



UNIVERSITY OF BERGEN

Department of Clinical Medicine
Section of Ophthalmology

Professor Eyvind Rødahl
Department of Clinical Medicine
University of Bergen
5020 Bergen

Faculty of Medicine
University of Bergen
5020 Bergen

Bergen 9 April 2019

Evaluation report for PhD thesis

Enclosed please find the evaluation report for Øystein Vesterli Tveitens PhD thesis “Long term quality of life and symptoms in patients with vestibular schwannoma”.

Yours sincerely,

Eyvind Rødahl

Visiting address:
Jonas Lies vei 72
N-5021 Bergen

Postal address:
The Faculty of Medicine
Section of Ophthalmology
P.O Box 7804
N-5020 Bergen

Telephone:
+ 47 55974110

Telefax:
+47 55974143

PhD thesis
Long term Quality of Life and symptoms in patients with vestibular schwannoma
Øystein Vesterli Tveiten, MD

The thesis consists of 93 pages and is divided into

- Title page, description of scientific environment and Acknowledgements
- Table of contents
- List of papers
- Abbreviations
- Abstract/summary
- Introduction
- Methods
- Aims
- Summary of results from each of the four studies
- Discussion
- Conclusions
- Papers 1-4
- Future perspectives
- References

Title page, preface with acknowledgements, a list of the four original manuscripts behind the thesis, a list of abbreviations and an abstract constitute the formal structure around the core texts. Thesis language is English throughout. The table of contents with headings and subheadings provides the reader with an immediate, comprehensive overview of the contents.

The reference list contains 318 references, which cover the literature of the thesis topic.

The 4 original papers included in the thesis are all published in highly recognized peer-reviewed neurosurgical journals:

1. Long-term quality of life in patients with vestibular schwannoma: an international multicenter cross-sectional study comparing microsurgery, stereotactic radiosurgery, observation, and nontumor controls. Carlson ML, Tveiten OV, Driscoll CL, Goplen FK, Neff BA, Pollock BE, Tombers NM, Castner ML, Finnkirk MK, Myrseth E, Pedersen PH, Lund-Johansen M, Link MJ. *J Neurosurg.* 2015 Apr;122(4):833-42. doi:10.3171/2014.11.JNS14594.
2. Long-term Auditory Symptoms in Patients With Sporadic Vestibular Schwannoma: An International Cross-Sectional Study. Tveiten OV, Carlson ML, Goplen F, Vassbotn F, Link MJ, Lund-Johansen M. *Neurosurgery.* 2015 Aug;77(2):218-27; discussion 227. doi:10.1227/NEU.0000000000000760.
3. Audiovestibular Handicap and Quality of Life in Patients With Vestibular Schwannoma and "Excellent" Hearing. Tveiten OV, Carlson ML, Link MJ, Lund-Johansen M. *Neurosurgery.* 2017 Mar 1;80(3):386-392. doi: 10.1227/NEU.0000000000001238.
4. Patient- versus physician-reported facial disability in vestibular schwannoma: an international cross-sectional study. Tveiten ØV, Carlson ML, Goplen F, Myrseth E, Driscoll CLW, Mahesparan R, Link MJ, Lund-Johansen M. *J Neurosurg.* 2017 Nov;127(5):1015-1024. doi: 10.3171/2016.8.JNS16707.

In a separate file, there is a brief statement from the candidate's supervisor regarding the candidate's independent contribution to the papers. He is first author of three (no.s 2,3 and 4) of the papers.

The thesis

Introduction

The introduction starts with a general and historical description of vestibular schwannoma (VS), histological and radiological diagnosis described in short passages. Epidemiological aspects, considerations on assessment of tumour size and –growth and the natural course of VS are presented a little more extensively, A larger part of the introduction is dedicated to description of subjective symptoms, objective symptoms (findings) and quantitation of neuro-deficits (mainly hearing and facial nerve function). QOL/patient-perceived outcome is explained in much more detail; correctly stating the increasing focus on QOL issues. The second part of the introduction describes the three management modalities (surgery, ionizing radiation and observation with serial imaging). This is divided into historical accounts of surgery, radiosurgery, and observation followed by the current status of the treatment modalities; the indications, tumour control rates, outcomes, complications and QOL studies.

The introduction is very readable and contains all the relevant information based on thorough review of the literature, which is also reflected in the extensive list of references applied to the introduction. The length and extent of information in the numbered sections of the introduction has a weight, which is nicely balanced to suit the subject and aim of the thesis. The historical sections about the treatment modalities are not only interesting, but also provide a good background for understanding why current treatment preferences and recommendations evolved. There is a natural “division” of the introduction into the first part dealing with “the tumour” and the second part dealing with “the treatment” which could have been emphasized. Overall, the introduction naturally merges towards the purpose of the conducted studies – and would even more, if the first two paragraphs of the section on study purpose (1.3.3.2) had instead been given its own heading; e.g. “limitations of knowledge in current treatment recommendations”.

Overall purpose of the study

This section is the final part of the introduction, however, the aim is usually presented in a paragraph of its own to emphasize that - in contrast to the introduction summarizing research results of others - the purpose, aims and objectives are generated by the candidate himself. In fact most of this section deals with other aspects including more background information and considerations about methodology – although very relevant, these parts belong to the introduction and the method section, respectively.

The stated overall purpose is to observe long term outcome of sporadic, unilateral VS < 3 cm following one treatment decision. Further aims are specified in a section headed “Aims”

- 1) to describe the long-term QOL outcomes following initial treatment for small- and medium-sized VS
- 2) to analyse and review the long-term auditory symptoms (hearing and tinnitus) following initial treatment
- 3) to elucidate the relationship between QOL and hearing difficulties in a VS population
- 4) in part to inform about the long-term facial nerve outcomes following initial treatment, but more importantly to explore the potential implications of a facial nerve paresis to the patient and to review potential differences of the facial nerve outcomes from that of the doctor's perspective.

Methods

The study is observational, non-randomized. In the preceding section on “purpose”, the candidate reflects on the choice of data accumulations through “an observational cohort” with non-randomized treatment as the only viable alternative to obtain this knowledge “given the strong opinions on the subject, the rarity of the tumour, the lack of international collaboration and ... the lack of economic gains”.

Working hypotheses defined a priori to the study design are thus lacking, i.e. the thesis and its studies are entirely observational. A risk of losing direction and obtaining imprecise, ambiguous answers is inherent in observational studies without predefined direct research questions, but here this weakness is appropriately counteracted by sticking to just three outcome parameters (hearing, facial function, QOL); quantitatively comparing self-assessed to objective outcome; and relating specified outcome scores to each of the three treatment modalities – and for all of these selecting outcome measurement tools which are validated for the tested domains.

It is a strength that the studies contain a large cohort, that all three treatment principles were available and used in both participating institutions and that a non-VS control cohort was included. There is a superficial reflection on statistical power by combining the retrospective cohorts from two institutions; the candidate’s own in Norway and a large centre in the USA, but there is no further calculation or assessment of actual statistical power.

It is stated, that both institutions had a prospective VS database from which the data were extracted. It is briefly mentioned, which types of data, the databases contain, but without a better description of the databases, how comparable they are, and how data were accrued, it is not clear if the design is a mixture of retrospective, prospective and ambi-directional methodology. How study design may influence data-quality and analysis is not discussed at all.

Results

This section contains a good summary of each of the four original papers including the one figure from each paper best suited to illustrate the most important findings. 642 patients were included – for the survey self-assessment studies, there was a response rate of 79%, which is very good – but there is no contemplation if the 20% non-responder group could mean a possible selection bias; e.g. if basic epidemiological parameters (general bias questions) or tumour characteristics (disease related bias) differ between responders vs. non-responders or if the non-responder rate was different between the two institutions (selection bias in socio-cultural perspective) . There was a USD 30 payment to participants when completing the survey – could there be a bias introduced by this?

The mean observation time was 7 years for all studies (SD 2-2.5 years). As outcome including number and severity of complications is influenced by observation time – in opposite directions after surgery (with the potential for long term improvement) vs. radiosurgery (with the potential for late radiation damage) – an extended comment on the impact on data quality and symptom burden by setting/not setting a minimal time limit for follow-up duration is warranted. This is particularly relevant when comparing studies, which often differ in length of follow-up periods and in different definitions of follow-up period. This important aspect is only very briefly mentioned in the Discussion on auditory symptoms (page 74, section 5.2.1.)

Discussion

In the discussion, the candidate balances the observations and conclusions from his own four papers against existing publications with the same themes; QOL, hearing/auditory symptoms and facial function. Figures from selected papers are very relevantly used to make these comparisons. The discussion is generally a balanced re-assessment of the results and the current status – very little is offered for future perspectives.

Limitations inherent in the study design are pointed out to be mainly in patient and physician factors favouring selection of certain types of treatment for certain situations. It is correctly concluded that this could influence the strength of outcome differences between the three management types. It is mentioned – and indirectly assumed – that a randomized controlled trial could provide different and statistically more definite results; but the assumption is not further described or challenged. Based on the existing literature, the historical evolution of VS treatment and current logistics, the candidate is quite probably right by stating that a “scientifically optimal” randomized controlled trial is not realistic in the current world of VS treatment. The best possible alternative is then to conduct transparent studies on correctly accumulated data with reliable data-quality and consensus on data definitions, which is the philosophy behind the papers and the thesis. It would have been interesting to speculate how many patients would theoretically have to be included in an RCT to demonstrate statistically significant differences between the management modalities – and to extrapolate this estimate to assess if an RCT could contribute enough additional practical value to justify the difficulties, time and resources going into such a study.

Paper 1. Long-term quality of life in patients with vestibular schwannoma: an international multicenter cross-sectional study comparing microsurgery, stereotactic radiosurgery, observation, and nontumor controls. Carlson ML, Tveiten OV, Driscoll CL, Goplen FK, Neff BA, Pollock BE, Tombers NM, Castner ML, Finnkirk MK, Myrseth E, Pedersen PH, Lund-Johansen M, Link MJ. *J Neurosurg.* 2015 Apr;122(4):833-42.

The candidate is 2nd author. The paper uses self-reported survey scores to compare defined outcome domains in VS patients vs. non-VS controls. Outcome domains are also compared for VS patients with each for the three management options. The figure chosen to illustrate the results in the paper summary in the thesis comprehensively shows the better scores of the “VS specific” domains (facial, hearing, balance) in the control group vs. the entire VS group. Although differences between the treatment groups are much smaller, by using a combination of several scores, it becomes convincing, that the less invasive treatments give better functional scores. However, for the general “well-being” domains like anxiety, energy and general health perception, response patterns are quite mixed. The authors are appropriately cautious in their data interpretation and in their conclusions relating their own observations to other reports on VS growth; patients with non-growing tumours are better off observed; patients with growing tumours who need treatment are marginally better off with radiosurgery; open surgery is primarily for larger tumours with mass effect.

The stated purpose to investigate long-term HRQOL following all three types of treatment is fulfilled and well documented by quantitative scores in relevant VS specific and general domains. Figures, tables and statistics are appropriate with good captions and show the results in a very accessible manner. The authors do not overextend their conclusions and recommendations beyond the statistical strength of the results.

Paper 2. Long-term Auditory Symptoms in Patients With Sporadic Vestibular Schwannoma: An International Cross-Sectional Study. Tveiten OV, Carlson ML, Goplen F, Vassbotn F, Link MJ, Lund-Johansen M. *Neurosurgery*. 2015 Aug;77(2):218-27; discussion 227.

This paper is reporting very poor results on hearing preservation but surprisingly does not really question the reason why the results are so poor compared to several major series from the literature. It is stated in the Methods section: “As a general rule at our centers, small- and medium-sized tumours were initially observed and only given treatment after proven growth on serial imaging studies.” It seems that the authors are not really considering their “strategy of management” as potentially the cause of the poor results. The strategy has been to wait and do radiosurgery or resection only when the tumour is growing and/or the hearing declining. Several radiosurgery (SRS) series of papers that are not quoted, have shown major impact of both the timing of the treatment (Yomo et al 2012, Régis et al 2004, Hillman et al 2010) and the nature of the first symptom. Younger patients (Lobato et al, Régis et al 2004), treated with better hearing (Tamura et al 2009, Régis et al 2004) with no hearing perceived handicap (Mousavi et al 2016), discovered on the occasion of a tinnitus (Régis et al 2004) or unbalance (and not due to hearing loss) are much more likely to preserve hearing on the long term. Thus, upfront proactive SRS is reported to better preserve hearing functionality than wait and see in several series (Régis et al 2013, Yomo et al 2012, Hillman et al 2010, Akpinar et al 2016). Waiting until the tumour is growing leave time for the hearing to decline. The window of opportunity of hearing preservation is missed when treatment by SRS or MS is delayed until the hearing preservation chances are quite gone. Observing a decline and concluding that observation is better is a sophism.

The patient groups are significantly different. The SRS group is older with worse hearing and bigger lesions, which are known factors of bad prognosis for hearing. The authors write that the hearing status after SRS and MS is equally poor but in fact the patients proposed to SRS have a worse hearing status.

The observed hearing depends on the “audiological aggressiveness” of the disease and the toxicity of the management, which are two different phenomena important to distinguish.

Some additional comments :

1 abstract “Patients with <3 cm VS “: Although traditional, defining a limit based on a 2D max diameter is very crude (specially due to the high variability of posterior fossa size) and Koos classification would in our opinion be more relevant.

2 abstract: “At a mean time of 7.7 years after initial treatment, patients were surveyed via mail with the use of the Hearing Handicap Inventory for Adults (HHIA) and the Tinnitus Handicap Inventory. » The mean is not enough, also the minimum and maximum should be presented.

3 abstract: “Overall, the hearing prognosis was poor, because more than 75% of all patients had non serviceable hearing at the last clinical follow-up “. It would have been good to add in the abstract the quality of the hearing at the time of the primary care.

4 There is no clear information about the quantification of the tumour growth and hearing deterioration during the waiting period. This is a key point.

5 In the method section “Patients that were treated with GKS typically received 12 Gy to the 50% isodose line with a high (>95%) tumour coverage and selectivity. Typically, shots and dosages were deliberately planned so that the facial nerve and cochlea are spared from higher doses.” 12 Gy for patients with a good hearing is slightly high. The doses to the modulus of the cochlea are not reported. Real quantification would be much better (median, min max

prescription dose, dose to the cochlea, etc.). Since this is a historical cohort starting in 1998 the claim of homogeneity of the radiosurgical technic is questionable.

6 The risk of loss to follow up evaluated to 25% in case of wait and see strategy (Bakkouri et al 2009, Hillman et al 2010) should have been discussed.

Paper 3. Audiovestibular Handicap and Quality of Life in Patients With Vestibular Schwannoma and "Excellent" Hearing. Tveiten OV, Carlson ML, Link MJ, Lund-Johansen M. *Neurosurgery*. 2017 Mar 1;80(3):386-392.

1 The term excellent hearing is here used to depict GR1 (Los Angeles A) hearing, which is questionable as long as this class is covering a quite heterogeneous group of patients with level of hearing performance of a quite a large range.

2 In the statement “we identified 296 patients with either bilateral class A (AA) hearing or 1 good ear and 1 deaf ear (AD)”, it is not mentioned at what time point this hearing status refers to. Is it at the first contact, when the kind of management was decided, or at the last follow up of the study? When the patients completed the questionnaire, was an audiometric assessment also made?

3 The authors are quoting mainly the papers supporting their claims but not those reporting data in contradiction with their observations. This is a pity since it could have opened different perspectives of interpretation of the work,

General comments:

Very few data exist about patient-reported audio-vestibular handicap and QoL in patients presenting with Vestibular Schwannomas (VS). In this sense the work of Tveiten and coworkers is of high interest. The paper however is not addressing a specific scientific question and is more a “fishing party style”. In reason of a number of methodological weaknesses no strong conclusion can be drawn from this work. The paragraph “limitation” of the discussion section is listing, on purpose, some of the main weaknesses: the compared groups are significantly different in age, size of tumours medical management, the objective hearing status of the control is unknown, the time between the first symptoms and the evaluation is unknown, the nature of the first symptoms leading to the diagnosis is not reported....

In spite of all these limitations we found the observation of the hearing handicap in class AA patients very interesting. The mean HHI is 3 time higher in AA compared to control with 33% of the patient having a middle or significant handicap. However, here again, these data are to be taken with caution as long as the control group is having a unknown hearing status.

Anyway, this observation is not surprising due to the heterogeneity of class A hearing patients in term of audiometric performances. It is a pity that the authors are not discussing more how to address better in the future this limitation (sub-classification). They don't either quote some interesting papers from the literature like the one from Mousavi entitled: “Hearing sub-classification may predict long-term auditory outcomes after SRS for VS patients with good hearing » (2016 JNS). This paper is showing, in GR1 patients, a dramatically higher chance of long term functional hearing preservation in patients with “subjective perception” of the absence of hearing handicap.

Paper 4. Patient- versus physician-reported facial disability in vestibular schwannoma: an international cross-sectional study. Tveiten ØV, Carlson ML, Goplen F, Myrseth E, Driscoll CLW, Mahesparan R, Link MJ, Lund-Johansen M. *J Neurosurg*. 2017 Nov;127(5):1015-1024.

The candidate is 1st author. The primary objective was to compare patient self-assessment of facial function to physician-documented facial nerve status. A healthy control group was included. In addition, facial nerve disability following treatment via observation, radiosurgery surgery or open surgery was compared. In parallel with paper 1, the less invasive the management, the less the disability. 20% of patients undergoing open surgery had facial dysfunction resulting from the operation, vs. 0-2% in the other two management groups. Self-reported facial dysfunction is comparable to objectively assessed dysfunction in a yes/no fashion,

The stated purpose to compare long-term 7th nerve function following all three types of treatment is fulfilled. It is surprising – but well documented - that physician scored HB degree is not reflected in patient perceived severity of facial disability. Also surprising that general social well-being seems quite unrelated to degree of facial dysfunction. The authors conclude that other elements of 7th nerve function than facial muscle function undetected by the HB score contribute to the patient's feeling of facial disability.

Conclusion and assessment

The present PhD-thesis comprises an overview and four original papers published in international peer-reviewed journals. The thesis is an excellent overview of the literature establishing the relevance of these studies. In the thesis text, the study findings are discussed in a balanced way, both in relation to existing knowledge.

Although VS is one of the less common intracranial tumours, constituting less than 10% of intracranial tumours and reported incidence figures are variable, the number of diagnosed VS is increasing. Adding to this that VS is a benign tumour with almost “unlimited survival” (except in a small percentage of treatment related fatality), the symptom burden of untreated vs. treated tumours becomes extremely important. This is one of the major perspectives and a major relevance of the thesis. As the candidate correctly mentions in the introduction, increasing numbers of persons with very few and/or ambiguous symptoms or other reasons for imaging undergo neuroradiological examinations. This will result in an increasing number of “incidentaloma” VS being diagnosed which further extends the relevance of the studies.

The literature review and the presented results provide a very sound guidance for managing and advising patients with VS and for predicting the statistical average outcome.

Individualized outcome assessments are not possible with the existing material.

The individual studies in the original papers are in the general design of combined retro-anteprospective (ambi-directional) data collection. Strengths and weaknesses of such a design are not discussed. The thesis presents studies conducted with the currently best available methodology using trustworthy data-quality. It would have been very interesting to include in the future perspectives a re-assessment whether “the statistically ideal” randomized control study would indeed be worthwhile by adding more than the currently accumulated evidence.

In conclusion, this is a valuable contribution and the committee finds the thesis worthy to be defended for the degree of PhD. The decision of the evaluation committee is unanimous.

Marseille/Copenhagen/Bergen

9 April 2019

Professor Jean Marie Régis
Aix-Marseille Université

Professor Marianne Juhler
Copenhagen University

Professor Eyvind Rødahl
University of Bergen