen for cerebellopontine angle lesions in symptomatic patients. *Clin Otolaryngol* 2018; 43(3): 805-811.

Scholte M, Hentschel MA, Hannink G, Kunst HPM, Steens SCA, Rovers MM, Grutters JPC. In search of the most cost-effective monitoring strategy for vestibular schwannoma: a decision analytical modelling study. *Clin Otolaryngol* 2019; 44(4): 525-533.

Hentschel MA, Rovers MM, Markodimitraki L, Steens SCA, Kunst HPM. An international comparison of diagnostic and management strategies for vestibular schwannoma. *Eur Arch Otorhinolaryngol* 2019; 276(1): 71-78.

Hentschel MA, Rovers MM, Steens SCA, Hannink G and Kunst HPM. Development of a diagnostic model to identify patients at high risk for cerebellopontine angle lesions. *Submitted*.

Hentschel MA, Hannink G, Steens SCA, Mulder JJS, Rovers MM, Kunst HPM. Development of a model to predict vestibular schwannoma growth: an opportunity to introduce new wait and scan strategies. *Submitted*.

CHAPTER 14

Dankwoord

Misschien schrijf je het dankwoord met een reden als allerlaatste. Na het jarenlang ontwikkelen van je schrijfvaardigheden mag je de belangrijkste tekst (die hoogstwaarschijnlijk ook door de meeste mensen gelezen wordt) als laatste op papier zetten. *Save the best for last*. En terecht, want zonder onderstaande mensen was het leven van deze promovenda onmogelijk en/of een stuk minder leuk.

Prof. dr. Rovers, lieve Maroeska, wat heb ik respect voor jouw werkzaamheden. Ik bewonder hoe je, ondanks de variatie aan specialismen waar je mee samenwerkt, zicht blijft houden op het grote geheel en de klinische toepasbaarheid van onderzoek niet uit het oog verliest. Je weet altijd tot de kern van het probleem te komen en de resultaten te interpreteren aan de hand van de klinische praktijk. En dat alles in nog meer vakgebieden dan de KNO alleen… Daarnaast is het heel fijn dat evenveel van je aandacht uitgaat naar de persoon achter het onderzoek als naar het onderzoek zelf. Ik ben dankbaar voor je leerzame begeleiding en de ruimte die je me geboden hebt voor mijn persoonlijke ontwikkeling.

Prof. dr. Kunst, beste Dirk, wat is het leuk om met jou te werken met je tomeloze enthousiasme voor schedelbasispathologie. Je zit altijd vol nieuwe ideeën en weet verder te denken dan de standaard gebaande paden. Dit zorgde voor een inspirerende tijd. Bedankt! Ook voor de gezelligheid bij borrels, congressen en het schedelbasisdinertje. Ik kijk er naar uit om met je te blijven samenwerken tijdens de rest van mijn opleiding.

Dr. Steens, beste Stefan, bedankt dat je me wegwijs hebt gemaakt in de wereld van de radiologie. Met eindeloos geduld heb je gezorgd dat ik als bijna-leek heel wat MRI's heb weten te beoordelen. Daarnaast was het leerzaam om jouw kennis over kwaliteit en veiligheid te betrekken bij het onderzoek. Bedankt voor de fijne samenwerking: je snelle reactie op stukken en optimisme waren erg fijn!

Graag wil ik de betrokken patiënten van de commissie brughoektumoren van Stichting Hoormij bedanken voor hun waardevolle input. Prof. dr. Willem A.J. van Daal en Marco van Stralen, de gesprekken met jullie verruimden mijn blik en zorgden voor nieuwe ideeën, veel dank daarvoor.

Beste Gerjon, mijn R- en predictiemodelcoach… Hoe vaak zouden wij een 'Koffie en R' afspraak gehad hebben in de afgelopen jaren? Ik ben de tel inmiddels kwijt. Hoe jij alle statistische puzzels binnen de modellen wist op te lossen snap ik nog steeds niet altijd, maar zonder jou geen predictiemodel! Bedankt voor je geduld, de eindeloze syntax aanpassingen en gezelligheid.

Lieve Janneke, het was erg leuk om samen met jou Mirre te begeleiden. Je analytische blik op de modellen was van grote waarde. Bedankt voor deze leerzame tijd.

Beste Jef en Jerome, Dirk en ik hebben jullie heel wat keren geconsulteerd als experts binnen het vakgebied. Bedankt voor jullie interesse in het onderzoek en jullie hulp bij het doorhakken van knopen. Jef, jij natuurlijk ook veel dank voor je inhoudelijke bijdrage aan het follow-up stuk.

Prof. dr. Grotenhuis en dr. Mark ter Laan, heel fijn dat jullie tijd wilden maken voor interviews met studenten en mee wilden denken over diverse dilemma's.

Beste heren van het Gamma-Knife Centrum Tilburg, bedankt voor de inspirerende middagen waarbij er ideeën met betrekking tot onderzoek werden uitgewisseld. Dat was zeer waardevol.

Special thanks to the European Skull Base Society, who provided us with contact details of their ENT members. Without the ESBS, chapter 3 of this thesis would not exist.

Beste Andy, bedankt voor al je uitleg over de BERA en het evenwichtsonderzoek. Helaas waren er uiteindelijk te weinig onderzoeken beschikbaar om deze data te gebruiken binnen de modellen. Gelukkig kan Stan alsnog verder met het onderzoek naar de kosteneffectiviteit van de BERA.

Stafsecretariaat KNO, in het bijzonder Loes, jullie stonden soms voor de onmogelijke taak om een afspraak in te plannen met Maroeska, Dirk en Stefan samen. Bedankt voor al jullie hulp!

Chris-Jan, bedankt voor al je hulp bij het verzamelen van data uit AudiologicX, dat heeft mij heeeel veel tijd en moeite bespaard!

Beste Sven Lafebre, bedankt voor alle hulp. Zowel het vinden van alle benodigde MRI-onderzoeken als het ontwerpen van de software om de beelden in te kunnen beoordelen waren onmisbaar voor dit proefschrift.

Dames van de poli, bedankt voor jullie hulp met het in- en uitchecken en opruimen van de enorme stapel papieren dossiers. Jullie moeten net zo blij geweest zijn als ik toen dat project eindelijk klaar was…

Lieve Mirre en Stan, samen onderzoek doen is een stuk leuker dan alleen! Bedankt voor alles wat jullie me hebben geleerd over HTA en de fijne samenwerking tijdens jullie stages. Erg leuk (en terecht!) dat jullie de EBS groep nooit meer hebben verlaten…

Laura en Beau, bedankt voor jullie bijdrage aan dit boekje en de fijne samenwerking.

Beste stafleden KNO, in het bijzonder prof. dr. Marres en dr. van den Hoogen, bedankt voor de leerzame en bovenal gezellige opleiding die jullie weten neer te zetten. Ik ben vereerd dat ik hier deel van uit mag maken en kijk heel erg uit naar de rest van mijn opleiding in Nijmegen.

Beste heren uit het CWZ, jullie ook bedankt voor het onvergetelijke jaar. Wat heb ik veel van jullie geleerd, de tijd vloog voorbij!

Lieve Evidence Based Surgery groep, bedankt voor alle gezelligheid en inspiratie. Ook al werkten we fysiek in een andere ruimte, de vis-vrijdagen, EBS lunches, heidagen en borrels (met als summum de eerste EBS skireis) maakten dat we regelmatig van gedachten konden wisselen over ons onderzoek en minder serieuze zaken.

Lieve (oud-) A(N)IOS en onderzoekers van de KNO, wat is het gezellig om elke dag met jullie te mogen werken! Ik kan me geen leukere collega's wensen. Ondanks de, naar het soms leek, never ending stapels dossiers die uit het archief naar boven kwamen en ingeklopt moesten worden, ging ik altijd met plezier naar mijn werk en dat komt mede door jullie! Dat er nog maar vele borrels, skireizen, weekenden, etc. mogen volgen. Lieve mede-schedelbasisonderzoekers, ik denk met veel plezier terug aan de ESBS congressen, met onder andere gezellige etentjes, whisky arrangementen en obscure kelderdiscotheken. Lieve Loftbewoners, ondanks de zeer minimalistische inrichting was het dankzij jullie toch altijd gezellig in het appartement. Ik denk met veel plezier terug aan de kopjes koffie die 's ochtends klaar stonden, Thai-bestel-avonden, maar ook de avonden temptation island kijken op de bank en ons uitje op de pannenkoekenboot.

Lieve vriendjes en vriendinnetjes: BOB, Halphies, Melt/Wir sind Berliners en anderen die niet passen binnen een van deze groepen. Wat ben ik blij met jullie! Bedankt dat jullie er altijd voor me zijn, ook tijdens minder leuke tijden.

Lieve Rosa, wat begon op het terras van het UMCU, wachtend op een informatieavond over coschappen in het buitenland, werd al snel een onvergetelijke reis naar India (twee groentjes voor het eerst buiten Europa) en een hechte vriendschap. Ook al verlopen onze carrièrepaden in andere steden en binnen een ander specialisme, ze vertonen veel gelijkenissen en dat maakt het extra fijn om hoogte- en dieptepunten te delen. Bedankt voor je jarenlange steun en alle gezelligheid! En bedankt dat je als paranimf aan mijn zijde wil staan op de grote dag.

Lieve Ria & Henk, Leoni & Thomas, Thijs & Floor, wat bof ik met jullie als schoonfamilie. Bedankt voor jullie interesse in de voortgang van mijn onderzoek, maar vooral voor alle gezelligheid! Ik voelde me vanaf moment één thuis bij jullie en dat is ontzettend fijn.

Lief zusje, lieve Poek, onze vroegere vechtpartijen hebben gelukkig al ruime tijd geleden plaatsgemaakt voor een hechte band. Het is altijd weer gezellig met jou en ik kan alles met je delen. Ik ben blij dat jij naast me zult staan als paranimf. Lo, ook jij bedankt voor alle gezelligheid. Juli 2020 wordt een bijzondere maand!

Lieve Papa, wat had ik gehoopt dat jij dit moment nog zou meemaken. Ik heb diep respect voor je positieve levenshouding en doorzettingsvermogen, zeker de laatste jaren. Ondanks dat praten steeds moeilijker ging, bleef je altijd interesse tonen, ook in dit onderzoek. Dankjewel dat je zo'n goed voorbeeld voor me was, als wetenschapper, maar vooral als persoon. *Im Prinzip ist alles ganz einfach*… Ik mis je!

Lieve Mams, ik bewonder het enorm hoe je je de afgelopen jaren staande hebt gehouden. Ik ben je dankbaar voor hoe je altijd voor ons klaar staat. Je bent altijd geïnteresseerd in ons en denkt altijd met ons mee. We kunnen altijd bij jou terecht voor een wijs advies en dat is superfijn!

Lieve Joost, ik ben gek op je positiviteit, nuchterheid, flauwe woordgrappen en tomeloze energie. Met jou verveel ik me nooit! Wat ben ik blij dat jij er voor mij bent, zowel tijdens leuke als tijdens verdrietige tijden. Lieve Lars, zo klein en nu al zo stoer! Gelukkig hoef ik nu minder tijd achter mijn laptop door te brengen. Ik heb heel veel zin in de tijd die we met zijn drieën tegemoet gaan!

CHAPTER 15

Curriculum vitae

Mayke Hentschel werd in 1988 geboren te Leiden. Ze heeft hier haar hele jeugd gewoond samen met haar ouders en jongere zusje. Na het tweetalig onderwijs en gymnasium te hebben afgerond op het Rijnlands Lyceum in Oegstgeest, gecombineerd met het Pre-University College aan de Universiteit Leiden, is ze in 2006 haar studie Geneeskunde begonnen aan de Universiteit van Utrecht. In 2009 bracht ze drieënhalve maand in India door voor haar coschap Gynaecologie in het Kasturba Hospital Manipal. Twee jaar later onderbrak ze haar studie gedurende drie maanden voor een cursus Spaans in Buenos Aires gevold door een reis door Zuid-Amerika. Tijdens haar studie werd haar interesse voor de Keel-, Neus- en Oorheelkunde gewekt. Daarom volgde zij in haar laatste studiejaar een wetenschappelijke stage en semi-arts stage binnen dit vakgebied. Samen met coassistenten van andere universiteiten organiseerde ze het Nationaal Coassistenten Congres 2013. Haar studie rondde ze in 2013 af, waarna zij als ANIOS Chirurgie in het Maasstad Ziekenhuis te Rotterdam heeft gewerkt. Een jaar later startte ze haar promotietraject op de afdeling Keel-, Neus- en Oorheelkunde in het Radboudumc te Nijmegen, wat heeft geresulteerd in dit proefschrift. Onder supervisie van prof. dr. M.M. Rovers, prof. dr. H.P.M. Kunst en dr. S.C.A. Steens voerde ze diversie studies uit om de kosteneffectiviteit binnen de diagnostiek en follow-up van brughoektumoren te verbeteren. Dit combineerde ze met een opleiding tot epidemioloog, grotendeels gevolgd aan de Radboud Universiteit Nijmegen. Ze is in december 2017 aan haar opleiding tot KNO-arts begonnen, welke zij deels in het Radboudumc (opleider dr. F.J.A. van den Hoogen) en deels in het CWZ te Nijmegen (opleider dr. J.A.M. Engel) volgde. Zij woont samen met Joost en hun zoon Lars in Nijmegen.

CHAPTER 16

PhD portfolio

Institute for Health Sciences
Radboudumc

Name PhD student: M.A. Hentschel	***PhD period: 01-12-2014 – 01-07-2020***
Department: Otorhinolaryngology	***Promotor(s): Prof. dr. M.M. Rovers***
Graduate School: Radboud Institute for Health Sciences	***Co-promotor(s): Prof. dr. H.P.M. Kunst, Dr. S.C.A. Steens***

	Year(s)	ECTS
TRAINING ACTIVITIES		
a) Courses & Workshops		
- RIHS introduction course	2015	0.75
- Basic Course for Clinical Investigators (BROK) course	2015	1
- Prognosis research course	2015	1
- Course multivariable statistical methods (5E003)	2015	5.5
- Course clinical epidemiology (5E004)	2015	5.5
- Refresher course in statistics for PhD-students	2015	1.5
- Course management for PhD-students	2015	3
- Workshop EndNote	2015	0.2
- Course statistical analysis with SPSS and SAS (5E002)	2016	5.5
- 6th Leuven Course on Ear Imaging	2016	1
- Course scientific integrity	2016	1
- Course presentation skills	2016	1.5
- Course clinical prediction models	2016	1
- Course methods of data collection (5E001)	2016	5.5
- Workshop introduction to R	2016	0.1
- Course survival analyses for clinicians	2017	1.9
b) Seminars & lectures		
- Grand rounds and research rounds Radboudumc	2015 - 2017	0.2
- Seminar OOR-ON	2016	0.1
- Organizer seminar OOR-ON	2017 - 2019	2
c) Symposia & congresses		
- European Skull Base Society Congress (2 oral presentations)	2016	1
- Meeting of Dutch ENT-society (oral presentation)	2016	0.5
- European Society for Head and Neck Radiology Congress (poster presentation)	2016	0.5
- RIHS PhD retreat (oral presentation)	2016	0.5
- WEON (oral presentation and poster presentation)	2017	1
- Radiology research meeting, Radboudumc (oral presentation)	2017	0.25
- European Skull Base Society Congress (2 oral presentations)	2018	1
- Meeting of Dutch ENT-society	2015 - 2019	1.5
a) d) Other		
- Book & Journal Club for Junior epidemiologists	2015 - 2017	6
- Working visit Ministry of Health, Welfare and Sport	2016	0.2
TEACHING ACTIVITIES		
e) Lecturing		
- None		
f) Supervision of internships / other		
- Reviewer Audiology & Neurotology	2015	0.1
- Reviewer Acta Oto-Laryngologica	2016	0.1
- Master Student Health Technology Assessment	2016	2
- Master Student Medicine	2016	1
- Master Student Health Technology Assessment	2017	2
TOTAL		**55.9**

Institute for Health Sciences
Radboudumc