



Scand J Work Environ Health 2018;44(6):585-600

<https://doi.org/10.5271/sjweh.3725>

Published online: 15 Mar 2018, Issue date: 01 Nov 2018

Otoacoustic emissions versus audiometry in monitoring hearing loss after long-term noise exposure - a systematic review

by [Helleman HW](#), [Eising H](#), [Limpens J](#), [Dreschler WA](#)

Claims on the advantages of otoacoustic emissions (OAE) over pure-tone audiometry are mostly based on cross-sectional studies. This review is the first comparing both methods in longitudinal studies of noise-induced hearing loss. A notable outcome was the heterogeneity in the data, preventing a meta-analysis. Overall, changes in both methods were small. The studies agreed that OAE cannot classify individual shifts in audiometry.

Affiliation: Department of Clinical & Experimental Audiology, Academic Medical Center, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands. h.w.helleman@amc.nl

Key terms: [audiometry](#); [hearing loss](#); [long-term noise exposure](#); [noise](#); [noise exposures](#); [otoacoustic emission](#); [review](#); [systematic review](#)

This article in PubMed: www.ncbi.nlm.nih.gov/pubmed/29542804

Additional material

Please note that there is additional material available belonging to this article on the [Scandinavian Journal of Work, Environment & Health -website](#).



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

Otoacoustic emissions versus audiometry in monitoring hearing loss after long-term noise exposure – a systematic review

By Hiske W Helleman, MSc,^{1,2} Hilde Eising, MSc,^{1,3} Jacqueline Limpens, PhD,¹ Wouter A Dreschler, PhD¹

Helleman HW, Eising H, Limpens J, Dreschler WA. Otoacoustic emissions versus audiometry in monitoring hearing loss after long-term noise exposure – a systematic review. *Scand J Work Environ Health*;44(6):585–600. doi:10.5271/sjweh.3725

Objectives The objective of this systematic review was to compare otoacoustic emissions (OAE) with audiometry in their effectiveness to monitor effects of long-term noise exposure on hearing.

Methods We conducted a systematic search of MEDLINE, Embase and the non-MEDLINE subset of PubMed up to March 2016 to identify longitudinal studies on effects of noise exposure on hearing as determined by both audiometry and OAE.

Results This review comprised 13 articles, with 30–350 subjects in the longitudinal analysis. A meta-analysis could not be performed because the studies were very heterogeneous in terms of measurement paradigms, follow-up time, age of included subjects, inclusion of data points, outcome parameters and method of analysis. Overall there seemed to be small changes in both audiometry and OAE over time. Individual shifts were detected by both methods but a congruent pattern could not be observed. Some studies found that initial abnormal or low-level emissions might predict future hearing loss but at the cost of low specificity due to a high number of false positives. Other studies could not find such predictive value.

Conclusions The reported heterogeneity in the studies calls for more uniformity in including, reporting and analyzing longitudinal data for audiometry and OAE. For the overall results, both methods showed small changes from baseline towards a deterioration in hearing. OAE could not reliably detect threshold shifts at individual level. With respect to the predictive value of OAE, the evidence was not conclusive and studies were not in agreement. The reported predictors had low specificity.

Key terms DPOAE; hearing conservation; monitoring; noise-induced hearing loss; occupational hearing loss; TEOAE.

Exposure to loud noise may cause sensorineural hearing impairment, called noise-induced hearing loss (NIHL). Occupational NIHL is a hazard for workers on construction sites, in factories, on farms but also for railway workers, musicians, miners, navy or army personnel and in many other trades (1–3). Shooting, riding motor-bikes or repeated exposure to loud music in nightclubs, concerts or personal music players might cause hearing loss as well (4). Prolonged, repeated exposure to loud sounds damages the delicate structures in the cochlea. Noise exposure induces metabolic and mechanical changes causing cell death and physical loss of integrity of

hair cells (5–8). The damage starts with the outer hair cells (OHC), which form the cochlear amplifier, before damaging the inner hair cells and other structures in the cochlea (7). Functionally, the noise-induced damage results in loss of hearing sensitivity for certain frequencies, typically starting around 4 kHz, impaired understanding of speech especially in noise, and can be accompanied by decreased sound tolerance and/or ringing in the ears (tinnitus) (1, 2, 9–12). These functional deficits can result in social isolation, depression and more workplace-related accidents and injuries (13, 14). With long-term continuous exposure to noise, the dete-

¹ Clinical and Experimental Audiology, ENT department, Academic Medical Center (AMC) Amsterdam, The Netherlands.

² Department of Otorhinolaryngology and Head and Neck Surgery, Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, Utrecht, The Netherlands.

³ Pento Audiological Center, Amersfoort, The Netherlands.

Correspondence to: Hiske Helleman, Department of Clinical & Experimental Audiology, Academic Medical Center, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands. [Email: h.w.helleman@amc.nl]

rioration is gradual and increases most during the first 10–15 years of exposure (2). Damage may also occur acutely as a result of exposure to a short, high intensity sound (15). Noise exposure can cause a temporary hearing loss, ie, a temporary threshold shift (TTS), or a permanent threshold shift (PTS). Nordmann et al suggest that the underlying mechanisms for PTS and TTS are different (10).

Many countries worldwide have rules and regulations in order to protect employees from damaging their hearing. An example is the European Directive (2003/10/EC) that provides both exposure limit values and exposure action values with respect to daily and weekly exposures (16). It also specifies the allowed peak sound pressure level. The employer has to assess or measure the noise levels to which workers are exposed. The exposure limit value is 87 decibels, taking into account the attenuation provided by personal hearing protection equipment. The exposure action value is fixed at 80 decibels (lower value) and 85 decibels (upper value). The risks arising from this exposure have to be minimized by choosing methods or equipment producing less exposure to noise, instructions on the correct use of the equipment, technical measures (shielding, noise absorption) or organizational measures that reduce duration and intensity. If these measures cannot prevent the risk, the employer must provide individual hearing protection devices (HPD) and provide access to periodical audiometric screening. For exposures >85 decibels, the EU places the responsibility on the employer to ensure that hearing protection is being used. One of the main goals of hearing conservation programs is to detect hearing loss as soon as possible and halt further deterioration (2). A key role in such a program is measurement of hearing status, traditionally assessed by pure-tone audiometry. It tests the detection threshold per frequency and thus the entire auditory pathway and requires active cooperation of the subject. The presence or absence of a noise notch, or a "bulging" audiogram plays an important role in medicolegal cases, although notches at 6 kHz can also be found in subjects not exposed to noise (9, 17, 18). Other methods that are used in hearing testing in occupational settings, either separately or in conjunction with audiometry, are speech-in-noise testing (19) and the measurement of otoacoustic emissions (OAE) (1, 20–22). Since OAE are related to the functionality of the OHC, it is not surprising that the relationship of NIHL and OAE has been investigated since the discovery of OAE by David Kemp (23).

OAE are very soft sounds originating mainly from the micromechanical properties of the normal functioning OHC in the cochlea (23). They can be spontaneous or evoked by sound stimulation and can be recorded in the external ear canal. Transiently evoked OAE (TEOAE) are elicited by broad band clicks and reflect

the OHC's activity throughout the length of the basilar membrane in the cochlea, stimulus-frequency OAE (SFOAE) are emitted in response a continuous tone. Distortion product OAE (DPAOE) are evoked by two simultaneously presented pure-tone stimuli and reflect the OHC's activity at specific positions on the basilar membrane. Spontaneous OAE (SOAE) exist as well. Emissions can be classified according to the mechanism creating the emission: they can be caused by linear reflection within the cochlea (for example SOAE or low-level SFOAE or TEOAE) or arise by non-linear distortion (ie, DPOAE) (24, 25). Higher level stimuli create a combination of both types.

Measuring OAE does not require active cooperation from the subject and is therefore an objective tool. This is an important benefit when compared to pure-tone audiometry. A disadvantage is the dependency of middle-ear status for the transmission of the stimulus and response through the ear canal. A suboptimal transmission of sounds through the middle ear reduces the small stimulus and even weaker response of the inner ear and results in absent or low-level emissions, also in case of an intact cochlear amplifier (26).

OAE are currently applied in neonatal hearing screening programs worldwide, but can also be used in a more diagnostic manner such as monitoring hearing status in subjects exposed to noise or to ototoxic agents (27). Over the years, there has been much interest in a presumed role for OAE in detecting hearing loss at an earlier stage than audiometry and the hypothesized potential in predicting future hearing loss (28). Hamernik and colleagues have shown in histopathological studies that OHC damage in animals can occur without an increase in hearing thresholds (29). Such findings have led to the term OHC redundancy, implying that loss of OHC does not directly lead to loss of detection sensitivity. Several cross-sectional studies found differences in emission levels between noise-exposed and non-exposed subjects while the audiometric thresholds were within the same limits (21, 28, 30–34). For a more detailed discussion of these studies, see the review by Lapsley Miller & Marshall (28). Such findings led to the hypothesis that OAE might be a more sensitive test for cochlear function and that they might be able to detect so-called preclinical damage. There are two aspects that should be taken into account before this can be concluded from these studies. First, as emphasized by Sisto and co-authors (35), there could be a difference between two groups of subjects in audiometric thresholds even when they are both within normal limits depending on the definition (usually ≤ 25 or 20 dB HL). In their study, Sisto et al found OAE to be capable of detecting even mild hearing losses (10–20 dB HL). The second limitation is that differences on group level as detected by cross-sectional studies cannot always be regarded

as signs of future hearing loss. For an actual predictive value, longitudinal studies are required. This is the same when the goal is to identify subjects that are more vulnerable than others. In the abovementioned review from 2007, Lapsley Miller & Marshall (28) called for more large-scale longitudinal studies and emphasized that more knowledge was required about optimal OAE parameters. From our own knowledge in this field, it was felt that the results and conclusions from the few longitudinal studies since then have not shown consistent results. The lack of consistency among these studies was the basis for this review, which aims to provide a comparison between different longitudinal studies on OAE and audiometry on behalf of policy-makers, audiologists and occupational hygienists. Its focus is the role of OAE in monitoring NIHL after long-term exposure compared to audiometry by investigating and structuring available data in a well-defined, reproducible and systematic manner. We compared the setup, methodology, and quality of different studies, before we analyzed the outcomes on group-averaged data and on a subject level (individual shifts). The focus of this systematic review was on the possibility of (i) replacing audiometry by OAE in hearing conservation programs and (ii) early detection in the form of a predictive value or identification of vulnerability. We sought for agreement between studies with respect to these issues and possible overarching trends.

Methods

Protocol and registration

This review was prospectively registered in the Prospero database under number [PROSPERO 2015:CRD42015027111](https://www.crd.york.ac.uk/PROSPERO/record/CRD42015027111) and reported according MOOSE guidelines (36).

Literature search

A medical librarian (JL) performed a comprehensive search in OVID MEDLINE, OVID Embase and the non-MEDLINE subset of PubMed from inception to 14 March 2016 to identify studies on the use of OAE to monitor NIHL. The search included both freetext and controlled terms (ie, Mesh in MEDLINE) for OAE and NIHL or activities known to be related to NIHL (certain occupations and leisure activities). No language or other restrictions were applied. The entire MEDLINE search strategy is shown in Appendix A (www.sjweh.fi/show_abstract.php?abstract_id=3725). On completion, citations identified in each database were imported into EndNote and de-duplicated. Forward and backward snowballing of the identified relevant papers was applied and the search was adapted in case of additional relevant

studies. Corresponding authors were contacted via email if their studies could not be obtained otherwise.

Eligibility criteria

Original studies were included in which subjects were exposed to noise (continuous or impulse) and hearing status was assessed on more than one occasion (longitudinal or repeated measures approach) with both audiometry and evoked OAE (TEOAE and/or DPOAE). Studies on animals, infants or neonates were excluded. Studies on OAE and audiometry with noise *and* another intervention (ototoxicity or preventive strategy as in antioxidants) were excluded except when there was a control group with noise as the sole intervention.

Selected articles

Two authors (HH and HE) independently screened titles and abstract of all included studies. When disagreement occurred whether or not to exclude a paper in this stage, consensus was reached through discussion or by consultation of a third reviewer (WD). The same two authors independently examined the fulltext articles, again with consensus through discussion and/or subsequent consultation of the third reviewer. Screening was conducted with support of Covidence systematic review software (Veritas Health Innovation, Melbourne, Australia). After full-text screening only studies reporting on long-term (weeks/years) exposure and permanent effects on hearing were included.

Quality assessment

A modified Downs & Black checklist (37) was used to assess quality, including reporting, external and internal validity, risk of bias. The checklist originally consists of 27 items with a maximum count of 32 points. For this review, questions that were not relevant or applicable were omitted and other items were adjusted slightly, with more attention on the reporting of confounders and appropriate statistical tests (multiple comparisons with ears and frequencies). The adapted checklist (see Appendix A) consisted of 14 items with a total count of maximum 16 points. Items 1–3, 5–7, 11, 16, 17, 20, and 26 were incorporated from the original list with an explicit question relevant for this topic. Items 4, 9 and 18 were modified. If no explicit hypothesis was given, the objective(s) of the study had to be clear. Confounders that we felt needed to be addressed, were age, middle ear status, previous and recent noise exposure, and use of hearing protection. Reporting on these items in a sufficient manner would yield two points, whereas partial or unclear reporting on these items yielded one point. Any attempt to address variation in the published

data was rewarded with one point. Standard errors (SE) of the mean were considered sufficient, although standard deviations (SD) or confidence intervals (CI) are preferred from a statistical point of view (displaying actual spread or estimated effect size). The final item on the reporting section deals with test statistics. It was required to have a full description of test statistics, degrees of freedom, and P-value. The original Downs & Black checklist only requires the P-value to be given explicitly. External validity was assessed with the question whether the participating subjects were representative of the entire population from which they were recruited. Four items dealt with internal validity, with a maximum of five points. The final question concerned selection bias: were losses to follow-up taken into account? This question was made more explicit by asking whether the amount of excluded data points was reported if inclusion was based on a certain signal-to-noise ratio (SNR) criterion.

The first two authors independently assessed quality and, when the scores on separate items were different, a consensus score was reached through discussion. Studies were not excluded based on the outcome but the overall quality assessment was used in the narrative comparison of the included studies.¹

Summary measures and synthesis of the data

A narrative approach was used to qualitatively compare studies in descriptive characteristics (population, exposure, age, gender, etc.), aim, methodology, outcome and conclusion. The principal (quantitative) outcome measures were hearing threshold levels and emission amplitude and the change from baseline values for these measures. Both the vulnerability assessment and individual analyses were discussed in a narrative manner. A simplified approach allowed comparison of the size of change from baseline across different longitudinal studies by ignoring possible effects of initial hearing status or the inclusion criterion applied and by averaging changes across frequencies. This approach was chosen because of anticipated difficulties in combining numerical outcomes from different studies, but we realize that this provides only a first-order approximation. Although the typical noise-notch occurs around 4 kHz, we expected some studies to look at broader frequency ranges than at 4 kHz only. In order to take this specific region into account and to be able to compare across studies, the changes between 2 and 8 kHz were averaged (ie, one octave above and one octave below the noise-notch area). This allowed comparison across studies with different frequency regions,

¹Articles written by two of the four authors of this review (HH and WD) were scored in a similar matter by the first two authors (HH and HE) of this review. It was felt that it would be of more value to address these articles in a similar and reproducible manner, than to omit assessing these articles.

possibly at the cost of underestimating the maximal effect. Besides the reported frequencies, there could be differences in the way thresholds were derived (manually or automatically) and the step size or test resolution chosen (usually 5 dB).

For OAE, the measurement paradigm may differ across studies, eg, with respect to the use of SNR or emission amplitude as outcome measure, the level of stimuli, the frequency resolution, single or average measures, and exclusion criteria applied. For DPOAE, the reported emission levels (not SNR) between 2 and 8 kHz was averaged into one single measure, regardless of the chosen inclusion SNR criterion. Similarly, for TEOAE the response in the 4 kHz region was used in this analysis. The actual emission amplitude is a direct measure of cochlear function whereas SNR also reflects measurements conditions.

Results

Study selection

The search identified 657 references. The PRISMA flowchart (38) summarizing the data collection process, number of records in each step and reasons for exclusion is presented in figure 1. Based on fulltext, 120 articles were assessed and 105 studies dealing with short-term (hours/days) and temporary effects on hearing were excluded. Only 15 studies described effects of prolonged noise exposure on hearing, and 2 of these long-term studies were excluded.

Study characteristics

General description. All 13 eligible studies reported changes in hearing status of subjects occupationally exposed to noise. They described 11 unique populations. Both Seixas and Helleman and their co-authors (39–42) reported on (more or less) the same population in two different manuscripts. Seixas and his co-authors presented data on the same population but with a different follow-up time (three versus ten years) and these papers were analyzed separately. Helleman and co-authors reported on the same group, with the same follow-up but with a different approach in the analysis (individual versus group results) and these papers were considered as one study with respect to the outcomes. There was considerable variation in methodological and other properties of included studies. This made it difficult to describe overarching trends.

Table 1 addresses the major descriptives of the included studies. Three studies dealt with impulse noise (43–45), but the majority assessed the effects of pro-

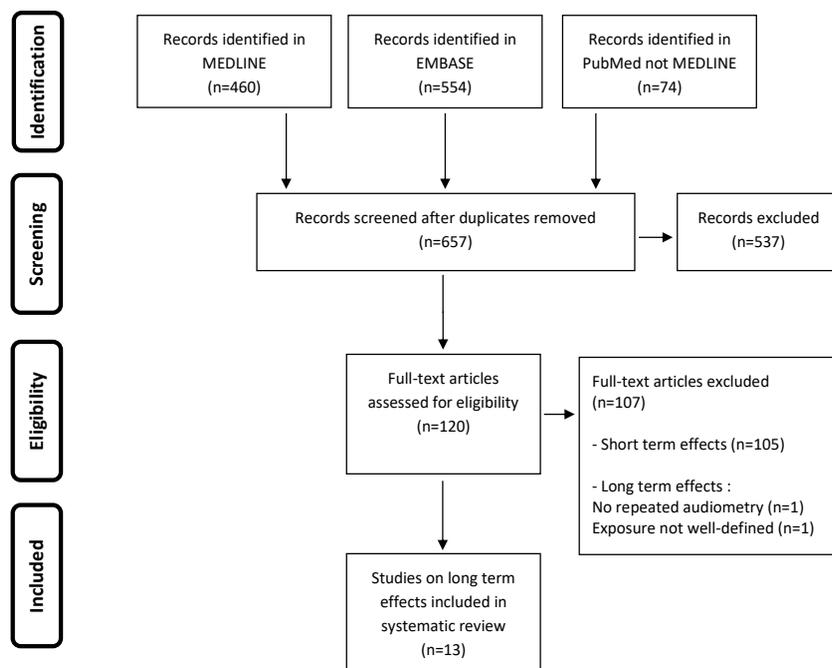


Figure 1. Flow diagram of the conducted search (based on the PRISMA 2009 Flow Diagram).

longed, continuous noise exposure (39–42, 46–51). Changes in hearing status of professionals were reported for both OAE (TEOAE and/or DPOAE) and pure-tone audiometric thresholds. All studies were observational and had no control over the noise exposure, with six studies having a non-exposed group serving as control group. Some studies only report group-averaged results, while others perform an analysis on a particular subgroup or perform analyses on individual changes in OAE and audiometry. Six studies explicitly test the hypothesis whether (a form of) OAE are suited to predict individual susceptibility and look at the predictive value (P) (43, 45, 47–49, 51). See table 1 for more details.

The age ranges of the subjects included differed between studies from a narrow range (18–20 years) on young army recruits to a broad range (19–61 years) for other studies. A smaller age range with no previous noise exposure form a relatively homogenous group in initial hearing status. Older subjects with a known history of noise exposure might enter the studies with a pre-existing hearing loss.

Some studies initially report a few hundred subjects but the numbers in table 1 are the actual amount of subjects contributing to the longitudinal analysis and these numbers are generally much smaller. There were studies combining the results for left and right ears, studies using ear as factor in analyses and studies presenting changes for left and right ears separately whilst performing other analyses on the ears combined. In order to combine this for all included studies, left and right ears were combined in the first order, overall

analysis presented in this review. Thus, the number of ears measured at baseline and at one of the follow-up measurements ranged from 56–518 and the number of subjects ranged from 30–350.

Most studies report a baseline measurement and one follow-up although there were also studies that measured hearing status repeatedly during the duration of the study, ranging from six months for impulse noise to ten years for continuous noise (41, 43, 46). The period between baseline and final measurement ranged from several weeks in the case of high-level impulsive noise to ten years in the case of exposure to more continuous industrial noise. There could be large – and unknown – differences in the actual sound levels to which subjects were exposed. These differences could be caused by the following confounders: nature and level of the noise sources, duration between initial and final measurement and the use and quality of hearing protective devices.

Quality assessment. Risk of bias was assessed for all studies included. The range of items met on the modified Downs & Black scale was 9–14 with a mean of 11.8 (SD 1.3). See Appendix B (www.sjweh.fi/show_abstract.php?abstract_id=3725) for a more detailed explanation of the scored items and the questions written in full. The scored items per study are found in Appendix C (www.sjweh.fi/show_abstract.php?abstract_id=3725).

In the reporting section (items 1–8, 9 points maximum), the scores of the individual articles ranged from 5–9, (mean 7.13, SD 1.30). The objectives are reported in table 1. One study did not mention clear goals, aims or

Table 1. General description of the included studies. Unless otherwise stated, the reported numbers correspond to numbers in the study group (exposed subjects) not for the control group or the baseline group. Follow-up time is the time between the first and last measurement used for group analysis. Please note that the number of subjects does not provide complete information because in some studies subjects with incomplete datasets were excluded, while ears or single data points were excluded in others. [OAE=otoacoustic emissions; DPOAE=distortion product OAE; SOAE=spontaneous OAE; TEOAE=transiently evoked OAE; IaDP=index of abnormality DPOAE; B=baseline; G=group analysis, I=individual analysis, PV=predictive value or vulnerability analysis, M=male, F=female, yrs=years, wks=weeks, mos=months; LE=low exposure, HE=high exposure, I/O=input-output curve.]

Study	Study aim	Study population	Analysis (G/I/PV) baseline	Subjects (N) per analysis				Age at baseline (years)			Follow-up time	Control group	OAE type	
				B	G	I	PV	Median	Mean	SD				
Duvdevany & Furst 2007 (43)	Characterization of changes in audiograms and TEOAE over time and investigation of PV ear vulnerability	Combat soldiers	G PV	84 M	30 M	-	30 M	-	-	18	-	2 yrs	No	TE
Konopka et al 2005 (44)	Comparison of the sensitivity of TEOAE with standard and extended audiometry in early evaluation of acoustic trauma	Soldiers	G	92 M	92 M	-	-	-	-	19	-	1 yr	Yes	TE
Marshall et al 2009 (47)	Investigation of vulnerability based on low-level or absent OAE in future hearing loss after impulse noise exposure	Marine recruits	G I PV	401 ^a M	60 M	285 M	285 M	17.4-28.1	19.2	-	-	3 wks	Yes	TE DP
Helleman & Dreschler 2012 (39)	Comparison of group-averaged and individual changes for audiometry, TEOAE, DPOAE in order to see if they exhibit the same pattern in time	Workers in printing office	I	320 (316M 4F ^b)	-	233 M	-	23-60	42	-	-	17 mos	No	TE DP
Helleman et al 2010 (40)	Exploring boundary conditions in which the use of OAE might have a contributing role in a general hearing conservation program	Workers in printing office	G	320 (316M 4F ^b)	233 M	-	-	23-60	42	-	-	17 mos	No	TE DP
Job et al 2009 (48)	Investigation of vulnerability for future hearing loss with initial DPOAE	Pilots	G PV	521 M/F ^c	350 M/F ^c	-	160 M/F ^c	20-40 ^d	-	-	-	3 yrs	No	DP IaDP
Lapsley Miller et al 2006 (49)	Assess changes in audiometric threshold and OAE in sailors after 6 months of hazardous noise exposure on an aircraft carrier	Sailors	G I PV	338 (303M 35F)	75 (67M 8F)	338 (303M 35F)	-	18-46 ^d	22M 21F	-	-	9 mos	Yes	TE DP
Lapsley Miller et al 2004 (46)	Investigation of OAE in a hearing conservation program to find if, when and how OAE change with hearing thresholds	Navy personnel	G I	474 ^e M/F ^c	69 (57M 12F)	69 (57M 12F)	-	14-49 ^a	-	24.5F 25.8M	5.1F 5.9M	1 yr ^f	Yes	TE DP SOAE
Moukos et al 2014 (50)	Comparison of initial and the final audiological evaluation in a long follow-up with conventional audiometry and DPOAE	Workers tobacco company	G I	76 (58M 14F ^g)	34 (29M 5F)	34 (29M 5F)	-	32-54	-	43.7	6.9	9 yrs	Yes	DP
Murray et al 1998 (47)	Investigation if OAE testing gives earlier warning of potential hearing loss than PTA	Musicians	G I PV	119 (61M 58F)	74 M/F ^c	74 M/F ^c	74 M/F ^c	19-61 ^a	-	-	-	5 yrs	No	TE Repro
Seixas et al 2012 (41)	Characterization of the effects of noise exposure including intermittent and peaky exposure, on hearing damage as assessed by PTA and OAE	Construction apprentices	G	372 ^h M/F ^c	258 (229M 29F)	-	-	-	-	27.6	6.6	10 yrs	Yes	DP
Seixas et al 2005 (42)	Evaluation of the relations between noise exposure related changes over time for a range of frequencies and DPOAE primary intensities	New construction workers	G	372 M/F ^c	268 (240M 28F)	-	-	-	-	27.2 LE 27.8 HE	6.7 LE 6.8 HE	3 yrs	Yes	DP I/O
Shupak et al 2007 (51)	Comparison of changes in TEOAE, and DPOAE with audiometric thresholds during the first 2 years of occupational noise exposure	Ship engine room recruits	G P	135 M	42 M	-	42 M	18-20 ^d	-	-	-	2 yrs	Yes	TE DP

^a This number includes the control group; ^b Female participants were excluded to avoid gender effects; ^c Number male/female not mentioned; ^d Age only reported for baseline group; ^e This number is for the study group; ^f Time between measurements in individual analysis may > 1 year; ^g Number of male/female given after excluding four subjects; ^h This number is not mentioned in the article, but copied from the original study by the same authors.

hypotheses (45). It was deduced from the results and similar studies by the same authors. Four studies mentioned their outcome measures for the first time in the results section of their article while they were not described in the introduction or methods (46, 47, 50, 51). Two did not clearly mention the noise exposure in terms of (estimated) levels and/ or durations (39, 40). Six studies obtained two points for describing the confounders (41, 42, 45, 46, 49, 51), seven studies partially addressed confounders, and obtained one point (39, 40, 43, 44, 47, 48, 50).

With respect to reporting on variability in the data, studies have used SE, SD and/ or CI with different P-values to report their data. This did not allow for a comparison. Two studies did not address the variability in the data (44, 48). Eight studies fulfilled the criterion for reporting the statistics appropriately (39, 40, 43, 45, 46, 49–51).

Although the source from which subjects were drawn was usually quite clear (eg, soldiers, construction workers etc.), only two studies were explicit about the way subjects within this group were recruited (42, 46). The range for the scores on the internal validity (four items) was

3–5. Six studies reported explicitly how the complexity of left and right ears and the repeated nature of frequencies were taken into account, resulting in two points for question 12 (39, 40, 43, 44, 46, 50). Six studies did not report on the loss to follow-up (41, 43, 44, 46, 50, 51).

Test characteristics

Table 2 provides information on the characteristics of the stimuli and measurement protocols for the audiometry and OAE measurements. It can be seen that there are differences in frequency span, resolution and properties of the OAE-stimuli such as the used stimulation level.

An important difference between studies is the required SNR for an emission [(either transiently evoked (TE) or distortion product (DP)] to be entered into the study. Only the study by Duvdevany & Furst did not mention whether or not data points, ears or subjects, were excluded based on SNR (43). Seixas and co-authors explicitly mentioned that they did not impose an inclusion criterion (41, 42). In contrast, Shupak et al (51) only included data points with an SNR ≥ 6 and Job

Table 2. Set-up parameters for pure-tone audiometry (PTA) and otoacoustic emissions (OAE). All studies were conducted with frequency ratio of f2/f1=1.2. [Stim int=stimulus, intensity; stab=stability; oct=octave; SNR=signal-to-noise-ratio, NFS=noise floor substitution, L=linear, NL=non-linear, L1=level of first primary, L2=level of second primary. NM=not mentioned; DPOAE=distortion product OAE; TEOAE=transiently evoked OAE].

Authors	PTA		TEOAE				DPOAE			
	Frequency	Inclusion	Frequency range, Δ (kHz)	Stim int (dB pSPL)	Mode	No of average	Inclusion	Frequency range f2 (kHz)	Frequency res Δ (total)	L1/L2 (dB SPL)
Duvdevany & Furst (43)	0.25, 0.5, 1, 2, 3, 4, 6, 8	NM	overall	84	NL	260	-	-	-	-
Konopka et al (44)	0.25-12	SNR ≥ 0	1-5 (1 kHz)	80	NL	260	-	-	-	-
Murray et al (47)	0.5, 1, 1.5, 2, 3, 4, 6, 8	stab $\geq 80\%$	overall/wave repro	80	NL	260	-	-	-	-
Helleman et al (39,40)	0.5, 1, 2, 3, 4, 6, 8	SNR ≥ 0 NFS	1-4 (1/2 oct)	80	NL	280	SNR $\geq 0^a$ NFS	0.8-8.0	1/8 oct (27)	75/70
Lapsley Miller et al (49)	0.5, 1, 2, 3, 4, 6	SNR ≥ 0 NFS	0.7-4 (1/2 oct)	74	NL	260	SNR $\geq 0^a$ NFS	1.8-4.5	1/6 oct (9)	57/45, 59/50, 61/55, and 65/45
Lapsley Miller et al (46)	0.5, 1, 2, 3, 4, 6, 8	SNR ≥ 0 NFS	0.7-4 (1/2 oct)	L: 80, 74, 68, 62 NL: 80, 74	L NL ^b	540	SNR $\geq 0^a$	1.0-8.0	31 frequency spaced evenly in log2 steps	70/60, 65/55, 65/45, 60/50, 60/40, 55/45, and 55/35
Marshall et al (45)	0.5, 1, 2, 3, 4, 6	SNR ≥ 0 NFS	0.7-4 (1/2 oct)	74	NL	260	SNR $\geq 0^a$ NFS	1.8-4.5	1/6 oct (9)	57/45, 59/50, 61/55, and 65/45
Shupak et al (51)	0.5, 1, 2, 3, 4, 6, 8	SNR $\geq 6^d$	1-4 (1/2 oct)	80	NM	256	SNR $\geq 6^d$	0.9-12.0	1/3 oct (12)	65/55
Job et al (48)	0.25, 0.5, 1, 2, 3, 4, 8	-	-	-	-	-	SNR ≥ 2	0.6-8.9	1/8 oct (32)	65/55
Moukos et al (50)	0.25, 0.5, 1, 2, 3, 4, 6, 8	-	-	-	-	-	SNR $\geq 0^a$ NFS	1.0-6.0	1/2 oct (7)	65/55
Seixas et al (41)	0.5, 1, 2, 3, 4, 6, 8	-	-	-	-	-	No criterion ^c	2.0-8.0	1/2 oct (5)	65/55
Seixas et al (42)	0.5, 1, 2, 3, 4, 6, 8	-	-	-	-	-	No criterion	1.0-10.0	21 log-spaced f2 frequency	65/55

^a Relative to 2 standard deviations of the noise floor estimate.

^b NL mode not analyzed.

^c Only in sub analysis SNR ≥ 6 .

^d For the baseline measurement only.

et al (48) required $SNR \geq 2$. The other studies required $SNR \geq 0$ for an emission at a certain frequency to be entered in the data set (39, 40, 44, 45, 46, 49). With one exception, so-called noise floor substitution was applied in these studies (39, 40, 45, 46, 49). This approach allows more data to be used when initially present emissions drop below the noise floor in follow-up measurements. Such substitution possibly underestimated the actual effects, see for a more detailed description one of above-mentioned original studies (39, 40, 45, 46, 49).

Outcomes

The results were first analyzed for the group-averaged data and compared with the control groups if available. The second step was to look at individual shifts in audiometry and OAE. The third step focused on the possibility of predicting a shift in threshold from OAE-parameters or the possibility of identifying vulnerable subjects. Finally, the general conclusions on the use and role for OAE from the studies were compared.

Outcomes: comparison of group changes. Seven studies reported significant changes in audiometry and OAE (TE and/or DP) over time (39–41, 43, 44, 46, 48, 50, 51). The study on the musicians did not find any significant effect on either audiometry or TEOAE (47). The three remaining studies did not observe a significant change from baseline for audiometric thresholds but did find a significant change for both TEOAE and DPOAE (45, 49) and for DPOAE alone (42). For these studies, the frequency range for the TEOAE-effects was 1–3 kHz

or 1–4 kHz and the effect size approximately 1 dB SPL (45, 49). For the DPOAE, an effect size of 1.5 dB SPL was found at 2–4 kHz (49) and an effect size of 0.8 dB SPL at 2.5–3.6kHz (45). Seixas and co-authors found small but significant decrements (0.5 dB SPL) per year for a group of young construction apprentices, with a significant difference in response over time compared to controls in the 3–4 kHz region (42). In the follow-up study by the same group, the construction workers were measured after ten years and both audiometry and DPOAE were significantly deteriorated (41). The long-term changes were very small for both workers and controls and showed a similar time course. But in the region <6 kHz for audiometry and ≥ 3 kHz for DPOAE, the changes for the construction workers were larger than for the controls. It was computed that per 10 dBA increase in exposure level, the hearing thresholds increased with 2–3 dB HL and the DPOAE with 1 dB SPL during these ten years. In the nine year follow-up of Moukos et al (50), the changes in audiometry took place between 4–6 kHz, while the largest effects in DPOAE were found between 3–5 kHz. This was the same DPOAE region as in the study with pilots where the audiometric changes took place at and >3 kHz (48). Shupak et al (51) found significant DPOAE changes in the 4–6 kHz region that were accompanied by TEOAE changes between 2–4 kHz and audiometric changes between 4–6 kHz.

Both TEOAE and audiometry changed significantly in two studies on soldiers exposed to impulse noise, but there were differences in the frequency region where the effects took place: Duvdevany & Furst (43) observed a

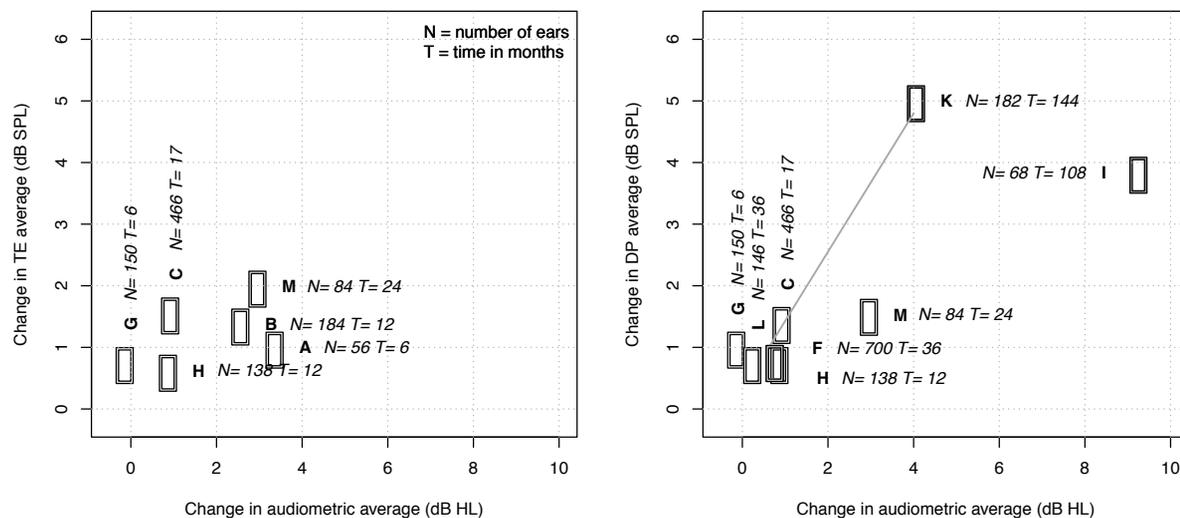


Figure 2. Shift in audiometry (dB HL) versus shift in emission level (dB SPL) for TEOAE in the 4 kHz band (A) and for the high frequency averaged DPOAE (B). Please note the ranges of the axes. The color of the circles corresponds to the duration, with darker colors representing longer durations. The dashed line represents the two studies by Seixas and colleagues going from 3 years follow-up (L31) to 10 years (L8).

decrease in wideband TEOAE at a group level accompanied by threshold increases at 1 kHz and higher frequencies but without a significant correlation for individual changes. Hearing thresholds of the soldiers in the study by Konopka et al (44) increased in the extended high-frequency range (≥ 10 kHz) parallel to a decrease in TEOAE level at 2–4 kHz, while these effects were not significant in the control group (44). Helleman et al (40) and Lapsley Miller et al (46) both observed that TEOAE showed a small decrease and in a broader frequency range than audiometry and DPOAE but this effect was either not significant or also occurred in the control group. Helleman et al (40) found significant changes for audiometry in the 6–8 kHz region, and for 1–2 kHz and 4–6 kHz in DPOAE. They also reported an increase in emission strength around 3 kHz for the DPOAE. Lapsley Miller et al (46) found a significant change of 2.0 dB HL in audiometric thresholds around 3–4 kHz, accompanied by an overall decrease in DPAOE-emission level of 2.3 dB SPL between 1–3 kHz.

Figure 2 is a simplified representation of the above-mentioned changes. The group-averaged changes (shifts) from baseline threshold were compared with the changes (shifts) from baseline emissions level with extra information regarding duration of exposure and the number of contributing ears.

Some concordance in the effects in OAE and audiometry can be seen. With one exception, all hearing thresholds increase and emission levels are generally lower in the follow-up measurement (49). Both effects imply a deterioration in hearing, which can be expected in a noise-exposed, ageing population. However, the

effects are rather small, amounting to 1–2 dB in audiometry up to three years and 4–9 dB HL for longer durations. Moukos et al (50) report the largest average change in audiometry from baseline with almost 10 dB HL after nine years in the tobacco industry versus almost 5 dB HL in the ten year study by Seixas et al (41). In contrast, the changes in DPOAE were larger for the younger construction workers in the study from Seixas and colleagues when compared with the tobacco workers from Moukos et al. Age, initial hearing status, compliance with the use of hearing protection devices and exposure levels might have affected the size of the observed changes in hearing. Changes in TEAOE have been followed for at most three years, amounting to a maximum shift from baseline of 2 dB SPL in the highest frequency band. DPOAE in the same time frame shift on average 2 dB SPL and up to 4–5 dB SPL for longer durations of exposure.

Outcomes: Individual shifts. Six studies investigated individual shifts in audiometry and OAE. Four of these looked at both TEOAE and DPOAE (39, 45, 46, 49), one only at DPOAE (50) and one at TEOAE (47). The first step in such an approach is to determine which change in audiometric thresholds and emission level qualifies as a real shift, the second step is to compare the number of threshold shifts with emission shifts and look for agreement. All different approaches and numerical values to define a significant shift are expressed in table 3.

The continuous data of change in emission amplitude or change in hearing threshold level is transferred in a dichotomous "yes" or "no", discarding information on

Table 3. Significant threshold shifts (STS) and significant emissions shift (SES), their size and origin. [PTA=pure-tone audiometry; DPOAE=distortion product otoacoustic emissions; TEOAE=transiently evoked otoacoustic emissions; NM=not mentioned; CI=confidence interval; SEmeas=standard error of measurement, Δ SEmeas=standard error of measurement of the difference

Authors	PTA: STS		TEOAE: SES		DPOAE: SES	
	Size (dB SPL)	Rationale	Size (dB SPL)	Rationale	Size (dB SPL)	Rationale
Helleman & Dreschler (39)	>14.5	Δ SEmeas 95% CI, 60 subjects (111 ears, test-retest)	4.0	Δ SEmeas 95% CI, 60 subjects (111 ears, test-retest)	7.0–12.4 ^a	Δ SEmeas 95% CI, 60 subjects (94–109 ears, test-retest)
Lapsley Miller et al (49)	>8.3–25 ^a	Δ SEmeas 98% CI, 28 subjects (56 ears, control)	3.2–7.5 ^a	Δ SEmeas, 98% CI, 28 subjects (35–46 ears, control)	4.6–8.5 ^a	Δ SEmeas, 98% CI, 28 subjects (33–43 ears, control)
Lapsley Miller et al (46)	>8.3–25 ^a	Avg+3 \times SEmeas 53 subjects (106 ears, control)	3.5–7.6 ^a	Avg+3 \times SEmeas 53 subjects (83–103 ears, control)	5.0–7.2 ^a	Avg+3 \times SEmeas 53 subjects (93–106 ears, control)
Marshall et al (45)	>8.3–15 ^a	Δ SEmeas 98% CI, 32 subjects (64 ears, control)	4.0–6.1 ^a	Δ SEmeas 98% CI, 32 subjects (36–54 ears, control)	6.0–8.7 ^a	Δ SEmeas 98% CI, 32 subjects (36–54 ears, control)
Moukos et al (50)	≥ 15	NIOSH (52)	-	-	5.8–12.4 ^a	Avg+3 \times SEmeas 33 subjects (58–65 ears, control)
Murray et al (47)	≥ 15	Australia/NZ 1269 subjects (53)	Wave repro >28%	From unpublished study	-	-
Duvdevany & Furst (43)	≥ 10	NM	-	-	-	-
Job et al (48)	>5	All frequency ≤ 0 dB HL change to ≥ 1 frequency ≥ 15	-	-	-	-
Shupak et al (51)	≥ 10	NM	-	-	-	-

^a Frequency dependent.

the spread of the data. Such a fence criterion can be obtained by adopting established criteria from literature or standards, creating a new criterion, or using a statistical criterion based on test–retest measurements. The majority of studies used the term significant threshold shift (STS) or significant emission shift (SES) and these terms were adopted in this review.

Table 3 shows that the significant shifts range from 5–25 dB HL for audiometry, 3.2–7.6 dB SPL for the TEOAE and from 4.6–12.4 dB SPL for the DPOAE. It should be noted that some studies have used an average of several frequencies where others use a shift at a single frequency. We refer to the original papers for more details on the used criterion, as the underlying frequencies and the reasons behind each choice.

The next step is to compare the number of shifts in audiometry (STS) with shifts in OAE (SES) and look at congruency between cases. The first general observation, regardless of the chosen criterion was that the number of significant shifts was low when compared to the total number of ears. This implies that for the majority of ears, the difference from baseline was not large enough to qualify as a significant individual shift, even though most group results were significant. Significant shifts may occur in both directions, but the main focus was on worsening of hearing sensitivity and thus on decrease of emission amplitude and increase in hearing threshold level.

The majority of studies reported the number of ears, not the number of subjects having significant shifts. Everything here is reported in number of ears, discarding information on left versus right ears. Based on the studies that mentioned both subjects and number of ears, there seemed to be more unilateral than bilateral shifts. Some studies reported the actual number of ears, others in percentages of the total valid data or percentage of ears with repeated measurements². Table 4 shows the number of significant shifts for audiometry, TEOAE and DPOAE per study.

The percentage of permanent STS ears (worsening of hearing threshold levels) ranged from 4.4–43% with differences in the reported frequency region. The minimal amount of shifts was 4.4% ($N_{STS}=18$) at the average of 2.3 and 4 kHz (49), increasing to 7.4% at combinations between 2–4 kHz ($N_{STS}=42$) (45), 8.7% at combinations between 2–6 kHz ($N_{STS}=12$) (46), 9.5% at 6 kHz ($N_{STS}=14$) (47), 13.7% ($N_{STS}=64$) at the average of 6 and 8 kHz (39) and finally to the maximum percentage of shifts of 43% at both 4 and 8 kHz ($N_{STS}=29$) (50). The latter amount deviates strongly from the other studies, while the used criterion is in the range of the others. The group-averaged threshold and emission shift of this study were also much larger than other studies and significantly

² It was not always straightforward to deduce the numbers from percentages and vice versa, caused by the vagueness of the exact number of ears in that particular computation.

larger than in the control group. Factors that could have played a role in this difference are the longer duration and the potential presence of temporary threshold shifts. The workers were measured during a workday, but after a noise-free period of at least one hour. Such an effect could have been also present in the baseline measurement making it impossible to estimate the effect. This argumentation is also valid for the study by Helleman & Dreschler where the measurements were performed in a similar matter (39, 40). A more valid approach to separate temporary from permanent shifts is to confirm the STS by remeasuring the audiogram after at least a few noise-free days as was done by several other authors (45, 46, 49). Although audiometric shifts in the opposite directions (decrease in hearing threshold level) were mentioned to occur incidentally, they were not investigated any further.

The total number of SES could be obtained in five of the above-mentioned studies, three on both TEOAE and DPOAE (39, 45, 49), one on TEOAE only (47), and one on DPOAE only (50). One study only mentioned the occurrence of SES in case of an STS (46). The numbers are given in table 4. In case of the significant emission shifts (decrease) for TEOAE, the range was 6.8–14%. The minimal amount of shifts was 6.8% ($N_{SES,TE}=10$) (47), increasing to 8.6% ($N_{SES,TE}=49$) (45), 12% ($N_{SES,TE}=41$) (49) and is maximal at 14% ($N_{SES,TE}=62$) (39). Two studies also observed significant increases in TEOAE-emission level, with 10% ($N_{SES,TE+}=47$) (39) and 24% ($N_{SES,TE+}=35$) significant shifts (47). In another study, some increases in emission level were seen in the 18 STS cases, but they were considered improvements and therefore not investigated any further (49).

For the DPOAE, the percentage of significant emission shifts ranged from 4–52% for decreasing emission levels. The minimal amount of shifts was 4% at 1.5 kHz ($N_{SES,DP}=20$) (39), increasing to 7.7% between 2.5 and 4 kHz ($N_{SES,DP}=44$) (45), 12% at 2.5 kHz ($N_{SES,DP}=41$) (49) and the largest number of shifts amounted to 52% at 5 kHz ($N_{SES,DP}=34$) (50). Again significant increases in emission level were found: 9% shifts for the DPOAE level at 3 kHz ($N_{SES+,DP}=41$) (39). In other studies, significant individual increases in DPOAE were mentioned but not investigated because they were considered as random error (50) or as improvements and only mentioned for STS ears (49).

For the majority of studies, the numbers of significant shifts and percentages with respect to the total number of ears were small. A possible explanation for low numbers of shifts can be found in emission data that were excluded based on a SNR criterion. Cases with emission levels dropping below this criterion would be excluded, leading to an underestimation of the actual number. This consideration was put forward by several authors and the previously mentioned noise floor substitution can partially resolve this (39, 40, 45, 46,

Table 4. Numbers of significant threshold shifts (STS) and significant emissions shift (SES) in number of ears and percentage, percentage of maximal agreement between STS and SES cases. - Indicates a decrease in emission level, + indicates an increase in emission. If not explicitly stated otherwise, the numbers in the table correspond to a decrease in emission level. [PTA=pure-tone audiometry; DPOAE=distortion product otoacoustic emissions; TEOAE=transiently evoked otoacoustic emissions.]

Authors	PTA: STS		TEOAE: SES		DPOAE: SES		Agreement STS/SES _{TE}		Agreement STS/SES _{DP}	
	N (ears)	%	N (ears)	%	N (ears)	%	N (ears)	%	N (ears)	%
Helleman & Dreschler (39)	64	13.7	SES 62 SES+ 47	14 10	SES 20 SES+ 41	SES 4.4 SES+ 9	SES- 3 SES+ 8		SES- 2 SES+ 10 ^a	0.4 2.1 ^a
Lapsley Miller et al (49)	18	4.4	41	12	41	12	-- ^b	-- ^b	-- ^b	-- ^b
Lapsley Miller et al (46)	12	8.7	-- ^c	-- ^c	-- ^c	-- ^c	4	2.9 ^d	2	1.5 ^d
Marshall et al (45)	42	7.4	49	8.6	42 ^e	7.4 ^e	12	2.1	7 ^e	1.2 ^e
Moukos et al (50)	29	43	--	-	34	52	--	--	13	19.6
Murray et al (47)	14	9.5	SES 10 SES+ 35	6.8 2.4	--	--	1 -- ^f	0.6 -- ^f	--	--

^a For the combination of STS and SES+ at 3kHz, SES- at 1.5 kHz.

^b Unknown number/percentage: only mentioned that 31% of STS ears had SESTE either increase or decline, and 50% had SESDP for any of the measured frequencies, either increase or decline.

^c Total number of SES not reported, only reported for ears with a permanent STS.

^d Percentage was not given in original text with different number of ears per analyses, computed for this table based on percentage STSs column.

^e For another DP setting the number of SES was 44, with an accompanying 6 STS/SES cases.

^f Number of ears with SES+ and STS not mentioned in original text.

49). But still, there might be cases where the substitution could not be applied since the emissions were low in both measurements. Such cases were not entered in the analysis and recorded as missing data.

Overall there were only a few cases that had both a STS and an STS. The final column of table 4 expresses the maximal agreement in the number of ears having both a shift in audiometry and OAE. The percentages of agreement are relatively small, ranging from 1–19% of the total number of ears. The study from Lapsley Miller et al (46) only reported the number of SES cases for the 12 ears ($N_{STS}=12$) that also had an STS. Several studies mentioned that no association with SES and STS status was found, with all combinations occurring (39, 45, 47, 50). This was emphasized by two studies providing a scatterplot of change in TEOAE amplitude (39) or TEOAE wave reproducibility (waverepro) (47) versus audiometric changes. These graphs showed the lack of agreement in SES and STS and the spread of the data³.

To summarize these findings: despite differences in criteria and frequency region, the actual numbers exhibiting shifts in both OAE and audiometry were very small, calling for caution and reluctance in the interpretation of the results. This was also emphasized by Marshall and co-authors (45). From the two papers providing a scatterplot, it can be seen that the number of ears exhibiting a shift might alter with a different fence criterion. These graphs also show the evident lack of a relationship between changes in OAE and audiometry both on a continuous scale and when classified as SES and STS. The scatterplots also show that there were increases in emission level and decreases

in hearing threshold level. When explicitly mentioned (39, 47), these changes in OAE were in the same order of magnitude, but nevertheless they were regarded as random variation or as outliers in other studies (49, 50).

Outcomes: predictive value and vulnerability. Six studies investigated a possible role for OAE in predicting future hearing loss as measured in audiometry. Table 3 expresses the size of change in audiometric threshold that was classified as a shift. Five approaches were based on low or absent initial OAE levels predicting a change in audiometric threshold. This could be done retrospectively (possible vulnerability) or prospectively (predictive value). Another approach was chosen by Duvdevany & Furst (43). They tested ear vulnerability retrospectively among soldiers with TEOAE. The group, all initial hearing thresholds ≤ 20 dB HL, was split into two subgroups: having threshold changes at any frequency ≤ 5 dB (no hearing loss, NHL) or ≥ 10 dB (slight hearing loss SHL). The authors found that the SHL group had less variation in TEOAE values in time and could relate this to having “medium” emission strength. This led the authors to conclude that subjects having normal audiograms in combination with either relatively strong or low emission strength had more “tough ears” than ears having medium emission strength.

Murray and co-authors looked if TEOAE results are able to provide a warning for potential hearing loss in an earlier phase than pure-tone audiometry (47). They predefined a target group (based on another control group) with hearing levels within normal limits (< 25 dB HL) and with low-level emission, defined as an initial waverepro $< 35\%$. There was no evidence for this parameter to be of predictive value since there were about

³ Counting the cases in the scatterplot lead to other numbers than were reported in the text, the numbers from the discussion in the text are adopted here.

as many cases exhibiting shifts in audiometry without changes in waverepro or without low initial waverepro. They concluded that further investigation was required.

Lapsley Miller et al noted that in the 16–18 ears exhibiting STS, there was a relatively high amount of OAE data either missing or already low at baseline (49). They looked further into this matter by examining positive predictive values of absent or low-level emissions as a predictor for the occurrence of an STS, and found that for ears with lower emissions, there was an increased risk for developing PTS of 17–20% with TEOAE, 14–17% for DPOAE. They concluded that OAE might be a diagnostic predictor for NIHL, showing damage to the inner ear before hearing loss is present in the audiogram (49). In the next study by this group, it was investigated whether low level or absent emissions increased the chance of developing a STS for subjects exposed to impulse noise (45). Per type of OAE there were 17–21 ears with a STS compared with 217–263 ears without. If both ears from one subject were measured, the worst ear (hypothetically the most susceptible) was chosen for the computation of the likelihood ratios. The increased risk compared to the baseline risk of getting a STS maximized to nine fold depending on the condition. The authors concluded that OAE are predictive of incipient NIHL but, in view of the small numbers in the study, the results are indicative only (45). This was in contrast to the conclusions by Shupak et al who reported that lower (minus 2 SD below average) initial OAE are inappropriate for predicting future elevations in pure-tone thresholds (51). When they used another, absolute criterion on the same data [ie, adopted from Priev et al and defined as signal repeatability <50%, or SNR <3 dB, or absolute emission level <5 dB SPL (54)] OAE could label ears as either resilient or vulnerable. But this came at the cost of a high false-positive rate and thus questions the practicality of such a tool (51). Job et al agreed with the group of Lapsley Miller et al (48) when they tried to identify vulnerable ears by creating an index of abnormality (the so-called IaDPOAE). Their goal was to predict ears shifting from normal hearing (≤ 10 dB HL) to having an increased threshold (>25 dB HL). The index was based on normative data from a control group, with a high abnormality corresponding to a low DPOAE amplitude. They concluded that ears with an index of abnormality $\geq 15\%$ had larger odds of changing from the normal hearing group to the group having one threshold >25 dB HL. They found a significant but low (0.27) correlation between initial IaDPOAE and final hearing thresholds. These findings led them to conclude that DPOAE could be a biomarker of vulnerability with continuing noise exposure (48). Finally, although not explicitly set out to investigate any predictive value, Helleman & Dreschler tried to verify these findings by

looking at the absence of emissions in the STS ears but the odds ratio was not significant (ie, 5 ears missing OAE-data with a STS, versus 20 missing OAE data without a STS) (39).

Discussion

Strengths and weaknesses

As far as the authors know, this systematic review is the first attempt to systematically combine results from several observational studies on longitudinal changes in hearing as measured with OAE and audiometry. Only long-term effects on emission amplitudes were investigated and compared, whereas short-term effects and (contralateral) acoustical suppression were not examined. A separate review could be conducted on studies investigating short-term effects; more than hundred studies were found and could be examined further at full-text level. When also looking at (contralateral) acoustic suppression of OAE, even more studies are available for analysis.

In the current situation, the heterogeneity in the long-term studies did not allow for a meta-analysis to compare changes or compare statements on enhanced probabilities. The first order attempt to combine changes in audiometry and OAE could be done by a simplification of the actual data. Frequency information, initial hearing status and for example hearing at time-points between baseline and final measurement were omitted. The constructed graph illustrates that the group-averaged changes were only small for all individual studies.

Potential biases in the review process. Some of the included studies aimed to explore the relationship between OAE and audiometry whereas others explicitly set out to investigate a form of predictive value. Publication bias is a risk in any field in the hypothesis-generation stage, especially when the a-priori possibility that the hypothesis is true, is small and/or the statistical power is low (55). Many relationships that were found are based on small numbers and would require further investigation in larger studies to be confirmed. Another point for discussion is the inclusion of two articles by two of the authors from this review (ie, HH and WD). Because of the limited amount of studies available, it was felt that omitting these studies would be less favorable than a potentially less objective assessment of the quality of own work. No studies were excluded based on the quality and the group-averaged result from these studies lies amidst the cluster of other studies, so no bias was expected.

Justification for exclusion. A limitation in this study is the exclusion of some papers that could not be obtained. This amount was minimized by several attempts to contact the corresponding author through email and ResearchGate. Nevertheless, there were three longitudinal studies in Polish by the same author that were not available. For two of these, details in the abstract were identical to a paper by the same author that was included (44). Consequently, no bias from omitting these papers was expected. The third unavailable Polish paper by the same author, concerned effect of jet engine noise on technical staff, impulse noise for soldiers and a control group (56). The conclusion from the abstract mentions that there were more changes in TEOAE after one year than in audiometry. These effects could not be assessed numerically and therefore could not be combined with the results in other studies.

Conference papers or grey data were not explicitly searched for in this review.

Assessment of quality of included studies. The quality as assessed by the Black & Downs checklist differed across studies but no studies were excluded based on this assessment. There were considerable differences in the domains of reporting data and analysis of repeated measurements. But comparison of raw data was considered to be important despite quality issues in the above-mentioned items. In retrospect, more stringent quality criteria for reporting and analyzing could have been applied to allow for more differentiation in the quality between the studies. As long as no studies were excluded based on quality, this would not have changed the overall conclusions but it might have given more weight to higher quality studies.

Consideration of alternative explanations for observed results. All studies were observational studies without control and exact knowledge over the actual noise exposure in terms of level, duration and the use of hearing protective devices. The prediction of a future threshold shift does not depend solely on initial hearing status, the (unknown) exposure is the actual cause for the damage that occurs. This makes it difficult to distinguish between ears that are inherently more sensitive and ears that have just been exposed more between measurements (by higher noise levels, longer duration of exposure or inconsistent use of hearing protection).

Generalization of the conclusions. The overarching conclusions of this review are that the studies are very different and heterogeneous in many aspect, and that the overall change in both methods are relatively small for the time frame that hearing was followed. Besides, all studies agreed that all combinations between emission shifts and threshold shifts occur, ie, shifts in emission

without accompanying shift in audiometry, shifts in audiometry without accompanying shift in OAE, shifts in both (always lowest count) and finally, no shifts in both methods. In the studies included, the largest agreement was for the ears showing no shifts. Individual changes in both methods had no or very low correlations. So generally speaking, OAE and audiometry failed to identify the same subjects exhibiting significant shifts.

Hearing threshold level as measured in the pure-tone audiogram still is the reference standard. For the low number of threshold shifts that occurred, OAE could not be used to reliably detect a change in audiometry when based on a baseline test and one follow-up measurement. It cannot be ruled out that with a higher prevalence of shifts, OAE could be more able to identify them, for example after exposure to higher noise levels for a longer time. Whether an emission shift precedes the occurrence of a threshold shift could not be answered based on the studies included in this review. This question can only be answered by studies with at least three measurements available for analysis. Several papers reported that lower or absent emissions indicated a higher risk for future threshold shifts (45, 46, 48, 51). Different statistical parameters have been used to express this increased probability. Methods using the odds ratio, positive predictive values and likelihood ratios were based on the presence of a shift in a certain group versus another group. The chosen criteria for size and definition of a shift have a large impact on these statistics and thus on the proposed relationships. There were also many cases with low-level or absent emissions that did not exhibit audiometric shifts, thus creating many false positives.

Besides this dependency of the sensitivity for the chosen criteria and the false positives, some studies also presented other outcomes: Murray and co-authors did not find evidence for a predictive value (47) and Duvdevany & Furst reported that having either high or low TEOAE at baseline was an indicator for resilient ears (43). They reported that the ears with “medium” strength emissions were more at risk and that resilient ears could also be ears that had lower emissions at the start (43). So there was no consensus on the hypothesis that the ears with lower emissions at the start are more sensitive to noise-induced audiometric shifts.

Recommendation for further research

It would have been desirable if this review could end with clear, unequivocal recommendations for future research. The heterogeneity mentioned earlier does unfortunately not allow for such statements. The first goal should be to reduce the differences in setup between studies and resolve some methodological issues. General consensus in the field is needed concerning stimulus parameters

and measurement paradigms. A simple recommendation is the use of tympanometry to avoid changes caused by middle-ear pathology. Care should be taken to avoid TTS by introducing a suitably long noise-free period between latest exposure and measurements. Another recommendation is to use the emission strength and not SNR when working with emissions. The SNR is a useful measure of quality but the outcome itself is dependent upon measurement conditions whereas the emission amplitude itself reflects properties of the cochlea.

Next, we recommend more complete reports of raw data to allow the reader to make an assessment of the data for him or herself. We call for a more uniform approach in reporting emission and audiometric data by presenting the raw data with a measure of spread, i.e. standard deviation per frequency in a graph or table. Noise levels should be included as well to allow the reader to assess the measurement conditions. Correct statistics should take into account (some of) the dependency between frequencies and ears. Another example of required information is data on the exclusions: how many data points/ ears / subjects have been omitted for emissions, and for audiometry? How does this potentially affect the conclusions?

It is clear that several approaches were used to define an individual shift. When the underlying data is not presented, the effect of the size of the criterion cannot be assessed by the reader. Scatterplots may provide information on a continuous scale, whilst fence criteria are dichotomous. Agreement on the details how to define a shift, analysis for the effect the chosen fence, analysis on a continuous scale or a simple, uniformly accepted definition of a shift could also allow better comparison between studies and the prevalence of shifts that occur.

Concluding remarks

The studies were very heterogeneous making it impossible to perform a meta-analysis on the available data. There were several factors responsible for this heterogeneity such as the studied populations in terms of number of subjects, age, initial hearing status, and noise exposure (level and duration). Properties of the OAE formed another factor responsible for differences between the studies (for example level of primaries, method of including emissions). Quality assessment by the Black & Downs checklist made it clear that there were differences in reporting style, clearness of the applied statistical methods, missing data and subgroup analyses. With respect to the analysis of the data, there was a large variation in the applied statistical methods and the definitions used for individual shifts used for subgroup analyses.

The first-order attempt used in this study to pool the data required overlooking many sources of con-

founders and differences between the results of the individual studies. When looking at the overall results, both audiometry and OAE showed small changes from baseline towards a deterioration in hearing. There were many methodological complications in the definition of individual shifts. Nevertheless, it is safe to state there was no clear congruent behavior in the combined occurrence of audiometric and otoacoustic shifts in any of the studies. Therefore, the results of this study support the conclusions of several authors that the main contribution of OAE is in addition to pure-tone audiometry rather than instead of.

When low numbers of ears with threshold shifts were investigated more closely, some studies suggested that specific abnormal OAE-properties possibly indicated a higher risk for future hearing loss. But the underlying statistical methods are sensitive for the criteria chosen and there was no agreement on whether 'abnormal' emission were low-level, absent or abnormally high. These discrepancies imply that there is neither consensus nor clear evidence that OAE are able to predict future noise-induced threshold shifts. It would be interesting to compare all the data from the studies from this review and analyze the effect of the reported criteria. This could give more insight if generalization of the conclusions is possible, especially with respect to the role of OAE as being predictive, being an inherent biomarker for noise-induced hearing loss or just a symptom of it.

References

1. Feder K, Michaud D, McNamee J, Fitzpatrick E, Davies H, Leroux T. Prevalence of Hazardous Occupational Noise Exposure, Hearing Loss, and Hearing Protection Usage Among a Representative Sample of Working Canadians. *J Occup Environ Med.* 2017;59(1), 92-113. <https://doi.org/10.1097/JOM.0000000000000920>
2. Kirchner DB, Evenson E, Dobie RA, Rabinowitz P, Crawford J, Kopke R, et al. ACOEM Task Force on Occupational Hearing Loss, Occupational noise-induced hearing loss: ACOEM Task Force on Occupational Hearing Loss. *J Occup Environ Med.* 2012;54(1), 106-8. <https://doi.org/10.1097/JOM.0b013e318242677d>
3. Leensen MC, van Duivenbooden JC, Dreschler WA. A retrospective analysis of noise-induced hearing loss in the Dutch construction industry. *Int Arch Occup Environ Health.* 2011;84(5), 577-90. <https://doi.org/10.1007/s00420-010-0606-3>
4. Carter L, Williams W, Black D, Bundy A. The leisure-noise dilemma: hearing loss or hearsay? What does the literature tell us? *Ear Hear.* 2014 Sep-Oct;35(5):491-505. <https://doi.org/10.1097/01.aud.0000451498.92871.20>
5. Kurabi A, Keithley EM, Housley GD, Ryan AF, Wong AC.

- Cellular mechanisms of noise-induced hearing loss. *Hear Res.* 2017 Jun;349:129-137. <https://doi.org/10.1016/j.heares.2016.11.013>
6. Talaska AE, Schacht J. Mechanisms of noise induced damage to the cochlea. *Audiol med.* 2007;5(1):3-9. <https://doi.org/10.1080/16513860601158887>
 7. Henderson D, Bielefeld EC, Harris KC, Hu BH. The role of oxidative stress in noise-induced hearing loss. *Ear Hear.* 2006;27(1):1-19. <https://doi.org/10.1097/01.aud.0000191942.36672.f3>
 8. Sliwinska-Kowalska M, Jedlinska U. Prolonged exposure to industrial noise: Cochlear pathology does not correlate with the degree of permanent threshold shift, but is related to duration of exposure. *J Occup Health.* 1998; 40 (2), 123-131. <https://doi.org/10.1539/joh.40.123>
 9. McBride D, Williams S. Characteristics of the audiometric notch as a clinical sign of noise exposure. *Scand Audiol.* 2001;30(2):106-11. <https://doi.org/10.1080/010503901300112211>
 10. Nordmann AS, Bohne BA, Harding GW. Histopathological differences between temporary and permanent threshold shift. *Hear Res.* 2000;139(1-2):13-30. [https://doi.org/10.1016/S0378-5955\(99\)00163-X](https://doi.org/10.1016/S0378-5955(99)00163-X)
 11. May JJ. Occupational hearing loss. *Am J Ind Med.* 2000;37(1):112-20. [https://doi.org/10.1002/\(SICI\)1097-0274\(200001\)37:1<112::AID-AJIM9>3.0.CO;2-#](https://doi.org/10.1002/(SICI)1097-0274(200001)37:1<112::AID-AJIM9>3.0.CO;2-#)
 12. Chung DY, Mack B. The effect of masking by noise on word discrimination scores in listeners with normal hearing and with noise-induced hearing loss. *Scand Audiol.* 1979;8(3):139-43. <https://doi.org/10.3109/01050397909076314>
 13. Girard SA, Leroux T, Courteau M, Picard M, Turcotte F, Richer O. Occupational noise exposure and noise-induced hearing loss are associated with work-related injuries leading to admission to hospital. *Inj Prev.* 2015;21(e1):e88-92. <https://doi.org/10.1136/injuryprev-2013-040828>
 14. Hétu R, Getty L, Quoc HT. Impact of occupational hearing loss on the lives of workers. *Occup Med.* 1995;10(3):495-512.
 15. Axelsson A, Hamernik RP. Acute acoustic trauma. *Acta Otolaryngol.* 1987;104(3-4):225-33. <https://doi.org/10.3109/00016488709107322>
 16. Directive 2003/10/EC of the European Parliament and of the Council of 6 February 2003 on the minimum health and safety requirements regarding the exposure of workers to the risks arising from physical agents (noise) (Seventeenth individual Directive within the meaning of Article 16(1) of Directive 89/391/EEC). *Official Journal of the European Union* 46 PART 42, 38-44.
 17. Lutman ME, Coles RR, Buffin JT. Guidelines for quantification of noise-induced hearing loss in a medicolegal context. *Clin Otolaryngol.* 2016;41(4):347-57. <https://doi.org/10.1111/coa.12569>
 18. Coles RR, Lutman ME, Buffin JT. Guidelines on the diagnosis of noise-induced hearing loss for medicolegal purposes. *Clin Otolaryngol Allied Sci.* 2000;25(4):264-73. <https://doi.org/10.1046/j.1365-2273.2000.00368.x>
 19. Leensen MC, Dreschler WA. The applicability of a speech-in-noise screening test in occupational hearing conservation. *Int J Audiol.* 2013;52(7):455-65. <https://doi.org/10.3109/14992027.2013.790565>
 20. Forshaw C. Report of an International Expert Symposium on the usefulness of Otoacoustic Emissions (OAE) Testing in Occupational Health Surveillance 8-9th February 2011. Manchester: Corporate Medical Unit, Health & Safety Executive, 2011.
 21. Marshall L, Lapsley Miller JA, Heller LM. Distortion-Product Otoacoustic emissions as a screening tool for noise-induced hearing loss. *Noise Health.* 2001;3(12):43-60.
 22. Sliwinska-Kowalska M, Kotylo P. Otoacoustic emissions in industrial hearing loss assessment. *Noise Health.* 2001;3(12):75-84.
 23. Kemp DT. Stimulated acoustic emissions from within the human auditory system. *J Acoust Soc Am.* 1978 Nov;64(5):1386-91. <https://doi.org/10.1121/1.382104>
 24. Shera CA. Mechanisms of mammalian otoacoustic emission and their implications for the clinical utility of otoacoustic emissions. *Ear Hear.* 2004;25(2):86-97. <https://doi.org/10.1097/01.AUD.0000121200.90211.83>
 25. Shera CA, Guinan JJ Jr. Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAE. *J Acoust Soc Am.* 1999;105(2 Pt 1):782-98. <https://doi.org/10.1121/1.426948>
 26. Zhao F, Wada H, Koike T, Stephens D. The influence of middle ear disorders on otoacoustic emissions. *Clin Otolaryngol Allied Sci.* 2000 Feb;25(1):3-8. <https://doi.org/10.1046/j.1365-2273.2000.00312.x>
 27. Konrad-Martin D, Reavis KM, McMillan G, Helt WJ, Dille M. Proposed comprehensive ototoxicity monitoring program for VA healthcare (COMP-VA). *J Rehab Res Dev* 2014; 51(1), 81-100. <https://doi.org/10.1682/JRRD.2013.04.0092>
 28. Lapsley Miller J, Marshall L. Otoacoustic emissions as a preclinical measure of noise-induced hearing loss. In: Robinette M, Glatke T, editors. *Otoacoustic Emissions: Clinical Applications.* 3rd ed. New York: Thieme Medical Publishers; 2007. p. 321-41.
 29. Hamernik RP, Patterson JH, Turrentine GA, Ahroon WA. The quantitative relation between sensory cell loss and hearing thresholds. *Hear Res.* 1989;38(3):199-211. [https://doi.org/10.1016/0378-5955\(89\)90065-8](https://doi.org/10.1016/0378-5955(89)90065-8)
 30. Attias J, Horovitz G, El-Hatib N, Nageris B. Detection and Clinical Diagnosis of Noise-Induced Hearing Loss by Otoacoustic Emissions. *Noise Health.* 2001;3(12):19-31.
 31. Desai A, Reed D, Cheyne A, Richards S, Prasher D. Absence of otoacoustic emissions in subjects with normal audiometric thresholds implies exposure to noise. *Noise Health.* 1999;1(2):58-65.
 32. Xu ZM, Vinck B, De Vel E, van Cauwenberge P. Mechanisms in noise-induced permanent hearing loss: an evoked otoacoustic emission and auditory brainstem response study. *J Laryngol Otol.* 1998;112(12):1154-61. <https://doi.org/10.1017/S0022215100142720>

33. Attias J, Bresloff I, Reshef I, Horowitz G, Furman V. Evaluating noise induced hearing loss with distortion product otoacoustic emissions. *Br J Audiol*. 1998;32(1):39-46. <https://doi.org/10.3109/03005364000000049>
34. Engdahl B, Woxen O, Arnesen AR, Mair IW. Transient evoked otoacoustic emissions as screening for hearing losses at the school for military training. *Scand Audiol*. 1996;25(1):71-8. <https://doi.org/10.3109/01050399609047559>
35. Sisto R, Chelotti S, Moriconi L, Pellegrini S, Citroni A, Monechi V, et al. Otoacoustic emission sensitivity to low levels of noise-induced hearing loss. *J Acoust Soc Am*. 2007;122(1):387-401. <https://doi.org/10.1121/1.2737668>
36. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *JAMA*. 2000; 19;283(15)
37. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health*. 1998;52(6):377-84. <https://doi.org/10.1136/jech.52.6.377>
38. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*. 2015;1;4:1.
39. Helleman HW, Dreschler WA. Overall versus individual changes for otoacoustic emissions and audiometry in a noise-exposed cohort. *Int J Audiol*. 2012;51(5):362-72. <https://doi.org/10.3109/14992027.2011.653447>
40. Helleman HW, Jansen EJ, Dreschler WA. Otoacoustic emissions in a hearing conservation program: general applicability in longitudinal monitoring and the relation to changes in pure-tone thresholds. *Int J Audiol*. 2010;49(6):410-9. <https://doi.org/10.3109/14992020903527616>
41. Seixas NS, Neitzel R, Stover B, Sheppard L, Feeney P, Mills D, et al. 10-Year prospective study of noise exposure and hearing damage among construction workers. *Occup Environ Med*. 2012;69(9):643-50. <https://doi.org/10.1136/oemed-2011-100578>
42. Seixas NS, Goldman B, Sheppard L, Neitzel R, Norton S, Kujawa SG. Prospective noise induced changes to hearing among construction industry apprentices. *Occup Environ Med*. 2005;62(5):309-17. <https://doi.org/10.1136/oem.2004.018143>
43. Duvdevany A, Furst M. The effect of longitudinal noise exposure on behavioral audiograms and transient-evoked otoacoustic emissions. *Int J Audiol*. 2007;46(3):119-27. <https://doi.org/10.1080/14992020600937402>
44. Konopka W, Pawlaczyk-Luszczynska M, Sliwiska-Kowalska M, Grzanka A, Zalewski P. Effects of impulse noise on transiently evoked otoacoustic emission in soldiers. *Int J Audiol*. 2005;44(1):3-7. <https://doi.org/10.1080/14992020400022561>
45. Marshall L, Lapsley Miller JA, Heller LM, Wolgemuth KS, Hughes LM, Smith SD, et al. Detecting incipient inner-ear damage from impulse noise with otoacoustic emissions. *J Acoust Soc Am*. 2009;125(2):995-1013. <https://doi.org/10.1121/1.3050304>
46. Lapsley Miller JA, Marshall L, Heller LM. A longitudinal study of changes in evoked otoacoustic emissions and pure-tone thresholds as measured in a hearing conservation program. *Int J Audiol*. 2004;43(6):307-22. <https://doi.org/10.1080/14992020400050040>
47. Murray NM, EL LePage, N Mikl. Inner ear damage in an opera theatre orchestra as detected by otoacoustic emissions, pure tone audiometry and sound levels. *Aust J Audiol* 1998; 20 (2), 67-78.
48. Job A, Raynal M, Kossowski M, Studler M, Ghernaoui C, Baffioni-Venturi A, et al. Otoacoustic detection of risk of early hearing loss in ears with normal audiograms: a 3-year follow-up study. *Hear Res*. 2009;251(1-2):10-6. <https://doi.org/10.1016/j.heares.2009.02.008>
49. Lapsley Miller JA, Marshall L, Heller LM, Hughes LM. Low-level otoacoustic emissions may predict susceptibility to noise-induced hearing loss. *J Acoust Soc Am*. 2006;120(1):280-96. <https://doi.org/10.1121/1.2204437>
50. Moukos A, Balatsouras DG, Nikolopoulos T, Maragoudakis P, Yiotakis EI, Korres SG, et al. A longitudinal study of changes in distortion-product otoacoustic emissions and pure-tone thresholds in an industrial setting. *Eur Arch Otorhinolaryngol*. 2014;271(10):2649-60. <https://doi.org/10.1007/s00405-013-2754-z>
51. Shupak A, Tal D, Sharoni Z, Oren M, Ravid A, Pratt H. Otoacoustic emissions in early noise-induced hearing loss. *Otol Neurotol*. 2007;28(6):745-52. <https://doi.org/10.1097/MAO.0b013e3180a726c9>
52. National Institute for Occupational Safety and Health (1998) Criteria for a recommended standard: occupational noise exposure. DHHS (NIOSH) Publication No. 98-126. U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, Cincinnati
53. Australian/New Zealand Standard, Occupational Noise management, Standards Australia, 1998
54. Prieve BA, Gorga MP, Schmidt A, Neely S, Peters J, Schultes L, et al. Analysis of transient-evoked otoacoustic emissions in normal-hearing and hearing-impaired ears. *J Acoust Soc Am*. 1993;93(6):3308-19. <https://doi.org/10.1121/1.405715>
55. Ioannidis, J. P. A. Why Most Published Research Findings Are False. *PLoS Medicine* 2005, 2(8), e124. <https://doi.org/10.1371/journal.pmed.0020124>
56. Konopka W, Pawlaczyk-Luszczynska M, Śliwińska-Kowalska M. The influence of jet engine noise on hearing of technical staff. *Med Pr*. 2014;65(5):583-92. <https://doi.org/10.13075/mp.5893.00045>

Received for publication: 15 October 2017