Quality of Life after Stapes Surgery

A systematic review

An article-based master thesis

Mai Nayeli Hauge

Spring 2020

UNIVERSITY OF OSLO

Faculty of Educational Sciences

Department of Special Needs Education

June 30, 2020

Quality of Life after stapes surgery — A systematic review

2020

Quality of Life after stapes surgery - A systematic review

Mai Nayeli Hauge

http://www.duo.uio.no

University of Oslo

Abstract

Importance: Studies measuring outcomes after stapes surgery primarily direct their focus on audiometric outcomes. Recently, more studies have focused on measuring Quality of Life (QoL) in addition to audiological outcomes, extending the scope of how stapes surgery affects the patient in more ways than can be measured objectively.

Objective: To collate and evaluate results from individual primary studies on the degree of post-operative relation between audiometric outcomes and QoL after stapes surgery in adult patients with otosclerosis, presented in a systematic review without meta-analysis.

Methods: PubMed, Cochrane Library and OpenGrey were periodically searched for articles in the period January 3 - June 10, 2020. Dates of coverage for all databases were 2015-2020. Titles, abstracts and full texts of retrieved articles were screened for eligibility, on the basis of exclusion criteria illustrated in a PRISMA-style flow diagram. Additionally, article quality was assessed in accordance with GRADE and PRISMA guidelines.

Results: Of 37 articles screened, three articles were eligible for critical appraisal and qualitative synthesis. The three articles assessed showed a very low grade of quality of evidence and low grade of study design. Results suggest that post-operative audiometric outcomes should not be the sole indication for surgical success, patient benefit or QoL. There was inconclusive evidence regarding correlation between different QoL measures and components of audiometric results.

Conclusions and relevance: Disease-specific QoL measures should be developed, translated and validated to be implemented in future clinical studies. Combining objective and subjective measures pre- and post-operatively can provide a clearer picture of the degree of relation between audiometric and QoL outcomes.

Title: Quality of Life after Stapes Surgery – A Systematic Review
Journal: JAMA Otolaryngology: Head and Neck Surgery
URL: https://jamanetwork.com/journals/jamaotolaryngology

Preface

I am eternally grateful for having the opportunity to complete a master's degree at University of Oslo.

Thank you to Björn Lyxell for providing helpful and honest answers.

Thank you to Sissel Pellerud for sharing your clinical knowledge and experience.

Thank you to Ulrika Löfkvist for helping me when I was stuck. Your words of encouragement and wisdom are always appreciated.

Thank you to my friends for your encouragement, moral support and understanding.

Thank you to Nina for sharing your experience and wisdom. Your help has been essential.

Thank you to my dear and long-time friend Madeleine for taking the time to proofread.

Thank you to Thomas for lifting my spirits and believing in me. You've been a wonderful support and partner in crime.

Thank you to my dear father and sister, for believing in me and cheering me on whenever I felt lost.

Lastly I would like to thank my mother, whose wisdom, realistic and positive outlook on life has been an eternal inspiration, and who is the sole reason I chose to study educational audiology. This thesis is for you.

Oslo, June 2020

Mai Nayeli Hauge

Table of contents

Word count for theoretical framework: **5664**

Word count for article manuscript: **2525**

1 Introduction

1.1 Background and rationale for this thesis

The purpose of this review and thesis was to identify, evaluate and summarize all relevant findings of individual studies measuring and evaluating Quality of Life (QoL) (see point 2.3 in chapter 2) before and after stapes surgery in adult patients with otosclerosis. Further, the overall aim was to fill a small but significant gap in the research literature with the purpose of shedding light on which tools that are used to measure QoL in the specific patient group that is the otosclerotic population.

Acquiring a progressive hearing loss affects multiple aspects of life, as the gradual nature of the deterioration that occurs can go undetected for some amount of time before it is noticed. The patient is often led to continuously face new and progressively difficult obstacles and challenges; in everyday life, at work and in social settings, hearing loss could result in compensation mechanisms that eventually lead to fatigue (Hornsby, Naylor, & Bess, 2016).

Such mechanisms can be straining one's neck to improve perception of sound and/or avoiding social activity to avoid continually asking others for repetitions. Additionally, the 'invisibility' of hearing loss, and any cultural stigma associated with the disability and/or use of hearing aids can make it difficult to accept the condition and seek further help (Wallhagen, 2009). On this note, two Norwegian studies have found evidence suggesting that factors relating to hearing loss and fatigue can lead to prolonged sick leave and/or disability retirement in the work place (Svinndal, Jensen, & Rise, 2020; Svinndal, Solheim, Rise, & Jensen, 2018).

The majority of studies concerning the otosclerotic population with regards to pre- and post-operative follow-up have mainly focused on objectively measurable outcomes, such as audiometric outcomes. This is understandable with respect to validity and reliability. However, along with furthering the development of medical practice and follow-up procedures, it would be interesting to see the measurements of QoL gain more traction in future studies concerning stapes surgery. The patient perspective should receive more focus, even though maintenance and development of the technical aspects of surgery are essential. It is, after all, the patient who will be impacted and affected by the surgical intervention.

However short or long, a rehabilitation process follows surgery, and that process is experienced first-hand by the patient. Thus, in terms of success rates of surgical procedures, patient focus and satisfaction should be included and weighted when looking at contributing

factors for successful surgery (Fayers & Machin, 2016). Whatever deems surgery to be successful can indeed be gaged by the surgeon's technical skills, but without the patient there is no surgery. Therefore, the patient's QoL should be an integral part of the evaluation process, as audiometric outcomes could prove dissimilar and uncorrelated from the patient's reported experience of QoL (Weiss et al., 2019).

Seeing as there are currently no reviews observed concerning QoL after stapes surgery, I have chosen to explore this area of research through the method of a systematic review (without meta-analysis). The systematic review might fill a small but significant gap in the research literature concerning QoL after stapes surgery.

1.2 Thesis structure

This is an article-based master thesis, and is structurally divided into two parts: The first part, the theoretical framework of this thesis, provides the theoretical background and methodological considerations for the research, and the second part is the article manuscript. The intent behind choosing an article-based master thesis was to conduct research that will be sent to a peer-reviewed journal, and hopefully published. Consequently, as an author, I have the opportunity to reach a wider audience, and my research will be reviewed by other skillful research peers. The ultimate goal is to contribute with enlightening and useful research that can make a difference for future research and in clinical practice.

As described by Khan, Kunz, Kleijnen & Antes (2003), the chosen structure for this thesis will reflect the steps taken in conducting a systematic review (without meta-analysis), and are as follows: The first part of this thesis, namely the theoretical framework, will describe the framing of the theme (QoL after stapes surgery), where the first chapter herein provides a rationale and theoretical background for choosing this theme in particular. The second chapter will identify the relevant work, by shedding light on relevant research already conducted in the field of otosclerosis and stapes surgery, as well as QoL. This will provide a necessary theoretical backdrop for the discussion in the article manuscript. The third chapter is where any and all methodological considerations and issues are described, in addition to a detailed description of exclusion criteria, which is an important step in assessment of quality. The fourth and final chapter will include a short conclusion and thoughts on further implications for future research. Qualitative synthesis and appraisal of findings will be described in the second part of this thesis, the article manuscript.

The chosen journal is JAMA Otolaryngology – Head and Neck Surgery (URL: https://jamanetwork.com/journals/jamaotolaryngology). This journal has an impact factor of 3.5 and is, according to their website, the highest ranking otolaryngology journal in the world.

The reasoning behind choosing this esteemed journal were the benefits that come with this being a highly respected journal as well as a journal providing established article instructions (see Appendix 1), thus allowing for an opportunity to conduct research that is both valuable and necessary to the emerging field of QoL after stapes surgery, as well as having said research reviewed by esteemed peers who are specialized in the particular field.

2 Theoretical and empirical background

2.1 Hearing loss

According to Stach (2009), hearing loss is the term describing the ear's lack of ability to conduct airwaves to the inner ear, and/or the inner ear's inability to produce electric signals to the brain. Hearing loss is commonly divided into two main types, conductive and sensorineural. Conductive hearing loss is determined by any physiological hindrance in the outer or middle ear, such as atrophy of the ear canal, excessive cerumen or fluid in the middle ear. Hearing loss is categorized as sensorineural when the inner ear, most notably the cochlea and its hair cells are damaged, or fail to produce electric signals to the brain from the air waves passed on through the outer and middle ear. In some cases the hearing function works normally, yet the individual is still classified as suffering from hearing loss, as the brain is unable to deduce electric signals into what the individual interprets as meaningful sound. Hearing loss can be congenital, meaning the individual is born with it (or it manifest in very early childhood), or it can be acquired later in childhood or in adult age. For some acquired types of hearing loss, the loss is a progressive one that often takes its time in manifesting to the point of notice. One such type of hearing loss is the middle ear disease called otosclerosis.

2.1.1 Otosclerosis

Otosclerosis is a middle-ear disease, which gradually leads to new bone growth on the stapes, and can thus fixate the stapes in the oval window over time. This leads to a conductive hearing loss, although in progressive cases (also termed advanced otosclerosis), the bone growth permeates and damages the outermost part of the cochlea and the hair cells, thus determining the hearing loss as mixed (Heining, Banga, Irving, Coulson, & Monksfield, 2017). Otosclerosis is a common reason behind hearing loss in adults, but research concerning the specific genetic pathophysiology of otosclerosis remains inconclusive. Gentil et al. (2016) has found evidence of otosclerotic prevalence in women, as it can manifest during or after pregnancy, likely attesting to the hormonal changes that occur. Even though otosclerosis has been genetically linked, it occurs sporadically in roughly half the cases, although these areas of research on otosclerosis remain somewhat inconclusive (Ealy & Smith, 2011).

2.1.2 Tinnitus

Many otosclerosis patients suffer from the additional burden of chronic tinnitus, and many choose stapes surgery in an effort to tackle it (Danesh, Shahnaz, & Hall, 2018). According to Baguley, Andersson, McFerran & McKenna (2012), tinnitus is termed a subjective (albeit in rare cases, objective) perception of sound without the presence of external stimuli, and is often a symptom of hearing loss. Tinnitus can manifest as multiple sounds, in some cases at the same time, unilaterally or bilaterally. The sound characteristics may vary from a low frequency humming, a high frequency pure tone, a swishing sound or a crackling sound. Often dependent on the individual's physical and mental state, it can be perceived as stronger or louder. This can also be the case if the individual has been subjected to loud noise (Baguley et al., 2012).

Ismi et al. (2017) have found evidence that the estimated incidence of tinnitus in the otosclerotic population is 56-84,5%, which points to tinnitus as being a symptom of the hearing loss that otosclerosis patients experience. On the same note, chronic tinnitus could be the reason why many opt for surgical intervention, as it has been shown that a decrease in tinnitus improves QoL for otosclerosis patients (Drexler et al., 2016). This point is furthered by a study by Skarzynski et al. (2018), who note that perceived changes in tinnitus are associated with comorbid hearing difficulties, and importantly, that otosclerosis patients are not representative of the tinnitus population as a whole.

2.2 Stapes surgery

Although many new advances and surgical techniques are emerging in the field of stapes surgery (e.g. laser and robot-assisted surgery), the two most established and commonly used surgical techniques are stapedectomy and stapedotomy (Balu, Kumar, Nair, John, & Sreekumar, 2019; Sedwick, Louden, & Shelton, 1997). However, there is evidence found that stapedotomy yields better surgical outcomes (Cheng, Agrawal, & Parnes, 2018). Both surgical procedures aim to facilitate airwave conduction, and the main difference between the two techniques is in how much of the stapes footplate is removed and replaced by a prosthesis. In stapedectomy the entire footplate is removed, and in stapedotomy only a limited hole in the stapes footplate is produced.

2.2.1 Success and complication rates

In surgery, the aim is to achieve a reduced air-bone gap (ABG), which is the difference in air and bone conduction in hearing thresholds. As previously mentioned, in most

cases of otosclerosis the hearing loss is conductive, meaning the bone conduction thresholds are at a functional level. Generally the desired audiometric outcome after stapes surgery is a reduced ABG (<10 dB) (Weiss et al., 2019). Many studies have reported high success rates regardless of surgical method and type of prosthesis, however, few studies that focused on complication risks and follow-up procedures have been found. This could be attested to the fact that complications during and following stapes surgery is relatively uncommon (Antonelli, 2018).

Although stapes surgery is an established and safe form of surgical intervention for the otosclerotic population, certain complications during or following surgery can occur. Due to the frontal placement of the corda tympani, maneuvering surgical tools around it can cause damage to the nerve, leading to dysgeusia, meaning loss or distortion of the taste sense (Guder, Böttcher, Pau, & Just, 2012). Other complications are tinnitus (more elevated than the pre-operative stage), vertigo or hearing loss (Mohseni, Daneshi, Asghari, Mohebi, & Moradi, 2019). However, ensuring patient follow-up and care, there are methods for retraining the taste sense and cognitive retraining therapy for those suffering from tinnitus (Jastreboff, 2015).

2.2.2 Hearing aids and pre-operative evaluation

Not all cases of otosclerosis are treated surgically, and many opt for hearing aids instead. A common reason behind this is the progressive nature of the disease, and many will need revision surgery years after primary surgery (Eshraghi & Telischi, 2018). Generally, if the patient with otosclerosis is overall healthy, surgical treatment will be considered once they have had a pre-operative evaluation (Antonelli, 2018). Due to potential risks associated with an invasive surgical procedure, in combination with an evaluation of each patient's health, stapes surgery may not be a treatment alternative for everyone.

Additionally, practice guidelines for preoperative evaluation, care and follow-up can differ between countries and clinics, especially with respect to the patient's age, health, degree of hearing loss and perceived difficulties. As McElveen & Kutz (2018) note, in the field of pre-operative evaluation and post-operative care for otosclerosis patients, there are varying opinions as to how to conduct this optimally. An example worthy of mention here, are the national guidelines for hearing aids in Norway, where they are covered by the Norwegian National Insurance Act, with an upper cost limit (forskrift om stønad til

høreapparat og tinnitusmaskerer, 1997). Due to this variation in guidelines and pre-operative evaluation, these two components will not be evaluated further.

2.3 Quality of life

Within the scope of this review the abbreviation QoL was chosen to denominate the term Quality of Life. QoL is a broad definition, and many studies use other and more specific sub-terms such as Health Related Quality of Life (HRQoL), patient-reported outcomes and patient satisfaction (Fayers & Machin, 2016). Nonetheless, it has become evident that researchers view QoL as a valuable source of outcome in health-related research, despite lacking a common interpretation of a term which is already ill defined (Bowling, 2019; Fayers & Machin, 2016).

Ring (2007) quotes Socrates in saying that value should not be placed on living, but on living well. The most salient thing about QoL is that it is individual at the core for each person, and that makes it a difficult term to define in detail and thus, quantify and measure in a cohort. However, the relevancy of QoL in clinical trials has emerged as more prominent in the medical field, as researchers and practitioners view QoL as a holistic approach that sheds light on the psychosocial aspect of health; the acceptance of QoL as a valuable measure can lead to enhanced patient-information and care, as well as clinical decision-making. However, for this to happen the QoL measurement tools need to be reliable and sensitive to change (Ring, 2007).

2.3.1 Response shift and QoL psychometrics

What is important to consider when assessing QoL is something called response shift. According to Ring (2007), this term refers to someone's altered assessment of an idea or concept based on a change of internal values. To exemplify this she points to the fact that in terms of measured QoL, people with severe diseases view their own QoL as better than what is measured in people who are healthy. Further, she states that a response shift can have a masking effect on treatment outcomes, especially when measuring QoL, yet the response shift itself could be viewed as a sign of a patient's ability to adapt to a new way of life and its challenges (Ring, 2007). Similarly, the use of questions of symptoms or challenges in a QoL measurement, poses the belief that these components do have an effect (Fayers & Machin, 2016).

Schwartz & Rapkin (2004) argue that the appraisal process of QoL psychometrics has been wrongly applied, as there is no such thing as a "true" score for QoL because the phenomena related to it are not susceptible to nomothetic measurement. The problematization of the psychometric appraisal can be viewed in the same vein as the core notion of why measuring QoL is still complicated; subjectivity, feelings, sense of self and individual life experiences are abstract notions to anyone but the patient who is experiencing them, and to whom they remain intrinsically personal. On this note, Ring (2007) describes how previous measures were based on healthcare workers' assessment of patient QoL, although it is currently accepted that patients are in the best position to evaluate their own QoL. However, the possible outcome impact of questionnaires is relevant, as their questions are formulated by investigators rather than patients, perhaps as a way of circumventing the issue of a lack of a QoL definition (Fayers & Machin, 2016).

2.3.2 Measuring QoL

The challenge in clinically assessing a patient's QoL lies in the term's inherent subjectivity. Quality of life cannot be measured with easily applicable, objective criterion, yet it functions as an important goalpost when assessing the effects of a surgical intervention. The reason behind can perhaps be attributed to the fundamental notion about personal wellbeing being different across people and cultures (Fayers & Machin, 2016). Whilst objectively measurable components are an obvious and essential tool to gage the effect of an intervention (e.g. measuring audiometric results before and after stapes surgery), considering the broader effect on the patient's life should be equally important.

Developing a successful surgical technique and/or practice and ensuring a desired effect for each patient is of great importance, yet the two have seldom received equal focus in terms of research in the field of QoL after stapes surgery. More often than not the main focus lies on the objectively measurable, such as audiometric outcomes, however, favoring objective outcomes has been the modus operandi in medical research for a long time (Ring, 2007). This is to a certain degree understandable, as the objective measures and outcomes are less invasive and more easily interpreted before and after surgical intervention.

A surgical intervention such as stapes surgery is serious and does not come without risks of complications, and could therefore have a considerable impact—positive, negative or both—on a patient's QoL, both short-term and long-term. In a study on hearing outcomes after stapes surgery, Andersen, Ohman, & Sorensen (2015) argued that changes in audiometric outcomes between a short-term and long-term follow-up were clinically

insignificant. However, few patient-centered longitudinal follow-up studies have been conducted on QoL after stapes surgery, perhaps because many studies have been seen to only include an audiometric follow-up measurement <1 year post-operatively.

Considering the fast development in stapes surgery and the medical field in general, longitudinal studies with aged data (>10 years) provide an informative insight to long-term outcomes, yet one could raise the question of how much reliance can be placed upon studies of such design; Schwartz & Sprangers (1999) point to response shift not being accounted for in modern QoL measurement tools. Thus, factors relating to intrapersonal change, cultural change and interventional change can be seen as necessary to control for in any interpretation of QoL outcomes in future research (Bowling, 2019; Fayers & Machin, 2016).

2.3.3 QoL measurement tools

According to Bowling (2019) there exist a multitude of disease-specific QoL measures, although the level of standardization is below par. In the field of QoL after stapes surgery there has been observed some variation in QoL measurement tools used in addition to audiometric outcomes. There are especially three noteworthy QoL measurement tools to consider in this instance, as they are used to measure QoL in the articles to be reviewed. They are as follows: Tinnitus Functional Index (TFI); Glasgow Benefit Inventory (GBI); Stapes-Plasty-Outcome-Test (SPOT-25).

TFI is a validated questionnaire which focuses on diagnosing the impact of tinnitus into quantifiable measures. A UK study by Fackrell, Hall, Barry, & Hoare (2018) found the range of psychometric properties satisfying in terms of covering symptom domains, as well as good reliability. As mentioned previously (see point 2.1.2), many people with otosclerosis suffer from tinnitus, and TFI is very applicable for those studies focusing on the part of the otosclerotic population who have received surgical intervention because of tinnitus, or for the part who finds their tinnitus exacerbated by it.

GBI is a validated questionnaire specific to otorhinolaryngological surgical interventions, measuring patient benefit post-surgery (Robinson, Gatehouse, & Browning, 1996), and questions formulate the intervention as such, an otorhinolaryngological surgical intervention in general. The GBI questionnaire has been validated in Swedish also; a study by Redfors, Jonsson, Tideholnn, & Finizia (2019) found the Swedish version of GBI to be satisfactory in terms of psychometric properties.

SPOT-25 is a recent German addition to the plethora of QoL measurement tools (the term used for this tool is HRQoL), and is the first disease-specific measure to be validated for otosclerosis and stapes surgery (Lailach et al., 2017). The validation process has shown a high internal consistency, and the measurement tool can discriminate between healthy subjects and otosclerosis patients. However, being a recent addition it has yet to be validated in English, but is in the process of being validated in Dutch (Blijleven, Thomeer, Stokroos, & Wegner, 2019).

3 Method

3.1 Systematic review without meta-analysis

A way to phrase the meaning of systematic reviews is ‘to know what is known’ (Gough, Oliver, & Thomas, 2017). Conducting a systematic review means to compile relevant literature, with the aim to provide a more comprehensive and reliant body of evidence with which to answer a research question or achieve an overview of something specific. The need for systematic reviews, or perhaps all types of literature reviews, has arisen from the challenge of navigating through the ever expanding plethora of research which has been made readily available through numerous databases across the worldwide web.

In order to continually evolve and develop new areas of research and knowledge, it is evident that one needs to gage which way to go forward. As a means of attaining new depths of knowledge, and developing new and interesting hypotheses, it is necessary to find research holes and aim to fill them. To achieve this goal, it all depends on knowledge about and thorough research on what has previously been done on a research topic, and if any gaps exist in the literature. Conducting a systematic review is one such method of research.

Per author instructions from the chosen journal (JAMA Otolaryngology - Head and Neck Surgery), this systematic review will not include a meta-analysis, and is therefore termed a qualitative or narrative review which will make inferential claims based on theory rather than statistical analysis (Gough et al., 2017).

Gough et al. (2017) points to the framework in qualitative reviews allowing for a more open approach to themes or research questions, emergent concepts, less formal procedures, theoretical inference and an aim to impact on the basis of enlightenment. It must be stated here that the qualitative term does not mean that the synthesis of the review (in the article manuscript) is based purely on qualitative data input. All studies included in the review have results based on numerical data, however, the synthesis of the review is of a qualitative nature.

3.2 Theoretical considerations

In writing this thesis there were a number of limitations considered, with the aim to produce a valuable review. Firstly, the issue of including other reviews in this thesis was thought of at length. Reviews offer valuable overviews on relevant research, yet my aim with

this systematic review was to get as close to the empirical reality as possible. Consequently, there is to a certain degree a risk of bias in reviews (perhaps from the primary studies included), and I deliberately sought to avoid risk of bias as much as possible. This review did also not evaluate other reviews due to the apparent lack of reviews concerning QoL after stapes surgery in the adult otosclerotic population, which was also a reason for choosing to conduct a systematic review as a research method for this thesis.

A deliberate exclusion in this review was the issue of cost-efficiency of treating otosclerosis surgically. Within cost there lie many factors, such as national healthcare and benefits associated with it (Casazza et al., 2019). Due to the fact that the articles that were reviewed are country or area specific, factoring in healthcare would make for a complicated variable, although factors relating to cost are and will continue to be important for future research. Thus, the exclusion of cost will hopefully make this research more relevant on an international level.

The second limitation was the exclusion of studies concerning newer and less established surgical techniques, such as laser and robot-assisted. Surgical practice can vary across countries and among surgeons, as their practice changes and improves with time and experience (Nguyen, Bernardeschi, & Sterkers, 2018). Therefore, opting for the two most common and well established surgical techniques, namely stapedectomy and stapedotomy, would contribute to research that is generalizable yet sufficiently specific. The same principle as mentioned above applies to studies concerning stapes prosthesis, as these studies were too technically specific to be included in this thesis, which aims to focus on patient-related studies and outcomes.

The otosclerotic population is also highly heterogenous, especially with respect to age and disease trajectory (Van Den Bogaert et al., 2002). In some cases otosclerosis will manifest in children, and diagnostic procedures and clinical evaluation of the disease may differ from those concerning adults. Consequently, the process of deciding when surgery is necessary is different in each country, as previously mentioned, and therefore studies concerning juvenile otosclerosis were not included in the review.

With respect to disease trajectory, there were several issues to consider. Due to the progressive nature of otosclerosis as a disease and the heterogeneity of the otosclerotic population, each case is, as with most disease-specific issues, slightly different in each patient. A patient will sometimes need revision surgery due to recurring hearing loss and/or failure of primary surgery (Gros, Vatovec, Zargi, & Jenko, 2005). Therefore, studies with only revision cases were not included. On a similar note, studies concerning far advanced

cases of otosclerosis and mixed hearing loss were not included, as this raises the issue of cochlear implants (Abdurehim, Lehmann, & Zeitouni, 2016).

During the process of collecting literature and searching for primary studies, the theme of this thesis was narrowed over time. Originally it was planned that the thesis would focus on success rates of primary studies on otosclerosis, looking specifically at study design, hearing thresholds, tinnitus and QoL. After much deliberation, I came to the conclusion that this theme needed a narrower approach, and decided to solely focus on QoL after stapes surgery. Thus the theme would have been more patient-centered and relevant to special needs education and educational audiology. Using this angle still allowed for tinnitus and hearing thresholds to be included, as they are very much relevant, without relinquishing the patient-centered perspective.

3.3 Data search

Data search and collection was done by periodically searching in PubMed, Cochrane Library and OpenGrey, with the last search completed in June 10, 2020. Before starting the article search process, I sought and received guidance by a certified librarian at University of Oslo's Main Library, who helped me in regards to search string production and database search processes. To be able to conduct a search in a database that would lead to all (or as many as possible) relevant articles it is important to produce a search string that encompasses as many synonyms as possible. Researchers can use different terms to denote similar or identical objects/goals/outcomes. One such example is the term stapes surgery, which can be named stapedial surgery, stapesplasty, stapedotomy or stapedectomy, to name the most common ones.

The search string applied in every database search is as follows: otosclerosis AND quality of life AND (qol OR hrqol OR stapes OR stapedial OR stape* OR outcome* OR result* OR tinnitus OR hearing OR surgery OR surgeries OR operation OR operations OR stapedotomy* OR stapedectom* OR stapedotomy OR stapedectomy OR stapesplasty OR outcome* OR satisfaction). The search process was conducted six times in the period from January 3 to June 10, 2020, to avoid missing new and relevant articles that may have been published since the first database search was conducted.

As is customary in systematic reviews, searching multiple databases is necessary to minimize publication bias. The database PubMed contains citations from a plethora of

journals, and seemed a good option to start the search process. Cochrane Library is esteemed in terms of quality of content, and was a natural addition to my selection of databases. Additionally, as is important when conducting a systematic reviews, is to search the gray literature. Again, to minimize publication bias, it is essential to search for content that may not have been published due to an ongoing research process or a lack of "sufficient" effect size (Booth, Sutton, & Papaioannou, 2016).

3.3.1 Exclusion criteria

The database search range in terms of years was set to 2015-2020. For a systematic review this seems quite a short range, and the reason behind was the decision to exclude studies with data aged more than 10 years. Due to the development in the surgical field of stapes surgery, comparing data from 35 years ago with data from last year does would probably make for a review with mostly incomparable data. Setting the lower year limit to 2015 would allow the search to find studies with data from at most 5 years past (2010).

I made the conscious decision to exclude retrospective studies from the review articles. Many studies are found to apply a method of retrospectively analyzing data, which all together does not come without its issues. Useful and practical as it may be to apply analyzes to existing data sets (perhaps owing to a lack thereof), this method could lend itself to hypothesizing on the basis of an apparent trend, rather than hypothesizing before data collection and appraisal, as is the case in prospective studies. I did not want to risk skewing the result of the review by using retrospective studies.

With respect to language, any studies in languages other than English, Norwegian, Swedish and Danish were excluded. Only a small handful of studies in other languages with abstracts in English were observed in the search results, but including them on the basis of an abstract and/or recounted from a secondary source would entail a risk of bias that is too great.

3.4 Ethical implications

Conducting research that involves personal data comes with a great responsibility. Even in systematic reviews, where the data consists of data collected from other researchers, there is a need to ascertain the ethics applied in the collection of primary data used in the review are taken into account. The systematic review conducted here has only collected secondary data, and each article selected for critical appraisal has stated that study design and

conduct is in accordance with the Declaration of Helsinki (Dziendziel, Skarzynski, Gos, & Skarzynski, 2019; Lailach et al., 2018; Weiss, Schuldt, Oberhoffner, Ginzkey, & Mlynski, 2020).

3.5 Methodological considerations

3.5.1 Qualitative validity

When conducting research, especially in the realm of subjective measurements such as QoL, validity and reliability are of utmost importance; to produce applicable and clinically useful results in a systematic review, certain basic properties need to be satisfied. The first one is validity, which is relevant for this thesis.

In the overall sense, validity is the property that describes how well the measurement captures what is supposed to be measured. Applying this in a review poses the challenge of ensuring good validity in the studies chosen for the review. Additionally, as the author of the review, other steps need to be taken to ensure validity of the review itself. This is called qualitative validity, and is first and foremost accomplished by conducting thorough examination of the findings in the selected literature to be reviewed. Checking the accuracy of the findings in the selected studies can be operationalized through a set of procedures, with the aim of honing in to the researcher's assessable ability, and in turn, produce research that is convincing to the readers (Creswell & Creswell, 2018).

There are two procedures that need to be highlighted here. First and foremost, it is important to evaluate the results in each article in the light of risk of bias (Booth et al., 2016). To consider if study design, conduct, subject selection or measurement tools are used in any way distort or influence the analysis of the results, is essential. Secondly, and equally important, is to look for confounding elements, where one needs to assess the intervariable effect, and ascertain whether the articles to be reviewed account for any confounding variables at an early stage. According to Booth et al. (2016), a common example of such variables can be age, ethnicity and sex.

3.5.2 Qualitative reliability

To gage the trustworthiness and strengths of the results in the articles to be reviewed is important, and is what the term reliability refers to. Analyzing results that possibly happened due to chance is something which should be avoided, and one way to check for reliability is

by whether the study design allows for repeated testing by others (Booth et al., 2016). However, as this systematic review is conducted without a meta-analysis, which would entail an assessment of numerical and statistical data in the studies, the reliability in this instance relates more to the qualitative part of the analysis, and in checking if any results are ignored, distorted or exaggerated. What is a challenge, however, is assessing the generalizability of the results (Creswell & Creswell, 2018). What needs to be checked for generalizability is the study design, tools and procedures.

4 Conclusion and further implications

In writing this thesis and review I have evaluated three articles concerning QoL after stapes surgery. A limitation is the small selection of articles, however, few articles met the strict inclusion criteria. Findings show that few studies have used SPOT-25, despite being the most specific and QoL measuring tool for the otosclerotic population to date. However, it is a relatively new QoL measurement tool. With this I would like to encourage authors of future studies to translate and validate SPOT-25 in English, with the aim of implementing it when measuring QoL in the otosclerotic population after stapes surgery.

Considering studies which have hitherto used more general measuring tools, the findings of specific QoL aspects might be less reliable and generalizable than findings from studies using SPOT-25. On the question of how a surgical intervention such as stapes surgery impacts the patient's QoL, the qualitative analysis from the review indicates that studies would benefit from raising the question of how much reliance can be placed on objective measures and overly general QoL measurement tools. Seeing as there is low correlation between audiometric outcomes and the patient's QoL post-operatively, it reinforces the notion that audiometric outcomes need to be seen in the context of the patient's QoL, as measured before and after stapes surgery, both short-term and long-term.

References

Abdurehim, Y., Lehmann, A., & Zeitouni, A. G. (2016). Stapedotomy vs Cochlear Implantation for Advanced Otosclerosis: Systematic Review and Meta-analysis. *Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery, 155*(5), 764-770. doi:10.1177/0194599816655310

Andersen, S. A. W., Ohman, M. C., & Sorensen, M. S. (2015). The stability of short-term hearing outcome after stapedotomy: a prospective database study. *Acta oto-laryngologica, 135*(9), 871-879. doi:10.3109/00016489.2015.1035402

Antonelli, P. J. (2018). Prevention and Management of Complications in Otosclerosis Surgery. *Otolaryngologic Clinics of North America, 51*(2), 453-462. doi:https://doi.org/10.1016/j.otc.2017.11.015

Baguley, D., Andersson, G., McFerran, D., & McKenna, L. (2012). *Tinnitus: A Multidisciplinary Approach*(2. Aufl. ed.).

Balu, R., Kumar, M., Nair, P. G., John, S. R., & Sreekumar, S. (2019). Small Fenestra Stapedotomy Versus Large Fenestra Stapedectomy in Improving Hearing Ability in Patients with Otosclerosis: Our 10 Years Experience. *Indian J Otolaryngol Head Neck Surg, 71*(3), 304-308. doi:10.1007/s12070-019-01710-1

Blijleven, E. E., Thomeer, H., Stokroos, R., & Wegner, I. (2019). Protocol for a validation study of the translated stapesplasty outcome test 25 for measurement of disease-specific quality of life in Dutch patients with otosclerosis. *BMJ Open, 9*(12), e030219. doi:10.1136/bmjopen-2019-030219

Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic Approaches to a Successful Literature Review*: Sage.

Bowling, A. (2019). Health Status and Quality of Life Assessment. In C. D. Llewellyn, C. McManus, J. Weinman, K. J. Petrie, S. Newman, S. Ayers, & T. A. Revenson (Eds.), *Cambridge Handbook of Psychology, Health and Medicine* (3 ed., pp. 193-195). Cambridge: Cambridge University Press.

Casazza, G. C., Thomas, A. J., Dewey, J., Gurgel, R. K., Shelton, C., & Meier, J. D. (2019). Variations in Stapes Surgery Cost within a Multihospital Network. *Otolaryngol Head Neck Surg, 161*(5), 835-841. doi:10.1177/0194599819855055

Cheng, H. C. S., Agrawal, S. K., & Parnes, L. S. (2018). Stapedectomy Versus Stapedotomy. *Otolaryngol Clin North Am, 51*(2), 375-392. doi:10.1016/j.otc.2017.11.008

Creswell, J. W., & Creswell, J. D. (2018). *Research design : qualitative, quantitative & mixed methods approaches* (5th edition. ed.). Los Angeles, California: Sage.

Danesh, A. A., Shahnaz, N., & Hall, J. W., 3rd. (2018). The Audiology of Otosclerosis. *Otolaryngol Clin North Am, 51*(2), 327-342. doi:10.1016/j.otc.2017.11.007

Drexler, D., López-Paullier, M., Rodio, S., González, M., Geisinger, D., & Pedemonte, M. (2016). Impact of reduction of tinnitus intensity on patients' quality of life. *Int J Audiol, 55*(1), 11-19. doi:10.3109/14992027.2015.1072772

Dziendziel, B., Skarzynski, H., Gos, E., & Skarzynski, P. H. (2019). Changes in Hearing Threshold and Tinnitus Severity after Stapes Surgery: Which Is More Important to the Patient's Quality of Life? *ORL J Otorhinolaryngol Relat Spec, 81*(4), 224-233. doi:10.1159/000500992

Ealy, M., & Smith, R. J. H. (2011). Otosclerosis. *Adv Otorhinolaryngol, 70*, 122-129. doi:10.1159/000322488

Eshraghi, A. A., & Telischi, F. F. (2018). Otosclerosis and Stapes Surgery. *Otolaryngol Clin North Am, 51*(2), xvii-xix. doi:10.1016/j.otc.2017.12.001

Fackrell, K., Hall, D. A., Barry, J. G., & Hoare, D. J. (2018). Performance of the Tinnitus Functional Index as a diagnostic instrument in a UK clinical population. *Hear Res, 358*, 74-85. doi:10.1016/j.heares.2017.10.016

Fayers, P. M., & Machin, D. (2016). *Quality of life : the assessment, analysis, and reporting of patient-reported outcomes* (3rd ed. ed.). Chichester , West Sussex: John Wiley.

forskrift om stønad til høreapparat og tinnitusmaskerer. (1997). *forskrift om stønad til høreapparat og tinnitusmaskerer (LOV-1997-02-28-19-§10-7)*. Retrieved from https://lovdata.no/dokument/SF/forskrift/1997-04-18-337?q=h%C3%B8reapparat

Gentil, E., Santos, C., Parente, M., Martins, P., Natal, R., & Almeida, E. (2016). *Otosclerosis-ear disease in women* (R. N. Jorge, T. Mascarenhas, J. A. Duarte, I. Ramos, M. E. Costa, M. H. Figueiral, O. Pinho, S. Brandao, T. DaRoza, & J. Tavares Eds.).

Gough, D. A., Oliver, S., & Thomas, J. (Eds.). (2017). *An introduction to systematic reviews* (Second edition. ed.). Los Angeles: SAGE.

Gros, A., Vatovec, J., Zargi, M., & Jenko, K. (2005). Success rate in revision stapes surgery for otosclerosis. *Otol Neurotol, 26*(6), 1143-1148. doi:10.1097/01.mao.0000172414.64907.9d

Guder, E., Böttcher, A., Pau, H. W., & Just, T. (2012). Taste function after stapes surgery. *Auris Nasus Larynx, 39*(6), 562-566. doi:10.1016/j.anl.2011.10.020

Heining, C., Banga, R., Irving, R., Coulson, C., & Monksfield, P. (2017). Audiological outcome of stapes surgery for far advanced cochlear otosclerosis. *J Laryngol Otol, 131*(11), 961-964. doi:10.1017/s0022215117001815

Hornsby, B. W., Naylor, G., & Bess, F. H. (2016). A Taxonomy of Fatigue Concepts and Their Relation to Hearing Loss. *Ear Hear, 37 Suppl 1*(Suppl 1), 136s-144s. doi:10.1097/aud.0000000000000289

Ismi, O., Erdogan, O., Yesilova, M., Ozcan, C., Ovla, D., & Gorur, K. (2017). Does stapes surgery improve tinnitus in patients with otosclerosis? *Braz J Otorhinolaryngol, 83*(5), 568-573. doi:10.1016/j.bjorl.2016.07.001

Jastreboff, P. J. (2015). 25 years of tinnitus retraining therapy. *Hno, 63*(4), 307-311. doi:10.1007/s00106-014-2979-1

Khan, K. S., Kunz, R., Kleijnen, J., & Antes, G. (2003). Five steps to conducting a systematic review. *J R Soc Med, 96*(3), 118-121. doi:10.1258/jrsm.96.3.118

Lailach, S., Schenke, T., Baumann, I., Walter, H., Praetorius, M., Beleites, T., . . . Neudert, M. (2017). Development and validation of the Stapesplasty Outcome Test 25 (SPOT-25). *Hno, 65*(12), 973-980. doi:10.1007/s00106-017-0389-x

Lailach, S., Schenke, T., Baumann, I., Walter, H., Praetorius, M., Beleites, T., . . . Neudert, M. (2018). Living with otosclerosis: disease-specific health-related quality-of-life measurement in patients undergoing stapes surgery. *Eur Arch Otorhinolaryngol, 275*(1), 71-79. doi:10.1007/s00405-017-4798-y

McElveen, J. T., Jr., & Kutz, J. W., Jr. (2018). Controversies in the Evaluation and Management of Otosclerosis. *Otolaryngol Clin North Am, 51*(2), 487-499. doi:10.1016/j.otc.2017.11.017

Mohseni, M., Daneshi, A., Asghari, A. M., Mohebi, S., & Moradi, S. (2019). Determining the Incidence of Complications and Their Causes after Stapedotomy in Patients with Otosclerosis. *Journal of Pharmaceutical Research International, 28*(6), UNSP-50453. doi:10.9734/JPRI/2019/v28i630222

Nguyen, Y., Bernardeschi, D., & Sterkers, O. (2018). Potential of Robot-Based Surgery for Otosclerosis Surgery. *Otolaryngol Clin North Am, 51*(2), 475-485. doi:10.1016/j.otc.2017.11.016

Redfors, Y. D., Jonsson, R., Tideholnn, B., & Finizia, C. (2019). Psychometric properties of the Swedish version of the Glasgow Benefit Inventory in otosclerosis subjects. *Laryngoscope Investigative Otolaryngology, 4*(6), 673-677. doi:10.1002/lio2.320

Ring, L. (2007). Quality of life. In A. Baum, C. McManus, J. Weinman, K. Wallston, R. West, S. Newman, & S. Ayers (Eds.), *Cambridge Handbook of Psychology, Health and Medicine* (2 ed., pp. 178-182). Cambridge: Cambridge University Press.

Robinson, K., Gatehouse, S., & Browning, G. G. (1996). Measuring patient benefit from otorhinolaryngological surgery and therapy. *Ann Otol Rhinol Laryngol, 105*(6), 415-422. doi:10.1177/000348949610500601

Schwartz, C. E., & Rapkin, B. D. (2004). Reconsidering the psychometrics of quality of life assessment in light of response shift and appraisal. *Health and Quality of Life Outcomes, 2*(1), 16. doi:10.1186/1477-7525-2-16

Schwartz, C. E., & Sprangers, M. A. (1999). Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Soc Sci Med, 48*(11), 1531-1548. doi:10.1016/s0277-9536(99)00047-7

Sedwick, J. D., Louden, C. L., & Shelton, C. (1997). Stapedectomy vs stapedotomy. Do you really need a laser? *Arch Otolaryngol Head Neck Surg, 123*(2), 177-180. doi:10.1001/archotol.1997.01900020059008

Skarzynski, H., Gos, E., Dziendziel, B., Raj-Koziak, D., Wlodarczyk, E. A., & Skarzynski, P. H. (2018). Clinically important change in tinnitus sensation after stapedotomy. *Health Qual Life Outcomes, 16*(1), 208. doi:10.1186/s12955-018-1037-1

Stach, B. A. (2009). *Clinical audiology : an introduction* (2nd ed. ed.). New York: Delmar - Cengage Learning.

Svinndal, E. V., Jensen, C., & Rise, M. B. (2020). Working life trajectories with hearing impairment. *Disabil Rehabil, 42*(2), 190-200. doi:10.1080/09638288.2018.1495273

Svinndal, E. V., Solheim, J., Rise, M. B., & Jensen, C. (2018). Hearing loss and work participation: a cross-sectional study in Norway. *Int J Audiol, 57*(9), 646-656. doi:10.1080/14992027.2018.1464216

Van Den Bogaert, K., Govaerts, P. J., De Leenheer, E. M., Schatteman, I., Verstreken, M., Chen, W., . . . Van Camp, G. (2002). Otosclerosis: a genetically heterogeneous disease involving at least three different genes. *Bone, 30*(4), 624-630. doi:10.1016/s8756-3282(02)00679-8

Wallhagen, M. I. (2009). The Stigma of Hearing Loss. *The Gerontologist, 50*(1), 66-75. doi:10.1093/geront/gnp107

Weiss, N. M., Schuldt, S., Grossmann, W., Oberhoffner, T., Ginzkey, C., Schraven, S. P., & Mlynski, R. (2019). Stapes surgery leads to significant improvement in quality of life, independently from the surgical method: evaluation of stapes surgery using different prostheses and different quality of life measurements. *Eur Arch Otorhinolaryngol, 276*(11), 2975-2982. doi:10.1007/s00405-019-05577-4

Weiss, N. M., Schuldt, S., Oberhoffner, T., Ginzkey, C., & Mlynski, R. (2020). Prospective evaluation of disease-specific quality of life measurements after stapes surgery using Nitinol Head Prostheses. *Eur Arch Otorhinolaryngol, 277*(2), 377-384. doi:10.1007/s00405-019-05709-w

1 Quality of Life after Stapes Surgery - A Systematic Review

Author: Mai Nayeli Hauge

Journal: JAMA Otolaryngology: Head and Neck Surgery
URL: https://jamanetwork.com/journals/jamaotolaryngology

2 Abstract

Importance: Studies measuring outcomes after stapes surgery primarily direct their focus on audiometric outcomes. Recently, more studies have focused on measuring Quality of Life (QoL) in addition to audiological outcomes, extending the scope of how stapes surgery affects the patient in more ways than can be measured objectively.

Objective: To collate and evaluate findings from individual primary studies on the degree of post-operative relation between audiometric outcomes and QoL after stapes surgery in adult patients with otosclerosis, presented in a systematic review without meta-analysis.

Evidence review: PubMed, Cochrane Library and OpenGrey were periodically searched for articles in the period January 3 - June 10, 2020. Dates of coverage for all databases were 2015-2020. Titles, abstracts and full texts of retrieved articles were screened for eligibility, on the basis of exclusion criteria illustrated in a PRISMA-style flow diagram. Additionally, article quality was assessed in accordance with GRADE and PRISMA guidelines.

Findings: Of 37 articles screened, three articles were eligible for critical appraisal and qualitative synthesis. The three articles assessed showed a very low grade of quality of evidence and low grade of study design. Findings suggest that post-operative audiometric outcomes should not be the sole indication for surgical success, patient benefit or QoL. There was inconclusive evidence regarding correlation between different QoL measures and components of audiometric results.

Conclusions and relevance: Disease-specific QoL measures should be developed, translated and validated to be implemented in future clinical studies. Combining objective and subjective measures pre- and post-operatively can provide a clearer picture of the degree of relation between audiometric and QoL outcomes.

Key points

Question: To what degree is there a relation between audiometric and QoL outcomes after stapes surgery?

Findings: 3 articles were assessed in this systematic review, evaluating the relation between QoL outcomes after stapes surgery in adult patients (n=276) with otosclerosis. Findings suggest that patients experience higher QoL after stapes surgery independent from audiometric outcomes.

Meaning: Future studies should aim to develop, translate and validate more disease-specific QoL measures to widen the scope of disease-specific QoL measures after stapes surgery.

3 Introduction

Otosclerosis is one of the most common reasons behind conductive and progressive hearing loss in adults, and has been shown to have an impact on patient QoL, sometimes also termed Health-Related Quality of Life (HRQoL)[1,2]. Its main characteristic is ossification of the stapes, which can become fixed in the oval window, producing conductive and progressive hearing loss. Stapes surgery is an established surgical intervention, and many opt for stapes surgery due to the additional psychological burden of tinnitus[3,4] or lack of benefit from using hearing aids[5].

The prevalence of otosclerosis is highest in the Caucasian population, and is estimated to be 0.3-0.4%[6]. Past studies have been conducted regarding the effectiveness of stapes surgery through measuring audiometric outcomes, often reporting high success rates and low complication rates [7,8], and a common factor for desired audiometric outcomes is a reduced air-bone gap (ABG) (>10 dB)[2]. but fewer studies factor in patient QoL when assessing surgical success[9].

The reason behind treating otosclerosis surgically will differ from patient to patient according to their perceived difficulties related to hearing loss. Depending on whether their wish is to primarily reduce tinnitus or improve hearing, or both, current evidence suggests that the impact of stapes surgery on patient QoL does not necessarily correlate with audiometric outcomes[2], and it is the relation between audiometric and patient QoL outcomes after stapes surgery that will be covered in this review.

4 Methods

A systematic review without meta-analysis was conducted, using studies published from November 7, 2017, through November 23, 2019, that evaluated the relation between QoL outcomes and audiometric data after stapes surgery in adult otosclerosis patients. This systematic review was conducted using the preferred PRISMA reporting guidelines (see Appendix 4 for PRISMA checklist).

Database search and article selection

A search string was applied periodically in the following databases: PubMed, Cochrane Library and OpenGrey. Titles and abstracts were screened, and non-relevant studies were removed. After removal of duplicates, full texts were screened, removing studies not conforming to inclusion criteria (see PRISMA-style flow diagram and search string in Appendix 2). All database searches and selection of articles were done by the author.

Only primary and patient-centered studies focusing on QoL after stapes surgery were included. Reflecting the fast development in the medical field of stapes surgery, the date range was set to 2015-2020 so as to not include data >10 years of age. Although QoL is a broad term, choosing *QoL* + synonyms, in addition to *otosclerosis* and *stapes surgery* + synonyms would enable retrieval of any primary studies pertaining to this area of research. Guidance on search string development and database search was provided by a librarian from University of Oslo's Main Library.

Patients' audiometric outcomes were objectively measured pre- and post-operatively, and self-reporting of QoL were measured by questionnaires. Retrospective studies were excluded to minimize risk of bias. Additionally, studies with a population characterized my only revision cases or mixed hearing loss were excluded. Articles were excluded if presented in languages other than English, Norwegian, Swedish or Danish.

Quality assessment

Quality assessment of the primary studies were first and foremost made on the basis of the Quality Rating Scheme for Studies and Other Evidence (PRISMA). Additional quality and risk of bias assessment was based on GRADE guidelines,[10,11] in accordance with procedures for qualitative synthesis. Studies were independently assessed for quality (see Appendix 3).

5 Results

Search strategy and study selection

The data search resulted in 37 studies identified through application of search string. Two studies were duplicates and were thus removed. 35 studies were screened by title and abstract. 30 studies were removed in concordance with exclusion criteria. In total, five studies were screened for full-text, and two were removed in accordance with exclusion criteria. Thus, three [12-14] studies were eligible for critical appraisal and inclusion in qualitative synthesis. The 3 studies were published between November 7, 2017, and November 23, 2019, and met the inclusion criteria.

Assessing quality of studies

See study characteristics and assessment of quality presented in Appendix 3. All 3 selected studies were assessed for quality in accordance with GRADE criteria for quality assessment and risk of bias, using a scoring system (high = $\geq$ + 4, moderate = +3, low = +2, very low = less than +1). Specifically, assessment of quality was made on the basis of study design, population sample size, limitations and imprecision.

The three studies selected for critical appraisal were varied in study design. The first study by Dziendziel et al.[12] was a prospective clinical study with the largest sample size (n=191), the second study by Lailach et al.[13] was a clinical case study with prospective analysis and had the smallest sample size (n=37), and the third study by Weiss et al.[14] was a prospective longitudinal study with a smaller sample size (n=48). Stapedotomy was the surgical method used in two studies,[12,14] and one study[13] did not specify which surgical method was used. Total number of patients combined was 276, although missing data was an issue in all of three studies.

The first study[12] used Tinnitus Functional Index (TFI) and Abbreviated Profile for Hearing Aid Benefit (APHAB) to measure patient QoL. The second study[13] used the newly validated HRQoL measure Stapesplasty Outcome Test 25 (SPOT-25), Hearing Handicap Inventory for Adults (HHIA) and a translated but not validated German version of the Glasgow Benefit Inventory. The third study[14] used SPOT-25 and GBI. For this review topic, GBI, TFI and SPOT-25 are the QoL measures of interest.

All studies measured audiometric outcomes by conducting pure tone audiometry pre- and post-operatively. However, the presented pure tone range was varied (highest pure tone ranging 3-8 kHz), one study included cases of mixed hearing loss[12], and two studies included

revision cases.[13,14] All studies reported on the relation between patient QoL and audiometric outcomes.

All of the above factored into the assessment of quality, resulting in a GRADE score of -3 (very low), -5 (very low) and -3 (very low) in the first[12], second[13] and third study[14] respectively (see Appendix 3).

Outcomes

Dziendziel et al.[12] and Lailach et al.[13] reported difference between pre- and post-operative audiometric outcomes of air conduction thresholds (AC), bone conduction thresholds (BC) and air-bone gap (ABG), all of which were statistically significant. Weiss et al.[14] reported statistically significant change in AC and ABG, however, BC thresholds remained stable after stapedotomy.

In Dziendziel et al.,[12] pre- and post-operative TFI scores were also reported to be statistically significant, with a moderate correlation between TFI and GBI scores. They analyzed GBI scores individually, and reported a 92% increase in QoL, 2% no change, and 6% decrease.[12] Their regression analysis pointed to AC and TFI change (especially) having a significant impact on patient QoL.

Lailach et al.[13] reported a significant increase in all subscores from the SPOT-25 after stapes surgery. They also found change and a strong association between pre- and post-operative audiometric data and the overall SPOT-25 scores, with the exception of subscores pertaining to the areas of tinnitus, social restriction and mental condition. Their GBI outcomes showed a significant patient benefit, with all individual scores reporting benefit from surgical intervention. No considerable correlation between GBI and post-operative audiometric outcomes was observed.

Weiss et al.[14] also found indications of subjective patient benefit from stapedotomy in the SPOT-25 and GBI scores, as well as a positive correlation with post-operative ABG, with the exception of subscores pertaining to tinnitus and mental condition (SPOT-25) and social support (GBI). The change in tinnitus and mental condition scores were found to not correlate with the change in ABG.

6 Discussion

In this systematic review results are presented from three studies concerning the relation between audiometric outcomes and patient QoL. The most important finding is first and foremost that there seems to be an increase in patient QoL after stapes surgery, in concordance with what is reported in current literature.[2] However, stapedectomy and stapedotomy have been compared in the literature, and evidence points to stapedotomy yielding better audiometric outcomes, especially in relation to audiometric outcomes.[8] The findings in this review, however, were found on the basis of stapedotomy. Only one study did not specify surgical method.[13]

Another finding is how few prospective studies exist on the topic of the relation between patient QoL and audiometric outcomes, perhaps due to previous studies primarily focusing on hearing thresholds. There has not been found any highly rated study designs applied on this topic, perhaps due to very few disease-specific QoL measures available; SPOT-25 is the most specific QoL which applies to the part of otosclerotic population that opts for surgical intervention, and is validated in German and in the process of being validated in Dutch.[15,16]

Interestingly, Lailach et al.[13] and Weiss et al.[14] did not find much difference between audiometric data and QoL questionnaire subscores relating to tinnitus, pre- and post-operatively. In contrast, Dziendziel et al.,[12] concluded with reduction in tinnitus having the greatest impact on patient QoL, as they performed statistical tests on the correlation between TFI and GBI scores with the difference in AC thresholds. Weiss et al.[14] on the other hand, found no significant correlation between ABG and the tinnitus score of SPOT-25. These findings point to a risk of bias due to possibly skewed results in Dziendziel et al.,[12] perhaps explained by the heavy focus on tinnitus in their study.

Weiss et al.[14] noted very weak correlation between audiometric data and QoL measures, and points to the observation of different correlations depending on AC, BC, ABG threshold changes. Similarly, Lailach et al.[13] found no significant association between post-operative audiometric outcomes and GBI scores. These results suggested the need to implement one or several patient QoL measures in future studies, as studies evaluated in this review concluded with the notion that post-operative audiometric outcomes were not necessarily indicative of subjective measures of post-operative QoL, and vice versa. Furthermore, due to the studies' categorization of their population as a whole, this systematic review was unable to discern and evaluate the difference of revision surgery cases and cases

presenting mixed hearing loss to those cases with conductive hearing loss receiving primary stapes surgery.

In relation to long-term outcomes, the study by Lailach et al.[13] highlighted an important point in relation to how habituation processes in measures of patient (HR)QoL are seldomly taken into account; few studies on this particular review topic are longitudinal, perhaps because of this reason. The phenomenon pertaining to how each patient's internal views on personal health and well-being can change after acquiring a disease and being treated surgically is called response shift.[17] It has been argued that response shift needs to be controlled for in the development of future measures of patient QoL, so that they can be applied in a long-term setting.[18,19]

Apart from the overall arc in the results, which was that change in hearing thresholds indicate an increase in patient QoL, the studies had limitations in the form of considerable variation in study design, audiometric testing, post-operative follow-up period, statistical analyses and small sample sizes (in addition to missing data). These factors resulted in very low quality of evidence scores (GRADE) and low quality rating of study design (PRISMA) in all three studies. Thus it was difficult to synthesize all study outcomes into a conclusion which answers to what degree of relation there is between patient QoL and audiometric results after stapes surgery.

Limitations and future directions

The findings and conclusions of this systematic review need to be appraised in light of its limitations. A limitation in this systematic review is that it was conducted solely by the author, which lowers reliability considerably, as it would have been preferable to have at least two reviewers in the evaluation process. The sample size of articles in this systematic review was very small, as few articles met the strict exclusion and inclusion criteria, although this measure was consciously made with the aim to elevate the research quality.

On a stronger note, this systematic review highlighted a topic that spans across the subject areas of QoL, otosclerosis and stapes surgery; a topic that has yet to be reviewed, until now. Although results may not be as generalizable as preferred, this systematic review fills a small gap in the literature. The patient-centered perspective was implemented with the aim to conduct research that will result in continued development of clinical practice, enhanced communication between patient and surgeon/healthcare professional, better patient outcomes and more frequent monitoring of post-operative patient QoL.

In the future, researchers should aim to conduct studies with study designs of higher evidence rating, such as prospective controlled trials or prospective cohort trials using validated and disease-specific QoL questionnaires.

7 Conclusions

Disease-specific QoL measures should be implemented in future clinical studies, with the aim to evaluate audiometric outcomes' relation to patient QoL after stapes surgery in the otosclerotic population. Combining two types of objective and subjective measures can provide a clearer empirical picture of the relation between audiometric outcomes and patient QoL both short- and long-term. For this to be achieved, future QoL measuring tools need to be more surgery and/or disease specific, as well as translated and validated in a greater variety of languages to obtain reliable and generalizable results.

8 References

1. Arlinger S. Negative consequences of uncorrected hearing loss--a review. *Int J Audiol.* 2003;42 Suppl 2:2s17-20.
2. Weiss NM, Schuldt S, Grossmann W, et al. Stapes surgery leads to significant improvement in quality of life, independently from the surgical method: evaluation of stapes surgery using different prostheses and different quality of life measurements. *Eur Arch Otorhinolaryngol.* 2019;276(11):2975-2982.
3. Danesh AA, Shahnaz N, Hall JW, 3rd. The Audiology of Otosclerosis. *Otolaryngol Clin North Am.* 2018;51(2):327-342.
4. Bagger-Sjöbäck D, Strömbäck K, Hultcrantz M, et al. High-frequency hearing, tinnitus, and patient satisfaction with stapedotomy: A randomized prospective study. *Sci Rep.* 2015;5.
5. Chang CJ, Wen YH, Sun CH, Beltrame MA, Wu HP. Treatment of moderate-to-severe otosclerosis with simultaneous piston surgery and incus vibroplasty. *Ci Ji Yi Xue Za Zhi.* 2019;31(2):96-101.
6. Ealy M, Smith RJH. Otosclerosis. *Adv Otorhinolaryngol.* 2011;70:122-129.
7. Antonelli PJ. Prevention and Management of Complications in Otosclerosis Surgery. *Otolaryngol Clin North Am.* 2018;51(2):453-462.
8. Cheng HCS, Agrawal SK, Parnes LS. Stapedectomy Versus Stapedotomy. *Otolaryngol Clin North Am.* 2018;51(2):375-392.
9. Bowling A. Health Status and Quality of Life Assessment. In: Llewellyn CD, McManus C, Weinman J, et al., eds. *Cambridge Handbook of Psychology, Health and Medicine.* 3 ed. Cambridge: Cambridge University Press; 2019:193-195.
10. Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). *J Clin Epidemiol.* 2011;64(4):407-415.
11. Guyatt GH, Oxman AD, Montori V, et al. GRADE guidelines: 5. Rating the quality of evidence--publication bias. *J Clin Epidemiol.* 2011;64(12):1277-1282.
12. Dziendziel B, Skarzynski H, Gos E, Skarzynski PH. Changes in Hearing Threshold and Tinnitus Severity after Stapes Surgery: Which Is More Important to the Patient's Quality of Life? *ORL J Otorhinolaryngol Relat Spec.* 2019;81(4):224-233.
13. Lailach S, Schenke T, Baumann I, et al. Living with otosclerosis: disease-specific health-related quality-of-life measurement in patients undergoing stapes surgery. *Eur Arch Otorhinolaryngol.* 2018;275(1):71-79.
14. Weiss NM, Schuldt S, Oberhoffner T, Ginzkey C, Mlynski R. Prospective evaluation of disease-specific quality of life measurements after stapes surgery using Nitinol Head Prostheses. *Eur Arch Otorhinolaryngol.* 2020;277(2):377-384.
15. Lailach S, Schenke T, Baumann I, et al. [Development and validation of the Stapesplasty Outcome Test 25 (SPOT-25)]. *Hno.* 2017;65(12):973-980.
16. Blijleven EE, Thomeer H, Stokroos R, Wegner I. Protocol for a validation study of the translated stapesplasty outcome test 25 for measurement of disease-specific quality of life in Dutch patients with otosclerosis. *BMJ Open.* 2019;9(12):e030219.

17. Ring L. Quality of life. In: Baum A, McManus C, Weinman J, et al., eds. *Cambridge Handbook of Psychology, Health and Medicine.* 2 ed. Cambridge: Cambridge University Press; 2007:178-182.
18. Schwartz CE, Sprangers MA. Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Soc Sci Med.* 1999;48(11):1531-1548.
19. Schwartz CE, Rapkin BD. Reconsidering the psychometrics of quality of life assessment in light of response shift and appraisal. *Health and Quality of Life Outcomes.* 2004;2(1):16.

Appendix 1

JAMA Otolaryngology - Head & Neck Surgery: Instructions for authors

Source: https://jamanetwork.com/journals/jamaotolaryngology/pages/instructions-for-authors#SecSystematicReviews

Article Type	Description	Requirements
Systematic Review (without meta-analysis) full info	Critical assessments of the literature and data sources pertaining to clinical topics, emphasizing factors such as cause, diagnosis, prognosis, therapy, or prevention. Systematic Reviews without meta-analysis are published as Reviews; those with meta-analysis are published as Original Investigations (see Meta-analysis).	• 3500 words • 50-75 references • ≤5 tables and/or figures • A PRISMA-style flow diagram should be included as an online supplement • Include a table with ratings of the quality of the studies/evidence • Subtitle should be "A Systematic Review" • Structured abstract • Key Points • Follow EQUATOR Reporting Guidelines

Systematic Reviews are critical assessments of the literature and data sources pertaining to clinical topics, emphasizing factors such as cause, diagnosis, prognosis, therapy, or prevention. Systematic Reviews without meta-analysis are published as Reviews; those with meta-analysis are published as Original Investigations (see Meta-analysis). Systematic Reviews should address a specific question or issue that is relevant for clinical practice and provide an evidence-based, balanced, patient-oriented review on a focused topic. Follow EQUATOR Reporting Guidelines.

The basic structure of manuscripts reporting Systematic Reviews should include the following: Abstract (structured abstract of no more than 350 words); Introduction (150-250 words); Methods (150-250 words); Results (1000-1250 words, with the following subsections, if appropriate, depending on the specific question or issue addressed:

Pathophysiology, Clinical Presentation, Assessment and Diagnosis, Treatment, and Prognosis); Discussion (1000 words); and Conclusions (2-3 sentences).

Maximum length: 3000 words of text (not including abstract, tables, figures, acknowledgments, references, and online-only material), with no more than a total of 5 tables and/or figures and no more than 50-75 references. For an example of a published Systematic Review, see *JAMA*. 2014;312(6):631-640 and below for the general structure of a Systematic Review article.

Specific Components of a Systematic Review

Key Points (75-100 words)
This feature provides a quick structured synopsis of the Review, following 3 key points: Question, Findings, and Meaning. Limit to no more than 100 words. This is different from the Abstract.

Example:

> **Question:** What are the most effective medical treatments for adult chronic sinusitis?
>
> **Findings:** In this systematic review, symptoms of chronic sinusitis were improved with saline irrigation and topical corticosteroid therapy compared to no therapy. Compared with placebo, 3-week courses of systemic corticosteroids or oral doxycycline were associated with reduced polyp size, and a 3-month course of macrolide antibiotic was associated with improved symptoms in patients without polyps.
>
> **Meaning:** First-line therapy for chronic sinusitis should begin with daily topical intranasal corticosteroid in conjunction with saline irrigation; subsequent therapies should be based on the patient's polyp status and severity of symptoms.

Back to top

Abstract (350 words)
A structured abstract is required; Systematic Review articles should include a structured abstract of no more than 350 words using the headings listed below.

> **Importance:** Include 1 or 2 sentences describing the clinical question or issue and its importance in clinical practice or public health.
>
> **Objective:** State the precise primary objective of the review. Indicate whether the review emphasizes factors such as cause, diagnosis, prognosis, therapy, or

prevention and include information about the specific population, intervention, exposure, and tests or outcomes that are being reviewed.

Evidence Review: Describe the information sources used, including the search strategies, years searched, and other sources of material, such as subsequent reference searches of retrieved articles. Methods used for inclusion of identified articles and quality assessment should be explained.

Findings: Include a brief summary of the number of articles included, numbers of various types of studies (eg, clinical trials, cohort studies), and numbers of patients/participants represented by these studies. Summarize the major findings of the review of the clinical issue or topic in an evidence-based, objective, and balanced fashion, with the highest-quality evidence available receiving the greatest emphasis. Provide quantitative data.

Conclusions and Relevance: The conclusions should clearly answer the questions posed if applicable, be based on available evidence, and emphasize how clinicians should apply current knowledge. Conclusions should be based only on results described in the abstract Findings subsection.

Back to top

Introduction (150-250 words)

The first 2 to 3 sentences of the Introduction should draw in readers such that they want to continue reading the article and should establish the importance of the Review. Reviews should include the clinical question or issue and its importance for general medical practice, specialty practice, or public health. The first paragraph should provide a general summary of the clinical problem (eg, obesity). The next paragraph should focus on the specific aspect of the clinical problem the article will explore (eg, treatments for obesity). The epidemiology of the disease or condition should be briefly summarized and generally should include disease prevalence and incidence. The third paragraph should discuss exactly what material will be covered in the Review (eg, obesity treatments reported in trials with a minimum follow-up of 2 years including 80% of the original cohort).

Back to top

Methods/Literature Search (150-250 words)

The literature search should be as current as possible, ideally with end dates within a month or two before manuscript submission. A search of the primary literature should be conducted, including multiple bibliographic databases (eg, PubMed/MEDLINE, Embase, CINAHL, PsycINFO). This can be facilitated by collaborating with a medical librarian to help with the search.

Briefly describe characteristics of the literature searched and included in the review, following the PRISMA reporting guidelines, including the bibliographic databases and other sources

searched, search terms used, dates included in the search, date the literature search was conducted, screening process, language limitations, and inclusion and exclusion criteria. The rating system used to evaluate the quality of the evidence should be specified (see table below) and the methods used to evaluate quality should be described, including number of quality raters, how agreement on quality ratings was assessed, and how disagreements on quality ratings were resolved.

The highest-quality evidence (eg, randomized clinical trials, meta-analyses, systematic reviews, and high-quality prospective cohort studies) should receive the greatest emphasis. Clinical practice guidelines ordinarily should not be used as a primary component of the evidence base for the systematic review, although relevant guidelines should be addressed in the Discussion section of the article.

The search methods should be described in sufficient detail so the search can be reproduced based on the information provided in the manuscript. A summary of the methods of the literature search including this information should be included in the main article; details can be included in an online-only supplement. A PRISMA-style flow diagram showing this information should also be included as an online-only supplement. In addition, a completed PRISMA checklist should be submitted for the items completed that apply to systematic reviews (the checklist items that apply to meta-analyses do not need to be completed for systematic reviews without meta-analysis). The checklist will be used during review but will not be published.

Back to top

Results (1000-1250 words)

First, briefly report the results of the literature search, including the number of articles reviewed and included, numbers of various types of studies (eg, clinical trials, cohort studies) included, and the aggregate numbers of patients included in the reviewed studies. Also provide a brief summary of the quality of the evidence. Details of this information can be included in a PRISMA-style flow diagram and table(s).

Next, the subsections listed below should generally appear in the Results sections of most Reviews although all of these subsections may not be necessary for some topics, depending on the specific question or issue addressed. The word counts following each subsection are suggested to assist with keeping the overall Results section limited to 1000-1250 words.

> **Pathophysiology** (150-250 words). Provide a brief overview of the pathophysiology of the disease. The intent is to provide readers with sufficient background information about the underpinnings of a disease to provide context for the rest of the article.

Clinical Presentation (150-250 words). Briefly describe the clinical characteristics that result in a patient seeking medical care for the condition or what features of the disease should lead a clinician to evaluate or treat it.

Assessment and Diagnosis (250-300 words). Describe the clinical examination for evaluation of the disease and explain the most salient physical examination findings. If laboratory or imaging studies are necessary, provide the sensitivity and specificity and diagnostic accuracy of these tests and consider providing positive and negative likelihood ratios. Sequences of diagnostic tests are best presented as algorithms or in tables.

Treatment (250-500 words). Treatments should be based on the most recently available and highest level of evidence. Treatment options should be summarized in the text and presented in detail in tables along with an indication of the strength of evidence supporting the individual treatments. In general, treatment recommendations should be supported by a systematic review of the literature, either performed by the author of the Review or published in the form of a high-quality review or guideline. If possible, the costs for various treatments should be provided.

Prognosis (100-150 words). A section outlining the overall prognosis for the condition, once treated, should be included.

Discussion (Approximately 1000 words)

Key findings should be summarized in the first paragraph of the Discussion section. All statements made should be supported by evidence. It is very important to not simply list findings from the studies reviewed. This information is best presented in tables. The Discussion should provide a critical synthesis of data and information based on the results of the review, an assessment of the quality of studies summarized, and a description of how studies can be interpreted and used to guide clinical practice. The limitations of the evidence and of the review should be discussed, and gaps in evidence should be addressed. A discussion of controversial or unresolved issues and topics in need of future research also should be included.

Clinical Practice Guidelines: In the Discussion section, describe current clinical practice guidelines, relevant to the topic of the review, if available, and whether the conclusions of this review agree with, or disagree with, the current clinical practice guidelines. If this is done and there is more than 1 guideline, a table should be prepared comparing the major features that differ between the guidelines. Guideline quality should be discussed using the standards outlined for the JAMA Clinical Guidelines Synopsis.

Back to top

Conclusions

Include a 2- to 3-sentence summary of the major conclusions of the review.

Tables
Construct tables that summarize the search results. Tables summarizing treatments should have information organized by category of treatment and then by individual treatments. Columns should include the name of the treatment, strength of evidence supporting the treatment, the treatment's effect (preferably shown as the treatment's effect as compared to control on the measured outcome together with 95% confidence intervals), adverse effects, and very brief comments, if necessary. Lengthy text-based tables should be avoided. Additional or lengthy tables may be published online only, if justified.

Ratings of the quality of the evidence. Tables summarizing evidence should include ratings of the quality of the evidence. Use the rating scheme listed below with ratings of 1-5 for Reviews that include individual studies (modified from the Oxford Centre for Evidence-based Medicine for ratings of individual studies).

Quality Rating Scheme for Studies and Other Evidence	
1	Properly powered and conducted randomized clinical trial; systematic review with meta-analysis
2	Well-designed controlled trial without randomization; prospective comparative cohort trial
3	Case-control studies; retrospective cohort study
4	Case series with or without intervention; cross-sectional study
5	Opinion of respected authorities; case reports

There are several other preferred systems for rating the quality of evidence in Review articles. For Reviews that synthesize findings from numerous studies into a single summary recommendation, use the rating scale shown above or the Oxford Centre for Evidence-based Medicine's Levels of Evidence and Grades of Recommendation or the recommendations in the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. For reviews that include diagnostic studies, use The Rational Clinical Examination Levels of Evidence table.

Follow additional instructions for preparation and submission of Tables.

Figures
A PRISMA-style flow diagram should be included as an online supplement that summarizes the results of the literature search and the numbers of articles/records/studies and

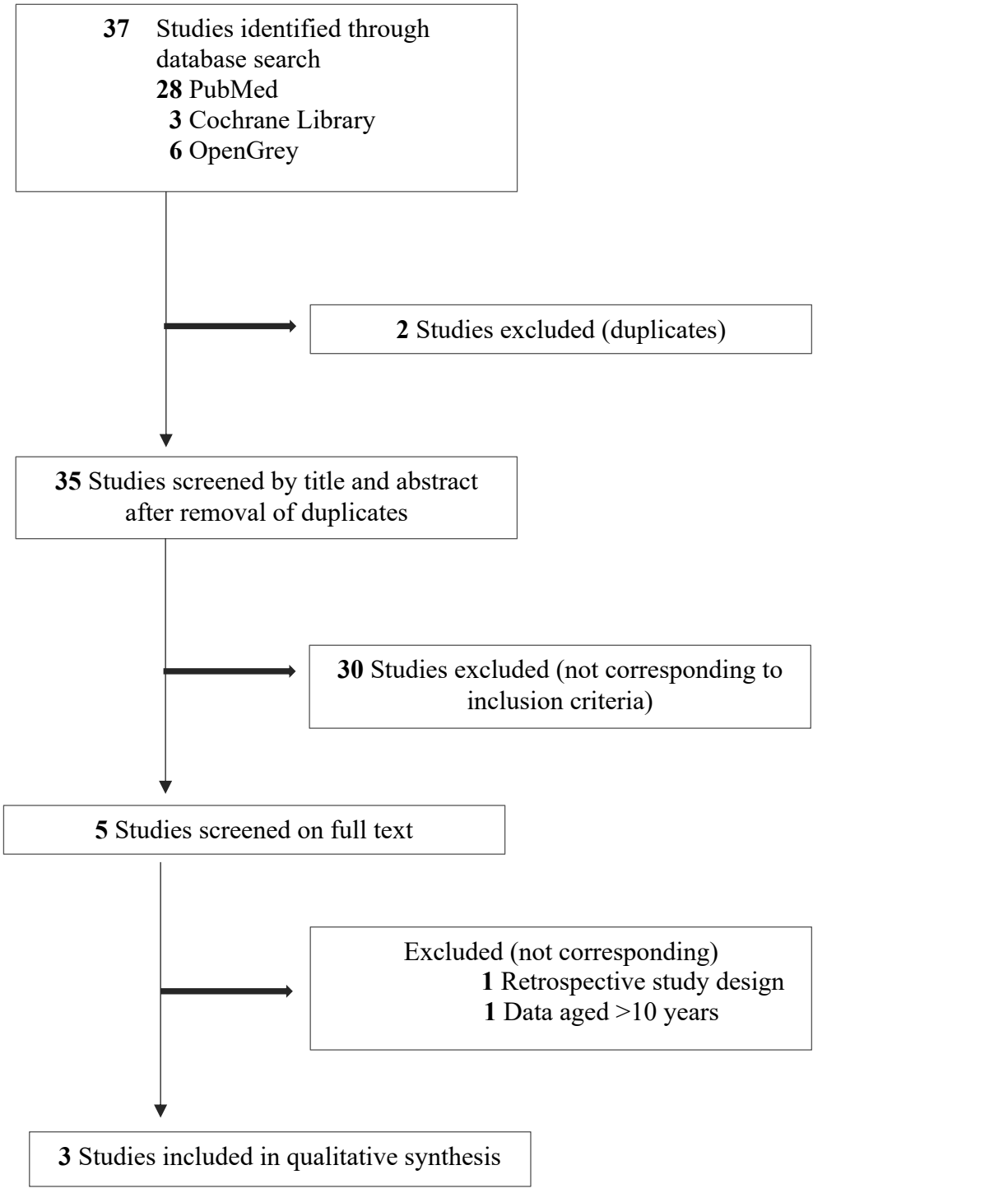
patients/participants represented in the studies identified, screened, eligible, and included in the final review.

Additional figures that illustrate pathophysiology or clinical presentation may be considered. We encourage videos, if appropriate, to illustrate a point made or process described in the Review.

Follow additional instructions for preparation and submission of Figures and Video.

Appendix 2

PRISMA-style flow diagram



Search string applied in database search:
otosclerosis AND quality of life AND (qol OR hrqol OR stapes OR stapedial OR stape* OR outcome* OR result* OR tinnitus OR hearing OR surgery OR surgeries OR operation OR operations OR stapedotomy* OR stapedectom* OR stapedotomy OR stapedectomy OR stapesplasty OR outcome* OR satisfaction)

Appendix 3

Study characteristics and assessment of quality

* GRADE score: (high = ≥ + 4, moderate = +3, low = +2, very low = less than +1)

Study	Population, No.	Study design & surgical method	Limitations	QoL Measurement tool(s)	Key findings	GRADE score*
Dzendziel et al., (2019)	191	Prospective clinical study Stapedotomy	Out of 191 patients, only 162 were available for post-operative audiometric results. Includes patients with mixed hearing loss.	Glasgow Benefit Inventory (GBI) Tinnitus Functional Index (TFI) Abbreviated Profile for Hearing Aid Benefit (APHAB)	QoL after stapes surgery depended on improvement in audiometric and self-reported hearing. The greatest impact on increase in QoL is reduction of tinnitus severity.	**-3** (very low)
Lailach et al., (2018)	37	Clinical case study, prospective analysis N/A	Smaller sample size. Low degree of information on population characteristics. Used non-validated version of GBI (German). Does not state type of surgical procedure. Includes revision cases.	Stapesplasty Outcome Test 25 (SPOT-25) Hearing Handicap Inventory for Adults (HHIA) Glasgow Benefit Inventory (GBI)	Hearing handicap and disease-specific QoL is significantly improved by stapes surgery. Audiometric parameters are not a sufficient indicator of social and mental well-being.	**-5** (very low)
Weiss et al., (2020)	48	Prospective longitudinal study Stapedotomy	Smaller sample size. 18 patients declined to complete questionnaires. Includes revision cases.	Stapesplasty Outcome Test 25 (SPOT-25) Glasgow Benefit Inventory (GBI)	Stapes surgery significantly improves hearing and HRQoL. Correlations between questionnaires and audiometric outcomes were inconclusive.	**-3** (very low)

Appendix 4

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	**1 & 2**
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	**2**
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	**3**
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	**4**
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	**4**
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	**Appendix 2**
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	**4 & 5**
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	**4**
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	**4**
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	**4**
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	

PRISMA 2009 Checklist

Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	**7 & 8**
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	**Appendix 2**
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	**Appendix 3**
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	**Appendix 3**
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	**6**
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	**7 & 8**
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	**8 - 10**
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097.

PRISMA 2009 Checklist

doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.