


Validating self-reporting of hearing-related symptoms against pure-tone audiometry, otoacoustic emission, and speech audiometry

Sofie Fredriksson, Oscar Hammar, Lennart Magnusson, Kim Kähäri & Kerstin Persson Waye


To cite this article: Sofie Fredriksson, Oscar Hammar, Lennart Magnusson, Kim Kähäri & Kerstin Persson Waye (2016): Validating self-reporting of hearing-related symptoms against pure-tone audiometry, otoacoustic emission, and speech audiometry, International Journal of Audiology, DOI: [10.1080/14992027.2016.1177210](https://doi.org/10.1080/14992027.2016.1177210)

To link to this article: <http://dx.doi.org/10.1080/14992027.2016.1177210>

 View supplementary material [↗](#)

 Published online: 19 May 2016.

 Submit your article to this journal [↗](#)

 Article views: 36

 View related articles [↗](#)

 View Crossmark data [↗](#)

Original Article

Validating self-reporting of hearing-related symptoms against pure-tone audiometry, otoacoustic emission, and speech audiometry

Sofie Fredriksson¹, Oscar Hammar¹, Lennart Magnusson², Kim Kähäri² & Kerstin Persson Waye¹

¹Department of Occupational and Environmental Medicine, Institute of Medicine, the Sahlgrenska Academy, University of Gothenburg, Sweden and ²Department of Audiology, Institute of Neuroscience and Physiology, the Sahlgrenska Academy, University of Gothenburg, Sweden



The British Society of Audiology



The International Society of Audiology



Abstract

Objective: To validate self-reported hearing-related symptoms among personnel exposed to moderately high occupational noise levels at an obstetrics clinic. **Design:** Sensitivity, specificity, and predictive values were calculated for questionnaire items assessing hearing loss, tinnitus, sound sensitivity, poor hearing, difficulty perceiving speech, and sound-induced auditory fatigue. Hearing disorder was diagnosed by pure-tone audiometry, distortion product otoacoustic emissions, and HINT (Hearing In Noise Test). **Study sample:** Fifty-five female obstetrics personnel aged 22–63 participated; including 26 subjects reporting hearing loss, poor hearing, tinnitus, or sound sensitivity, and 29 randomly selected subjects who did not report these symptoms. **Results:** The questionnaire item assessing sound-induced auditory fatigue had the best combination of sensitivity $\geq 85\%$ (95% CIs 56 to 100%) and specificity $\geq 70\%$ (95% CIs 55 to 84%) for hearing disorder diagnosed by audiometry or otoacoustic emission. Of those reporting sound-induced auditory fatigue 71% were predicted to have disorder diagnosed by otoacoustic emission. Participants reporting any hearing-related symptom had slightly worse measured hearing. **Conclusions:** We suggest including sound-induced auditory fatigue in questionnaires for identification of hearing disorder among healthcare personnel, though larger studies are warranted for precise estimates of diagnostic performance. Also, more specific and accurate hearing tests are needed to diagnose mild hearing disorder.

Key Words: Validation; sensitivity; questionnaire; hearing-related symptoms; audiometry; otoacoustic emission; speech audiometry; sound-induced auditory fatigue

In a recent report, the European Agency for Safety and Health at Work (2012) argues that risk assessment for noise exposure at work has been largely overlooked in traditionally female-dominated work areas such as the health sector. Carrying out risk assessment for hazardous occupational noise usually includes resource-demanding measurements. Screening with self-report of hearing-related symptoms and noise exposure can be a cost-efficient alternative, under the condition that the questionnaire items used are valid. Validation of screening items should in particular consider the diagnostic performance as estimated by sensitivity, specificity, and predictive values in order to assess how well the questionnaire items can identify individuals with or without disorder and how well self-report predicts a measured disorder (Altman & Bland, 1994a,b).

Validation studies assessing sensitivity and specificity of self-reported hearing-related symptoms have so far focused on hearing loss and hearing disability using pure-tone audiometry as diagnostic

method and predominantly included older age groups (e.g. Clark et al, 1991; Voeks et al, 1993; Nondahl et al, 1998; Sindhusake et al, 2001; Boatman et al, 2007; Deepthi & Kasthuri, 2012). Studies of working-aged populations exposed to occupational noise have predominantly included male participants with focus on self-reported hearing loss using pure-tone audiometry as diagnostic method (e.g. Gomez et al, 2001; Ahmed et al, 2004; Rosso et al, 2011; Hong et al, 2011). Pure-tone audiometry is considered the gold standard diagnostic method for assessing hearing loss. Other symptoms of hearing disorders such as sound sensitivity (hyperacusis), tinnitus, and sound-induced auditory fatigue, which are prevalent among personnel exposed to moderate to high sound levels predominantly from communication (Persson Waye et al, 2010) or music (Kähäri et al, 2003) have largely been overlooked. Additional diagnostic methods may be especially important when seeking to identify mild signs of hearing disorder and symptoms

Correspondence: Sofie Fredriksson, Department of Occupational and Environmental Medicine, Institute of Medicine, the Sahlgrenska Academy, University of Gothenburg PO Box 414, SE 405 30 Gothenburg, Sweden. E-mail: sofie.fredriksson@gu.se

(Received 3 June 2015; revised 1 April 2016; accepted 7 April 2016)

ISSN 1499-2027 print/ISSN 1708-8186 online © 2016 British Society of Audiology, International Society of Audiology, and Nordic Audiological Society
DOI: 10.1080/14992027.2016.1177210

Abbreviations

CI	Confidence interval
dB Cpeak	Decibel C-weighted peak sound pressure level
dB HL	Decibel hearing level
dB LAeq	Decibel A-weighted equivalent sound pressure level
dB SPL	Decibel sound pressure level
DPOAE	Distortion product otoacoustic emission
HINT	Hearing in noise test
NPV	Negative predictive value
PPV	Positive predictive value
PTA	Pure-tone average
SNR	Signal-to-noise ratio

other than hearing loss. Among standard diagnostic tests, distortion product otoacoustic emission (DPOAE) has been suggested to better detect mild signs of noise-induced hearing disorder compared to pure-tone audiometry (e.g. Prasher & Sulkowski 1998; Desai et al, 1998; Hall & Lutman 1999; Attias et al, 2001; Lucertini et al, 2002). A recent study by Engdahl et al (2013) has shown correlation between self-reported hearing disability and otoacoustic emission (OAE), but their analysis suggests OAE may not in general add more than pure-tone audiometry. Furthermore, speech perception in noise tests may be used to assess more complex hearing function associated with the ability to detect amplitude modulation and temporal processing of stimuli at supra-threshold levels (Ruggles et al, 2011), which may reflect noise-induced hearing disorder not showing as a pure-tone hearing loss (Kumar et al, 2012). Tinnitus and sound sensitivity have been related to both noise exposure (Palmer et al, 2002), and pure-tone hearing loss (Anari et al, 1999; Nondahl et al, 2011). To date though, validation studies have focused on the psychometric properties of questionnaires (Kuk et al, 1990; Khalfa et al 2002) rather than on diagnostic performance. Questionnaire items for sound-induced auditory fatigue have not yet been validated, but deemed viable in previous surveys (Persson Waye et al, 2010; Fredriksson et al, 2015). Sound-induced auditory fatigue has been described by sufferers (predominantly preschool and school teachers) as a sensation of fatigue from within the ear, with the affected individual seeking quietness after a day at work in a communication intense sound environment. Based on limited available research we hypothesize the causes of sound-induced auditory fatigue to be multifactorial, involving not only noise exposure and factors in the sound environment, but possibly also work-related stress, cognitive factors, and mental fatigue (Persson Waye et al, 2010; Sjödin et al, 2012; Kristiansen et al, 2014; Fredriksson et al, 2015).

The aim of the present study was to validate questionnaire items corresponding to different hearing-related symptoms reported by women exposed to moderately high occupational noise levels with high intensity speech communication, with the intention of capturing early signs of hearing disorder.

Materials and Methods

This study was approved by the ethics committee in Gothenburg, Sweden, No. 788-11.

Study participants

Fifty-five women were selected from a group of obstetrics personnel (midwives and assistant nurses, $n = 115$) that took part

in a cross-sectional questionnaire study three months prior to the present study (Fredriksson et al, 2015). The mean age of the initial population was 45 (SD: 11) ranging from 22 to 65 years and they had on average worked 12 years in obstetrics care (SD: 11) ranging from half a year to 40 years. The sound environment in the obstetrics ward was characterized by intensive speech communication and loud screams by mothers during delivery, and almost half of the 62 measured work shifts had A-weighted equivalent levels above 80 dB LAeq. Description of participants' age and years worked in obstetrics is shown in Table 1. The selection of participants for the current study, illustrated in Figure 1, was stratified such that it included 26 subjects out of 34 who had reported self-reported hearing loss, poor hearing, tinnitus, or sound sensitivity and 29 randomly selected subjects from 55 who had not reported those symptoms and who had also not reported difficulty perceiving speech. Subjects from both groups could report sound-induced auditory fatigue. The stratified selection of participants was used to ensure inclusion of individuals with probable measurable hearing disorder. Individuals in the initial study population who reported only difficulty perceiving speech or individuals who had missing data on hearing-related questionnaire items were not included in the present study. Lastly, one individual with diagnosed otosclerosis was excluded as the intention was to measure DPOAE amplitudes reflecting the status of the inner ear.

Questionnaire items

The questionnaire items assessed in this study are presented in detail in the supplementary table available in the online version of the journal. In the initial cross-sectional study (Fredriksson et al, 2015), self-reported hearing-related symptoms (hearing loss, tinnitus, sound sensitivity, poor hearing, difficulty perceiving speech, and sound-induced auditory fatigue) were collected via paper and computer questionnaire. No significant differences ($p > 0.05$) of reported hearing-related symptoms were seen between the response methods in a chi-square analysis (Fredriksson et al, 2015). The questionnaire items were constructed based on our previous studies (Persson Waye et al, 2010) or were modified from previously validated questionnaires (Nondahl et al, 1998; Kuk et al, 1990; Khalfa et al, 2002). In the analysis, the definition of reporting each symptom was as follows: reporting 'yes' for hearing loss, reporting 'once a week' or more often for tinnitus, sound sensitivity and sound-induced auditory fatigue, reporting 'bad' or 'very bad' for poor hearing and reporting 'yes' both at work and leisure time for difficulty perceiving speech. Additionally, a combined questionnaire item, which corresponds to reporting one or more of the six hearing-related symptoms, was also assessed. The questionnaire items assessing hearing loss, poor hearing, tinnitus, sound sensitivity, and speech perception were the same as those used in the selection of participants for the current study.

Hearing tests

The hearing assessment was performed in the selected study sample three months after the initial cross-sectional questionnaire survey. The assessment included otoscopy, tympanometry, and standard hearing tests: pure-tone audiometry, DPOAE, and HINT. The tests were performed by a licensed audiologist (first author) and two experienced occupational health care nurses trained for the project. The six-hour A-weighted equivalent sound level in the test room was 30 dB LAeq measured using Bruel & Kjaer 2260 with the

Table 1. Description of study participants.

	Participants with self-reported hearing loss, poor hearing, tinnitus, or sound sensitivity, $n = 26^a$			Participants without self-reported hearing loss, poor hearing, tinnitus or sound sensitivity, $n = 29^b$		
	Mean	SD	Range	Mean	SD	Range
Age, years	49	10	28 to 63	45	10	22 to 62
Worked in obstetrics, years	17	11	1 to 40	13	11	1 to 39
Audiometry left ear, dB HL ^c						
0.25 kHz	12	6	10 to 35	11	3	10 to 20
0.5 kHz	13	8	10 to 40	11	3	10 to 25
1 kHz	13	6	10 to 30	12	4	10 to 25
2 kHz	15	10	10 to 50	12	4	10 to 25
3 kHz	16	12	10 to 60	12	3	10 to 20
4 kHz	17	12	10 to 55 ^d	11	3	10 to 25
6 kHz	23	18	10 to 65 ^d	14	7	10 to 35
8 kHz	28	18	10 to 65 ^d	16	6	10 to 30
DPOAE left ear, dB SPL						
2–3.9 kHz	3	5	–4 to 15	6	5	–2 to 16
4–10 kHz	–3	4	–8 to 4	4	6	–4 to 20
HINT, dB SNR	–4	2	–9 to 2	–5	1	–10 to –3

^aOne subject had all four criteria symptoms, three subjects had three of the criteria symptoms, ten subjects had two of the criteria symptoms, and the remaining twelve had one of the criteria symptoms (tinnitus being most common). Also, seventeen subjects had sound-induced auditory fatigue.

^bSix subjects had sound-induced auditory fatigue.

^cPure-tone audiometry was measured at the lowest level of 10 dB HL. Hence, mean thresholds may appear higher in this sample compared to pure-tone thresholds measured down to –10 dB HL.

^dOne participant had thresholds above maximum output level of the audiometer.

microphone placed in the position of the test subjects head. Intermittent sounds were detected in the frequency range 0.25 to 1.6 kHz with one-third octave maximum level (LAFmax) 1 to 9 dB higher than recommended levels (ISO 8253–1, 1994), but only during short intervals. At most during three connected 30-second logs. Therefore, as a precaution, dampening insert earphones were used for DPOAE and Sennheiser HDA200 circumaural headphones were used for pure-tone audiometry and HINT. Calibration according to international standard (ISO 389-8, 2004) was ensured for HDA200. As an additional precaution, air conduction pure-tone audiometry was performed at a lowest presentation level of 10 dB HL and bone conduction was not performed. Also, testing was not performed when intermittent sounds were heard by the operator. The Grason-Stadler GSI 61 with a DVD-player connected to the audiometer was used for pure-tone audiometry and HINT. The pure-tone audiometry method essentially followed the Hughson Westlake ascending method (Carhart & Jerger 1959) at the standard audiometric frequencies 0.25, 0.5, 1, 2, 3, 4, 6, and 8 kHz, but with a fixed lowest presentation level of 10 dB HL. Masking of the opposite ear was not performed. HINT was measured binaurally using the Swedish version (Hällgren et al, 2006), with noise fixed at 65 dB SPL and the speech signal (female speaker) adaptively adjusted in 2-dB steps, starting at 0 dB signal-to-noise ratio (SNR). Two participants' results were excluded from HINT analysis as they did not have Swedish as their first language. DPOAE was measured in both ears using custom equipment generating two stimulus tones at 65 and 55 dB SPL (L1 and L2) with a frequency ratio (f_2/f_1) fixed at 1.23. DPOAE was registered as the cubic distortion product ($2f_1-f_2$) at 32 sets of primary input tones with f_2 ranging from 0.707 kHz to 10.374 kHz. Technical specifications can be found in the supplementary material available in the online version of the journal.

Eleven participants' results were excluded from the analysis of DPOAE data: seven due to negative tympanic peak pressure (≤ -100 daPa) or low middle-ear admittance (≤ 0.3 mmho) found in tympanometry, and four due to background noise during DPOAE measurement (average noise floor >2 SD from mean in frequencies below 2 kHz). The Otometrics MADSEN OTOflex 100 was used for tympanometry. Individuals with excessive cerumen were rescheduled and tested after the cerumen was removed. Description of hearing tests data underlying the analysis is shown in Table 1.

Definitions of diagnostic criteria

For statistical purposes, diagnosed hearing disorder was defined using two criteria for each of the three hearing tests: a strict cut-off value representing at least mild hearing disorder and a less strict cut-off value representing slightly more pronounced hearing disorder (worse hearing test result). Both cut-offs however were set with the intention of capturing mild signs of disorder. The cut-off values for failing the less strict criterion for each hearing test were as follows: one or more pure-tone thresholds at ≥ 40 dB HL in the 0.25 to 8 kHz range in either ear (labelled 40 dB HL), DPOAE SNR <3 dB in either one or both of the f_2 range 2 to 3.9 kHz and 4 to 10 kHz in either ear (labelled 3 dB SNR), and HINT SNR > -3 dB (labelled –3 dB SNR). The cut-off values for failing the strict criterion for each of the three hearing tests were as follows: two or more pure-tone thresholds ≥ 25 dB HL or one or more pure-tone thresholds ≥ 30 dB HL in the 0.25 to 8 kHz range in either ear (labelled 25/30 dB HL), DPOAE SNR <6 dB in either one or both of the f_2 ranges 2 to 3.9 kHz and 4 to 10 kHz in either ear (labelled 6 dB SNR), and HINT SNR > -4 dB (labelled –4 dB SNR).

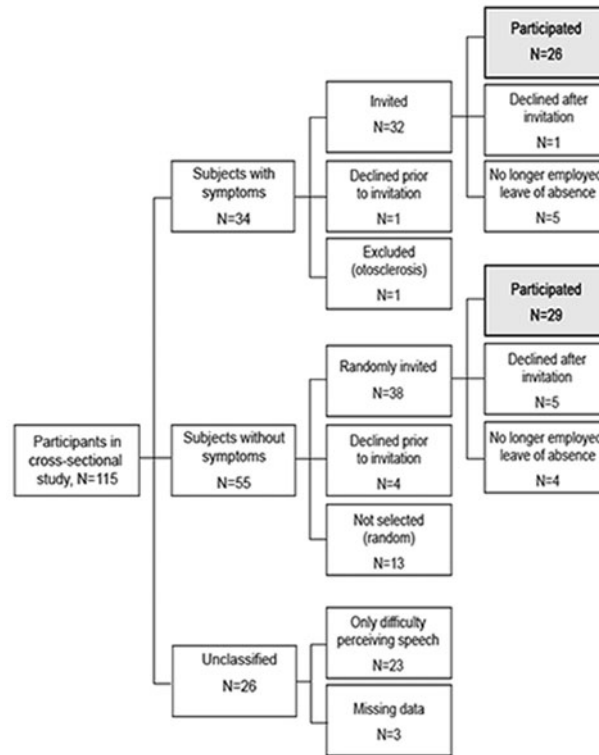


Figure 1. Flow chart illustrating the selection of the study sample.

Statistical analysis

The diagnostic performance of questionnaire items was assessed by calculating sensitivity and specificity and predictive values against standard diagnostic hearing tests. Based on the two diagnostic cut-off values for each hearing test and the definition of self-reported hearing-related symptoms using the questionnaire items, a subject was defined as either true positive, true negative, false positive, or false negative. Sensitivity was calculated as the proportion of correctly identified subjects (true positives) among all subjects who failed the diagnostic cut-off. Specificity was calculated as the proportion of subjects that was correctly dismissed (true negatives) among all subjects who passed the diagnostic cut-off (Altman & Bland, 1994a). Positive predictive value (PPV) was calculated as the proportion of true positives among all who reported the symptom, and negative predictive value (NPV) was calculated as the proportion of true negatives among all subjects that had not reported the symptom (Altman & Bland, 1994b). In order to obtain results representative for the underlying study population, the calculation of diagnostic performance of failing each diagnostic cut-off were weighted based on the stratified sampling of subjects with and without hearing-related symptoms (hearing loss, poor hearing, tinnitus, or sound sensitivity). As such, the weights balanced out the more frequent sampling of subjects with symptoms ($n = 26$, 47% of 34 in initial population) than subjects without those symptoms ($n = 29$, 53% of 55 in initial population), see Figure 1 for details. Weighted prevalence was calculated to reflect the prevalence in the underlying cross-sectional sample. Specific details for the calculation using weights can be found in the online supplementary material available in the online version of the journal. Standard

techniques for computing confidence intervals for the sensitivity, specificity, and predictive values were not applicable due to the use of weights. Therefore, we used percentile bootstrap to compute approximate confidence intervals (Efron, 1979), and 95% confidence intervals were generated by using 100 000 samples. Comparisons between groups were analysed using an independent sample t-test with Bonferroni correction for multiple tests. The significance level was set at 5% ($p = 0.05$) for all hypothesis tests.

Results

Diagnostic performance of questionnaire items

The diagnostic performance of questionnaire items are given in Table 2 for the less strict diagnostic cut-off values representing moderate or worse hearing disorder and for the strict diagnostic cut-off values representing at least mild hearing disorder. The questionnaire item corresponding to sound-induced auditory fatigue achieved the highest diagnostic performance, taking both sensitivity and specificity into account with an emphasis on high sensitivity. For the less strict diagnostic criteria sensitivity was 85 and 89% and specificity 70% in comparison to pure-tone audiometry and DPOAE. It was also the only item with a confidence interval ranging between 50–100%. The item assessing tinnitus achieved the second-highest estimate with sensitivity 78% and specificity 88% in comparison to pure-tone audiometry. The item assessing hearing loss had the third highest performance with sensitivity 67% and specificity 87% in comparison to pure-tone audiometry. However, confidence intervals for sensitivity extended below 50% for both

Table 2. Weighted calculations of diagnostic performance (sensitivity and specificity and 95% confidence interval) and weighted prevalence of diagnostic criteria.

	Diagnostic criteria, cut-off value					
	Audiometry		DPOAE ^a		HINT ^a	
	40 dB HL	25/30 dB HL	3 dB SNR	6 dB SNR	-3 dB SNR	-4 dB SNR
Prevalence, %	13	34	17	45	6	33
Hearing loss ^b						
Sensitivity, %	67 (0–100)	37 (15–60)	36 (0–75)	24 (6–44)	50 (0–100)	22 (5–43)
Specificity, %	87 (77–96)	93 (84–100)	88 (76–97)	89(76–100)	84 (73–94)	85 (72–95)
Tinnitus ^b						
Sensitivity, %	78 (43–100)	35 (14–57)	33 (0–67)	29 (11–50)	75 (0–100)	28 (8–50)
Specificity, %	88 (78–96)	87 (75–97)	84 (72–95)	89 (76–100)	84 (73–93)	84 (71–95)
Sound sensitivity ^b						
Sensitivity, %	44 (0–83)	26 (7–47)	22 (0–56)	25 (8–45)	25 (0–100)	28 (8–50)
Specificity, %	88 (78–96)	89 (78–98)	84 (72–95)	89 (76–100)	84 (73–93)	89 (77–97)
Poor hearing ^b						
Sensitivity, %	44 (0–83)	22 (5–41)	44 (10–80)	29 (11–50)	50 (0–100)	37 (15–60)
Specificity, %	85 (74–94)	82 (69–94)	82 (69–93)	82 (67–96)	82 (71–92)	89 (77–97)
Difficulty perceiving speech ^b						
Sensitivity, %	56 (17–100)	30 (11–52)	33 (0–70)	38 (18–59)	50 (0–100)	42 (19–65)
Specificity, %	81 (70–92)	80 (67–92)	77 (63–89)	86 (71–97)	79 (67–89)	86 (74–97)
Sound-induced auditory fatigue ^b						
Sensitivity, %	89 (60–100)	60 (38–81)	85 (56–100)	62 (41–82)	75 (0–100)	53 (29–76)
Specificity, %	70 (57–82)	74 (59–87)	70 (55–84)	79 (62–93)	63 (50–76)	68 (52–82)
Reporting any symptom ^b						
Sensitivity, %	100 (NA ^c)	69 (47–89)	85 (56–100)	66 (45–86)	100 (NA ^c)	67 (44–88)
Specificity, %	56 (43–70)	58 (42–74)	56 (41–71)	61 (42–80)	52 (38–65)	56 (40–72)

^aExcluded from DPOAE: seven subjects due to abnormal tympanometry, four subjects due to high background noise levels during measurement, and from HINT: two subjects who did not have Swedish as their first language.

^bNumber of positive responses on questionnaire items, hearing loss: $n = 7$ for audiometry and HINT, $n = 5$ for DPOAE; tinnitus: $n = 14$ for audiometry, $n = 10$ for DPOAE, $n = 13$ for HINT; Sound sensitivity: $n = 11$ for audiometry and HINT, $n = 9$ for DPOAE; $n = 13$ for audiometry and HINT, $n = 12$ for DPOAE; difficulty perceiving speech: $n = 16$ for audiometry, $n = 13$ for DPOAE, $n = 15$ for HINT; sound-induced auditory fatigue: $n = 23$ for audiometry and HINT, $n = 19$ for DPOAE; any symptom: $n = 32$ for audiometry, $n = 25$ for DPOAE, $n = 31$ for HINT.

^cNA = not applicable, confidence interval for estimates at 100% are not reliable.

hearing loss and tinnitus. The combined item, which represents self-reporting of one or more of the six individual symptoms had the highest sensitivity, but a low accompanying specificity and also wide confidence intervals. None of the questionnaire items performed satisfactorily for the strict diagnostic criteria, with particular regard to sensitivity. As seen in Table 3, predictive values were higher for strict than for the less strict cut-offs but none of the items had satisfying predictive values. For the strict diagnostic criteria, the item corresponding to hearing loss had the best combination of PPV and NPV; with PPV of 71% and NPV of 76% in comparison with pure tone audiometry, but with a wide confidence interval for PPV (38–100% for PPV and 63–88% for NPV). Sound-induced auditory fatigue had a PPV and NPV both at 71% and confidence interval 50–90% for PPV and 54–87% for NPV in comparison with DPOAE.

Comparing the groups with and without symptoms

A comparison was done of the results from the standard diagnostic hearing test between subjects reporting one or more of the six hearing-related symptoms and subjects reporting none of the symptoms (neither hearing loss, poor hearing, tinnitus or sound

sensitivity nor difficulty perceiving speech or sound-induced auditory fatigue). The mean age for subjects reporting any of the six symptoms was higher than for subjects not reporting symptoms, 50 years compared to 43 years ($p < 0.05$). The subjects reporting symptoms generally had worse results on the three hearing tests. As seen in Figure 2, mean thresholds at 4, 6, and 8 kHz were worse among subjects reporting symptoms and significantly so in both ears ($p < 0.05$) compared to subjects not reporting symptoms. After Bonferroni correction for multiple tests, significant difference remained for 6 kHz in both ears ($p = 0.006$ for right ear, and $p = 0.036$ for left ear) and 8 kHz left ear ($p = 0.018$). As seen in Figure 3(A), the mean amplitudes for DPOAE were somewhat lower among subjects reporting hearing-related symptoms for both f2 ranges 2 to 3.9 kHz and 4 to 10 kHz and significantly so for the higher frequency range ($p < 0.05$). After Bonferroni correction for multiple tests, only the left ear had significantly lower average amplitude in the 4 to 10 kHz range ($p = 0.048$). As seen in Figure 3(B), subjects reporting any symptom tended to require a higher level of the speech signal in relation to background noise, as compared to subjects reporting no symptoms. However, the difference was not significant ($p = 0.092$). Considering only the item speech perception, we did nonetheless find significantly worse

Table 3. Weighted calculations of diagnostic performance (PPV, NPV, and 95% confidence interval) and weighted prevalence of diagnostic criteria.

	Diagnostic criteria, cut-off value					
	Audiometry		DPOAE ^a		HINT ^a	
	40 dB HL	25/30 dB HL	3 dB SNR	6 dB SNR	-3 dB SNR	-4 dB SNR
Prevalence, %	13	34	17	45	6	33
Hearing loss ^b						
PPV, %	29 (0-63)	71 (38-100)	40 (0-80)	60 (20-100)	14 (0-40)	43 (11-75)
NPV, %	97 (91-100)	76 (63-88)	87 (75-97)	63 (47-78)	97 (91-100)	68 (53-81)
Tinnitus ^b						
PPV, %	50 (20-80)	57 (27-86)	30 (0-64)	70 (36-100)	23 (0-50)	46 (17-77)
NPV, %	96 (90-100)	72 (59-84)	86 (74-95)	60 (45-76)	98 (93-100)	70 (56-83)
Sound sensitivity ^b						
PPV, %	36 (0-71)	55 (20-88)	22 (0-57)	67 (30-100)	9 (0-33)	55 (20-88)
NPV, %	91 (83-98)	70 (57-83)	84 (71-95)	59 (44-74)	94 (87-100)	71 (58-84)
Poor hearing ^b						
PPV, %	31 (0-60)	38 (10-70)	33 (7-64)	58 (27-88)	15 (0-40)	62 (30-90)
NPV, %	91 (82-98)	67 (53-80)	88 (76-97)	58 (42-74)	96 (90-100)	74 (60-86)
Difficulty perceiving speech ^b						
PPV, %	31 (8-58)	44 (17-71)	23 (0-50)	69 (40-93)	13 (0-36)	60 (31-88)
NPV, %	92 (84-100)	69 (55-82)	85 (72-95)	62 (46-78)	96 (89-100)	75 (61-87)
Sound-induced auditory fatigue ^b						
PPV, %	31 (13-52)	54 (33-75)	37 (16-60)	71 (50-90)	12 (0-100)	45 (24-67)
NPV, %	98 (91-100)	78 (64-91)	96 (87-100)	71 (54-87)	97 (91-100)	74 (59-88)
Reporting any symptom ^b						
PPV, %	26 (11-42)	46 (28-64)	29 (12-48)	59 (39-78)	12 (0-25)	43 (25-62)
NPV, %	100 (NA ^c)	78 (62-93)	95 (84-100)	68 (48-87)	100 (NA ^c)	77 (60-92)

^aExcluded from DPOAE: seven subjects due to abnormal tympanometry, four subjects due to high background noise levels during measurement; and from HINT: two subjects who did not have Swedish as their first language.

^bNumber of positive responses on questionnaire items, hearing loss: $n = 7$ for audiometry and HINT, $n = 5$ for DPOAE; tinnitus: $n = 14$ for audiometry, $n = 10$ for DPOAE, $n = 13$ for HINT; Sound sensitivity: $n = 11$ for audiometry and HINT, $n = 9$ for DPOAE; $n = 13$ for audiometry and HINT, $n = 12$ for DPOAE; difficulty perceiving speech: $n = 16$ for audiometry, $n = 13$ for DPOAE, $n = 15$ for HINT; sound-induced auditory fatigue: $n = 23$ for audiometry and HINT, $n = 19$ for DPOAE; any symptom: $n = 32$ for audiometry, $n = 25$ for DPOAE, $n = 31$ for HINT.

^cNA = not applicable, confidence interval for estimates at 100% are not reliable.

results among individuals who reported difficulty perceiving speech, mean -3.6 dB SNR, compared to those not reporting difficulty perceiving speech, mean -4.7 dB SNR ($p = 0.049$).

Discussion

In the investigated study population, we showed that the questionnaire item assessing sound-induced auditory fatigue had the best diagnostic performance. The item identified 85 to 89% of study subjects with moderate signs of hearing disorder diagnosed by pure-tone audiometry or DPOAE, and correctly dismissed 70% of study subjects without disorder. The items assessing tinnitus and hearing loss performed next best. These items identified almost 70 to 80% of subjects within the study sample with moderate disorder diagnosed by pure-tone audiometry, and correctly dismissed almost 90% of subjects without disorder. Hence, the results point to the importance of including not only hearing loss and tinnitus but also measures of sound-induced auditory fatigue among health-care personnel working in communication-intense environments. A general observation though is that the precision of the estimates were affected by the limited sample size and the 95% confidence intervals were therefore wide. Of the mentioned items only sound-

induced auditory fatigue had a 95% confidence interval for sensitivity of less than 50% width, while the confidence intervals for specificity were narrower. The effect on precision was particularly obvious for sensitivity for the HINT criteria -3 dB SNR with a confidence interval from 0 to 100%, due to a very small sample of subjects failing that cut-off. Caution is thus needed in drawing conclusions from the results, and further studies with larger samples are warranted.

The diagnostic performances, particularly sensitivity, have varied in previous studies of self-reported hearing loss against pure-tone audiometry (Clark et al, 1991; Nondahl et al, 1998; Sindhusake et al, 2001; Ahmed et al, 2004; Rosso et al, 2011; Deepthi & Kasthuri 2012; Hong et al, 2011), and few studies have reported sensitivity and specificity in combination exceeding 80% for self-reporting of hearing loss in questionnaires. Comparison of our results to previous studies is hampered by differences in study populations, diagnostic criteria, and questionnaire items. Nonetheless, our less strict cut-off for pure-tone audiometry (one or more thresholds ≥ 40 dB HL) may be considered comparable to a PTA of ≥ 25 dB HL, and as such, our results for the item hearing loss agrees rather well with the results for 'at least a mild hearing loss' as assessed by Nondahl et al (1998). Given the similarity of

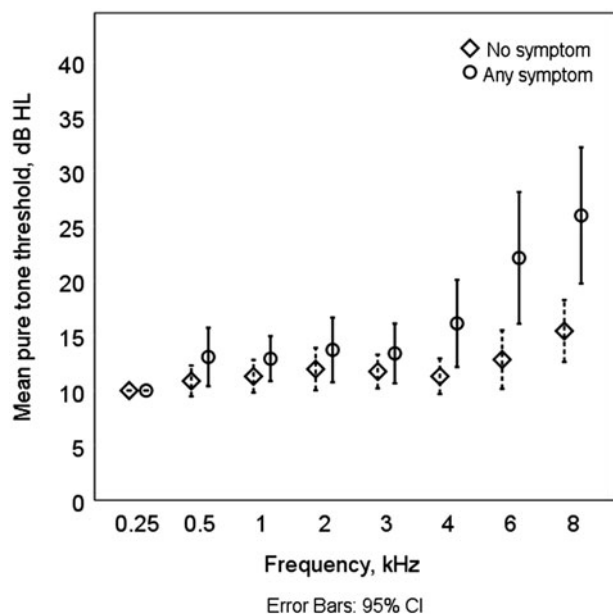


Figure 2. Mean pure-tone thresholds in dB HL and 95% CI of mean (error bar) for audiometric frequencies 0.25 – 8 kHz, left ear, for subjects reporting no symptoms (◇) and subjects reporting any hearing related symptom (○).

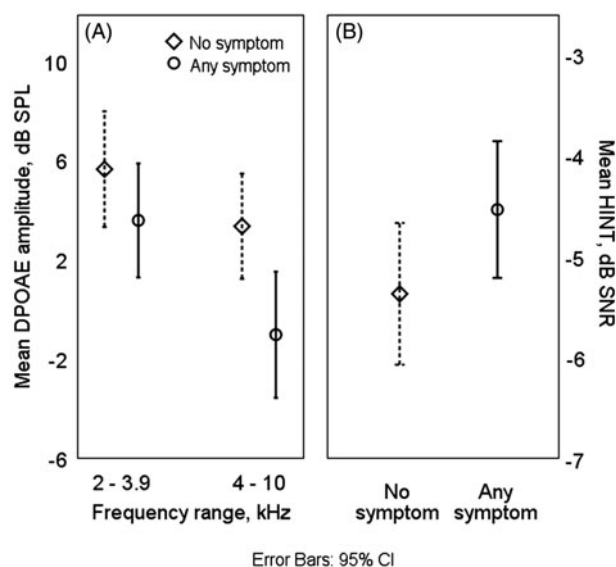


Figure 3. In A, mean amplitude of distortion product emission in dB SPL and 95% CI of mean (error bar) averaged over f_2 ranges 2–4 kHz and >4–10 kHz, left ear. In B, mean hearing in noise test in dB SNR and 95% CI of mean (error bar), for subjects reporting no symptoms (◇) and subjects reporting any hearing related symptom (○).

results, the validity of our findings is strengthened despite our limited study sample. Important for the assessment of diagnostic performance is also the fact that the desired level of sensitivity and specificity will differ depending on the intended use of a screening test. For example, Boatman et al (2007) refer to 80% as an

acceptable level of the ability of bedside tests to diagnose hearing loss, while Nondahl et al (1998) find a sensitivity of 71% acceptable for self-reporting in epidemiological surveys. If our results are correct, our questionnaire items assessing sound-induced auditory fatigue, hearing loss, and tinnitus may be used with reasonable satisfaction for identifying fairly moderate hearing disorder. The combined item (reporting any hearing-related symptom) resulted in very high sensitivity, but low specificity. Studies have shown increased sensitivity when a combination of tests or items are used (Boatman et al, 2007; Rosso et al, 2011), but as in our study the approach often entails low specificity. Therefore it is important to assess both sensitivity and specificity in combination. Nevertheless, the high sensitivity for reporting any symptom is supported by the result showing that those subjects had indications of worse measured hearing compared to subjects reporting none of the symptoms.

For the purpose of the analysis of diagnostic performance, the cause (such as age or noise exposure) is not as important as the ability of the individuals themselves to detect and report hearing loss. In general, measures of self-report may be better at identification of individuals with severe hearing disorder compared to mild signs of disorder. In our study this was reflected by higher performance of items in relation to the less strict diagnostic cut-offs, while the strict (more ‘difficult’) cut-off, for example, could mean having just one pure-tone threshold at 30 dB HL or worse, which may not be noticeable or reported by the individual. In addition, diagnostic performance may be affected by the diagnostic tests used. Standard hearing tests may be unable to diagnose early and mild signs of hearing disorder (Lindblad et al 2011), hypothesized to be reported as hearing-related symptoms. The accuracy of the underlying assumption of an association between these hearing-related symptoms and the standard hearing tests used is important for the interpretation of our results. Such associations have been reported in previous studies mainly for self-reported hearing loss or hearing disability (e.g. Pedersen & Rosenhall, 1991, Engdahl et al, 2013, Videhult Pierre et al, 2015). Notably, the item assessing difficulty perceiving speech did not achieve a satisfying sensitivity in our study sample using the speech perception test, HINT. This may be due to the HINT not having been developed as a diagnostic test of hearing disorder but rather for assessment of rehabilitation benefit (Nilsson et al, 1994; Hällgren et al, 2006). Our results could also be due to the cut-off, which was derived from average results of normal-hearing listeners (Nilsson et al, 1994; Hällgren et al, 2006). Self-perceived poor hearing and sound sensitivity both failed to reach a high sensitivity in our study sample, which if correct would render them less suitable in screening for measured hearing disorder in a similar population. Again though, it is important to consider that the hearing tests and the cut-offs used may not be appropriate for diagnosing all forms of self-reported hearing-related symptoms. This was pointed out by Sztuka et al (2010), who showed that tinnitus patients who also reported hyperacusis had increased instead of decreased DPOAE amplitude, which would not be shown as a disorder using our SNR based cut-off.

Predictive values assess the clinical use of a screening test. In our study, the positive prediction was 71% for both the item hearing loss as diagnosed by pure-tone audiometry and for the item sound-induced auditory fatigue as diagnosed by DPOAE using the strict cut-off value. Only sound-induced auditory fatigue had a reasonably narrow confidence interval. The result for hearing loss is

comparable to Nondahl et al (1998) who found a PPV of 68%. The diagnose criterion however differ from our study. Importantly, predictive values are influenced by diagnosed prevalence and therefore specific to the population under study. Therefore, our weighted prevalence which reflects the prevalence in the underlying cross-sectional sample is important to consider. The prevalence could explain why items in our study generally showed higher predictive values for a strict criterion with higher prevalence, while they generally had higher sensitivity for the less strict cut-offs. As Altman and Bland (1994b) point out, when prevalence is low in the population being screened, even a test with high sensitivity and specificity would incorrectly classify some individuals as having the disorder when in fact they do not (false positives).

According to what may be expected, a hearing disorder diagnosed by pure-tone audiometry (i.e. measured hearing loss) was best predicted in our sample by the item assessing hearing loss. Interestingly, DPOAE best predicted sound-induced auditory fatigue and tinnitus. Studies of tinnitus support abnormal DPOAE responses (e.g. Shiomi et al, 1997; Sztuka et al, 2010) and it is possible that inner ear pathology (as measured with DPOAE) is also involved in the aetiology or comorbidity of sound-induced auditory fatigue. However, further studies are warranted to support the hypothesis and more specific and accurate diagnostic tests are clearly needed to assess the symptoms reported by subjects working in a moderately noise environment with intensive speech communication. Promising possibilities of assessing minor auditory dysfunction have recently been suggested by Lindblad et al (2011).

As a limitation, our results may not be applicable to an unselected population as the initial survey population was specific and limited in size. The present analysis included a sub-sample of the initial population, and due to the beforehand chosen inclusion criteria subjects who reported only difficulty perceiving speech were excluded. We did however find similarities to large-scale population studies (e.g. Nondahl et al, 1998), which strengthen our findings.

Conclusions

The study indicated that self-reporting of hearing-related symptoms may identify individuals with fairly moderate hearing disorder, but not those with mild signs of disorder. In our sample of female obstetrics personnel working in moderately high noise levels with occasional loud screams, the best combination of sensitivity and specificity was achieved using the item sound-induced auditory fatigue. The item identified almost 90% of subjects with hearing disorder diagnosed using standard hearing tests pure-tone audiometry or DPOAE, and correctly dismissed 70% of subjects without a diagnosed hearing disorder. The items assessing tinnitus and hearing loss had the next best performance. However, the small study sample resulted in uncertain estimates. Sound-induced auditory fatigue was best predicted by failing the 6 dB SNR cut-off for DPOAE while self-reported hearing loss was best predicted by having at least a mild pure-tone hearing loss (25–30 dB HL or more). It would thus be interesting to study further whether inner ear pathology is involved in the mechanisms for sound-induced auditory fatigue.

In combination with our previous research, this study indicates that asking about sound-induced auditory fatigue in addition to hearing loss and tinnitus could make a valuable contribution to the assessment of hearing-related symptoms among individuals in

occupational sound environments with intense speech communication such as health care settings. Finally, the study shows that the present questionnaire items fail to identify very mild hearing disorders and the present standard hearing tests fail to diagnose some self-reported hearing-related symptoms. There is a need for continued research improving detection, diagnosis, and prevention of hearing disorder and hearing-related symptoms.

Acknowledgements

The authors would like to thank Lisbeth Birkhed and Johan Magnusson at the Occupational Healthcare Unit Hälsan & Arbetstlivet for assistance in measurements, participating personnel at the obstetrics care unit and especially Gullan Ahlgren for administrative support. This study was funded by grants from The Swedish Research Council for Health, Working Life and Welfare (2009-1741) and the Swedish organization AFA Insurance (090161).

Declaration of interest: The authors report no declarations of interest. All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare they received no support from any organization for the submitted work; have no financial relationships with any organizations that might have an interest in the submitted work in the previous three years; and have no other relationships or activities that could appear to have influenced the submitted work.

References

- Ahmed H.O., Dennis J.H. & Ballal S.G. 2004. The accuracy of self-reported high noise exposure level and hearing loss in a working population in Eastern Saudi Arabia. *Int J Hyg Environ Health*, 207, 227–234.
- Altman D.G. & Bland J.M. 1994a. Statistics notes: Diagnostic tests 1: Sensitivity and specificity. *BMJ*, 308, 1552.
- Altman D.G. & Bland J.M. 1994b. Statistics Notes: Diagnostic tests 2: Predictive values. *BMJ*, 309, 102.
- Anari M., Axelsson A., Eliasson A. & Magnusson L. 1999. Hypersensitivity to sound-questionnaire data, audiometry and classification. *Scand Audiol*, 28, 219–230.
- Attias J., Horovitz G., El-Hatib N. & Nageris B. 2001. Detection and clinical diagnosis of noise-induced hearing loss by otoacoustic emissions. *Noise Health*, 3, 19–31.
- Boatman D., Miglioretti D., Eberwein C., Alidoost M. & Reich S. 2007. How accurate are bedside hearing tests? *Neurology*, 68, 1311–1314.
- Carhart R. & Jerger J.F. 1959. Preferred method for clinical determination of pure-tone thresholds. *J Speech Hear Disord*, 24, 330–345.
- Clark K., Sowers M., Wallace R.B. & Anderson C. 1991. The accuracy of self-reported hearing loss in women aged 60–85 years. *Am J Epidemiol*, 134, 704–708.
- Deepthi R. & Kasthuri A. 2012. Validation of the use of self-reported hearing loss and the hearing handicap inventory for elderly among rural Indian elderly population. *Arch Gerontol Geriatr*, 55, 762–767.
- Desai A., Reed D., Cheyne A., Richards S. & Prasher D. 1998. Absence of otoacoustic emissions in subjects with normal audiometric thresholds implies exposure to noise. *Noise Health*, 1, 58–65.
- Efron B. 1979. Bootstrap methods: Another look at the jackknife. *Ann Stat*, 7, 1–26.
- Engdahl B., Tambs K. & Hoffman H.J. 2013. Otoacoustic emissions, pure-tone audiometry, and self-reported hearing. *Int J Audiol*, 52, 74–82.

- European Agency for Safety and Health at Work. 2012. New risks and trends in the safety and health of women at work. Available: <https://osha.europa.eu/en/publications/reports/new-risks-and-trends-in-the-safety-and-health-of-women-at-work>
- Fredriksson S., Hammar O., Torén K., Tenenbaum A. & Persson Waye K. 2015. The effect of occupational noise exposure on tinnitus and sound-induced auditory fatigue among obstetrics personnel: A cross-sectional study. *BMJ Open*, 5, e005793.
- Gomez M.L., Hwang S.A., Sobotova L., Stark A.D. & May J.J. 2001. A comparison of self-reported hearing loss and audiometry in a cohort of New York farmers. *J Speech Lang Hear Res*, 44, 1201–1208.
- Hall A.J. & Lutman M.E. 1999. Methods for early identification of noise-induced hearing loss. *Audiology*, 38, 277–280.
- Hong O., Ronis D.L. & Antonakos C.L. 2011. Validity of self-rated hearing compared with audiometric measurement among construction workers. *Nurs Res*, 60, 326–332.
- Hällgren M., Larsby B. & Arlinger S. 2006. A Swedish version of the hearing in noise test (HINT) for measurement of speech recognition. *Int J Audiol*, 45, 227–237.
- ISO 389-8. *Acoustics: Reference zero for the calibration of audiometric equipment. Part 8: Reference equivalent threshold sound pressure levels for pure tones and circumaural earphones*. Geneva: International Organization for Standardization, 2004.
- ISO 8253-1. *Acoustics – Audiometric Test Methods. Part 1. Basic Pure Tone Air and Bone Conduction Threshold Audiometry*. Geneva: International Organization for Standardization, 1994.
- Khalifa S., Dubal S., Veuillet E., Perez-Diaz F., Jouvent R. et al. 2002. Psychometric normalization of a hyperacusis questionnaire. *ORL J Otorhinolaryngol Relat Spec*, 64, 436–442.
- Kristiansen J., Lund S.P., Persson R., Shibuya H., Nielsen P.M. et al. 2014. A study of classroom acoustics and school teachers' noise exposure, voice load and speaking time during teaching, and the effects on vocal and mental fatigue development. *Int Arch Occup Environ Health*, 87, 851–860.
- Kuk F.K., Tyler R.S., Russell D. & Jordan H. 1990. The psychometric properties of a tinnitus handicap questionnaire. *Ear Hear*, 11, 434–445.
- Kumar U.A., Ameenudin S. & Sangamanatha A.V. 2012. Temporal and speech processing skills in normal hearing individuals exposed to occupational noise. *Noise Health*, 14, 100–105.
- Kähäri K., Zachau G., Eklöf M., Sandsjö L. & Möller C. 2003. Assessment of hearing and hearing disorders in rock/jazz musicians. *Int J Audiol*, 42, 279–288.
- Lindblad A.C., Hagerman B. & Rosenhall U. 2011. Noise-induced tinnitus: A comparison between four clinical groups without apparent hearing loss. *Noise Health*, 13, 423–431.
- Lucertini M., Moleti A. & Sisto R. 2002. On the detection of early cochlear damage by otoacoustic emission analysis. *J Acoust Soc Am*, 111, 972–978.
- Nilsson M., Soli S.D. & Sullivan J.A. 1994. Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am*, 95, 1085–1099.
- Nondahl D.M., Cruickshanks K.J., Wiley T.L., Tweed T.S., Klein R. et al. 1998. Accuracy of self-reported hearing loss. *Audiology*, 37, 295–301.
- Nondahl D.M., Cruickshanks K.J., Huang G.H., Klein B.E., Klein R. et al. 2011. Tinnitus and its risk factors in the beaver dam offspring study. *Int J Audiol*, 50, 313–320.
- Palmer K., Griffin M., Syddall H., Davis A., Pannett B. et al. 2002. Occupational exposure to noise and the attributable burden of hearing difficulties in Great Britain. *Occup Environ Med*, 59, 634–639.
- Pedersen K. & Rosenhall U. 1991. Correlations between self-assessed hearing handicap and standard audiometric tests in elderly persons. *Scand Audiol*, 20, 109–116.
- Persson Waye K., Agge A., Hillström J. & Lindström F. 2010. Being in a pre-school sound environment – annoyance and subjective symptoms among personnel and children. Internoise, 13-16 June, Lisbon, Portugal.
- Prasher D. & Sułkowski W. 1998. The role of otoacoustic emissions in screening and evaluation of noise damage. *Int J Occup Med Environ Health*, 12, 183–192.
- Rosso M., Agius R. & Calleja N. 2011. Development and validation of a screening questionnaire for noise-induced hearing loss. *Occup Med (Lond)*, 61, 416–421.
- Ruggles D., Bharadwaj H. & Shinn-Cunningham B.G. 2011. Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication. *Proc Natl Acad Sci*, 108, 15516–15521.
- Shiomi Y., Tsuji J., Naito Y., Fujiki N. & Yamamoto N. 1997. Characteristics of DPOAE audiogram in tinnitus patients. *Hear Res*, 108, 83–88.
- Sindhusake D., Mitchell P., Smith W., Golding M., Newall P. et al. 2001. Validation of self-reported hearing loss. The blue mountains hearing study. *Int J Epidemiol*, 30, 1371–1378.
- Sjödin F., Kjellberg A., Knutsson A., Landström U. & Lindberg L. 2012. Noise exposure and auditory effects on preschool personnel. *Noise Health*, 14, 72.
- Sztuka A., Pospiech L., Gawron W. & Dudek K. 2010. DPOAE in estimation of the function of the cochlea in tinnitus patients with normal hearing. *Auris Nasus Larynx*, 37, 55–60.
- Videhult Pierre P., Johnson A.-C. & Fridberger A. 2015. Subjective and clinically assessed hearing loss: A cross-sectional register-based study on a Swedish population aged 18 through 50 years. *PLoS One*, 10, e0123290.
- Voeks S., Gallagher C., Langer E. & Drinka P. 1993. Self-reported hearing difficulty and audiometric thresholds in nursing home residents. *J Fam Pract*, 36, 54–58.

Supplementary material available online