

Original Paper

# Self-Test Web-Based Pure-Tone Audiometry: Validity Evaluation and Measurement Error Analysis

Marcin Masalski<sup>1,2</sup>, MD, PhD; Tomasz Kręcicki<sup>1</sup>, MD, PhD

<sup>1</sup>Department and Clinic of Otolaryngology, Head and Neck Surgery, Wrocław Medical University, Wrocław, Poland

<sup>2</sup>Institute of Biomedical Engineering and Instrumentation, Wrocław University of Technology, Wrocław, Poland

**Corresponding Author:**

Marcin Masalski, MD, PhD

Department and Clinic of Otolaryngology, Head and Neck Surgery

Wrocław Medical University

Borowska 213

Wrocław, 50-556

Poland

Phone: 48 71 734 3700

Fax: 48 71 733 12 09

Email: [marcin.masalski@pwr.wroc.pl](mailto:marcin.masalski@pwr.wroc.pl)

## Abstract

**Background:** Potential methods of application of self-administered Web-based pure-tone audiometry conducted at home on a PC with a sound card and ordinary headphones depend on the value of measurement error in such tests.

**Objective:** The aim of this research was to determine the measurement error of the hearing threshold determined in the way described above and to identify and analyze factors influencing its value.

**Methods:** The evaluation of the hearing threshold was made in three series: (1) tests on a clinical audiometer, (2) self-tests done on a specially calibrated computer under the supervision of an audiologist, and (3) self-tests conducted at home. The research was carried out on the group of 51 participants selected from patients of an audiology outpatient clinic. From the group of 51 patients examined in the first two series, the third series was self-administered at home by 37 subjects (73%).

**Results:** The average difference between the value of the hearing threshold determined in series 1 and in series 2 was -1.54dB with standard deviation of 7.88dB and a Pearson correlation coefficient of .90. Between the first and third series, these values were -1.35dB±10.66dB and .84, respectively. In series 3, the standard deviation was most influenced by the error connected with the procedure of hearing threshold identification (6.64dB), calibration error (6.19dB), and additionally at the frequency of 250Hz by frequency nonlinearity error (7.28dB).

**Conclusions:** The obtained results confirm the possibility of applying Web-based pure-tone audiometry in screening tests. In the future, modifications of the method leading to the decrease in measurement error can broaden the scope of Web-based pure-tone audiometry application.

(*J Med Internet Res* 2013;15(4):e71) doi: [10.2196/jmir.2222](https://doi.org/10.2196/jmir.2222)

**KEYWORDS**

pure tone audiometry; computer-assisted instruction; self-examination

## Introduction

The development of Internet technologies combined with technological progress in the construction of personal computers, especially in terms of improving performance of sound cards, offer the possibility of conducting self-administered hearing tests at home. Moreover, in a group of elderly people who often suffer from hearing loss, there is an increasing number of Internet users. The validity of such hearing tests needs to be

evaluated before applying the test results in the diagnostic and therapeutic process.

The remote examinations of hearing conducted with the use of PCs have been carried out for about 10 years [1]. These examinations can be divided into surveys [2-4] and examinations done with the use of sound signals. The sound signals can be generated through a dedicated external device connected to a PC, usually an audiometer [5,6] or a PC sound card [3,4,7-12]. Among hearing examinations, we can distinguish screening

tests [3,7,11,12] and examinations such as pure-tone audiometry [4-6,8-10], which provide additional information.

The hearing screening tests done remotely with the use of a PC sound card are most often in the form of a speech-in-noise test [3,7,11,12]. The speech-in-noise test is preferred over surveys [3] as it improves the detection of hearing impairment among the population [7], especially after introducing a low-pass noise [12]. The evaluation of validity of pure-tone audiometry conducted in similar conditions, ie, with the use of a PC sound card and ordinary headphones [8-10] is ambiguous and depends on the adopted solutions, which include, eg, calibration, hearing threshold evaluation method, and presence or lack of a person supervising the test. In the supervised tests, the mean error concerning the determination of a hearing threshold on a specially constructed and calibrated PC-based device was below 2.3dB [8]. In unsupervised tests conducted after computer calibration performed by a person with good hearing, the greatest error occurred at the frequency of 2kHz and 4kHz and was -5.6dB and -5.1dB respectively, with standard deviation of 8.29dB and 6.9dB [10]. In unsupervised tests conducted without calibration, the maximum mean difference occurred at the frequency of 500Hz and was 11.3dB [9].

The potential application of pure-tone audiometry conducted on a PC depends on the value of the measurement error. The PC-based test will not substitute the clinical pure-tone audiometry. However, it can be applied for conducting self-administered check-ups in cases of limited access to clinical devices, eg, at the general practitioner. Alternatively, it can be used as an initial telemedical examination combined with a survey, which will help determine the direction of future treatment. The aim of the research was to identify the measurement error connected with determining the hearing threshold conducted by means of self-administered Web-based pure tone audiometry, as well as identify and analyze factors affecting its value.

## Methods

The hearing threshold evaluation was done in three series: (1) audiologist-performed tests on clinical audiometer at an audiology outpatient clinic, (2) self-administered Web-based tests on a specially calibrated computer at an audiology outpatient clinic under the supervision of an audiologist, and (3) self-administered Web-based tests conducted at home.

In series 1 and 2, 51 participants (32 men, 19 women), aged 11-60 years (the median age was 34) underwent examination. The research participants were selected from among the patients of an audiology outpatient clinic. The qualification criterion was the willingness to participate in the research, owning a PC, basic skills to operate it, and having an email account. 102 ears were examined from which 45 (44%) were ears without a hearing loss, ie, with the hearing threshold of 25dBHL and less, 17 (16.7%) were ears with hearing impairment below 40dBHL, 31 (30.4%) with hearing impairment above 40dBHL and below 70dBHL, and 9 (8.8%) were above 70dBHL. From the group of 51 patients examined in the first two series, series 3 was self-administered at home by 37 subjects (73%). The examinations in series 1 and 2 were conducted on the same day,

and examinations from series 3 were conducted up to 196 days later (median of 9 days). In the case of failure to conduct test in series 3 in a time of 2-3 weeks, the patients were reminded by a phone call and then after about 1 month, they were reminded by an email.

## Tests in Clinical Settings

All tests from series 1 were made in an acoustic cabin with the use of clinical audiometer Interacoustic AD229e and headphones TDH-39. The calibration of the audiometer was conducted according to ISO 389-1:1998. The hearing threshold in pure-tone audiometry was determined with the use of the ascending method, according to ISO 8253-1:2010. The level of the tone was reduced in 10dB steps until no further response occurred, and then it was increased in 5dB steps until the subject responded. The threshold was defined as the lowest level at which responses occurred in at least half of the series of ascending trials with a minimum of two responses required at that level [13]. The examinations were conducted at the frequencies of 250Hz, 500Hz, 1kHz, 2kHz, 4kHz, and 8kHz [13].

## Self-administered Web-Based Test

The examinations in series 2 and 3 were conducted on personal computers following their calibration. Both the calibration and the examination were performed at system volume set at the maximum level. The calibration consisted of determining the reference sound level by a person with good hearing. During calibration, two sounds differing in intensity were presented bilaterally in turns for 1 second at a frequency of 1kHz. The difference in sound intensity was stable and equalled 5dB. The task of the reference person was to set the volume in a way that ensured that only the louder of the two sounds presented was audible. The volume was controlled by means of a slider with the step of 1dB. The reference person could listen to the sound for unlimited time, adjusting the volume using the slider with 1dB step any number of times, in order to finally confirm the selected level with a button. The mean intensity of the two presented sounds determined the hearing level of the reference person and was marked 0dBRL-HL. The reference values at frequencies differing from the calibration frequency, ie, 1kHz, were calculated using the model based on the A-weighting filter [14].

The procedure of determining the hearing threshold during the test, both in series 2 and 3, differed from that applied during calibration. The task of the subject was to set the intensity of the presented pure tone modulated by sinusoidal envelope with a 2-second period at such a level that the sound was on the verge of audibility. Patients could listen to the sound for unlimited time and change the intensity of the presented sound signal by themselves using a slider with 5dB step any number of times, and then confirm the chosen level with a button. The hearing threshold was determined at frequencies as in series 1, ie, 250Hz, 500Hz, 1kHz, 2kHz, 4kHz, and 8kHz. For hearing thresholds below 0dBHL-RP, the value was assumed to be 0dBHL-RP. The time spent on calibrating and examinations was not recorded.

All the examinations in series 2 were conducted on a notebook Dell Vostro 1320 with operational system Microsoft Windows

7 and headphones Technics RP-F290 placed in an acoustic cabin of an audiology outpatient clinic and connected to the Internet. The calibration was repeated six times by 3 people whose hearing threshold in pure-tone audiometry at the calibration frequency was 0dBHL. Each person performed calibration twice. The final value of the calibration coefficient was determined as the mean of all values obtained from single calibrations and was the same for all examinations in series 2. Examinations were supervised by an audiologist, whose task was to train a patient and detect his/her mistakes, eg, changing the sides of headphones, omission of the frequency, or accidental marking of a threshold that was significantly different from the actual one.

Examinations in series 3 were conducted by the patients themselves on their own personal computers at home. Each examination was conducted on a different computer using different headphones. The test station was calibrated by other household members with no previous hearing problems. In that series, the hearing threshold of the reference person was not controlled. The trial participants were instructed to conduct the test at home in a quiet place, preferably in the evening or at night. Moreover, they were informed that the calibration should be performed by a person up to 35 years old, with no previous hearing problems.

## Results

### Comparison of the Results Between Series

The hearing threshold determined in series 1 was compared to hearing thresholds from series 2 and 3. If there was no response at a given frequency, measurement was not taken into account for further calculations. Figure 1 presents relationships between the series: the relationship between hearing thresholds with division into frequencies, the relationship between total hearing thresholds, and that between the mean hearing thresholds calculated on the basis of the value obtained at all the examined frequencies.

The mean difference between thresholds in series, its standard deviation, Pearson's correlation coefficient, and linear estimators (Table 1) were calculated for the relationships described above. Linear estimators were determined for Deming's regression because the explanatory variable, which constitutes the hearing threshold in series 1, is also burdened with measurement error. The mean difference in the hearing threshold between series 1 and 2 and between series 1 and 3 was  $-1.54\text{dB}$  and  $-1.34\text{dB}$  respectively, with standard deviation  $7.88\text{dB}$  and  $10.66\text{dB}$

respectively, and Pearson's correlation coefficients  $.90$  and  $.84$  respectively. In both comparisons, the lowest values of Pearson's correlation coefficient were obtained for the frequency of  $250\text{Hz}$  at the level of  $.88$  and  $.69$  respectively. The highest value of the standard deviation occurred at the frequency of  $8\text{kHz}$  ( $8.88\text{dB}$ ) and  $500\text{Hz}$  ( $12.03\text{dB}$ ) respectively. Pearson's correlation coefficients calculated for the mean threshold were at the level of  $.94$  and  $.89$  respectively (Table 1).

The mean difference between thresholds in series, its standard deviation, Pearson's correlation coefficient, and linear estimators of Deming's regression were determined for the hearing loss below  $40\text{dBHL}$  as well as greater than or equal to  $40\text{dBHL}$  (Table 2). The division was made on the basis of the mean value of the hearing threshold in two comparable series. The mean difference in the hearing threshold did not exceed  $2\text{dB}$  in any of the groups, and its standard deviation increases together with the increase in the hearing loss. Pearson's correlation coefficients reflect changes in the standard deviation and reach low values due to narrow ranges of random variables compared to their standard deviations. For the same reason, the values of linear estimators deviate from the values established for the whole group.

On the basis of the obtained results, the sensitivity and specificity of noise-induced hearing loss detection was calculated, according to the criteria proposed in [15] adopted for the purposes of this paper. The noise-induced hearing loss was detected when the hearing threshold exceeded  $30\text{dB}$  at one of the following frequencies:  $500\text{Hz}$ ,  $1\text{kHz}$ ,  $2\text{kHz}$ , or  $25\text{dB}$  at more than one, or when the hearing threshold exceeded  $50\text{dB}$  at  $4\text{Hz}$ . For series 2, the sensitivity was  $0.92$  with confidence interval of  $(0.81, 1.0)$  at  $P=.05$ , and the specificity was  $0.96$  with confidence intervals of  $(0.88, 1.0)$  at  $P=.05$ . For series 3, the respective values were sensitivity  $0.89$  ( $0.74, 1.0$ ) and specificity  $0.89$  ( $0.76, 1.0$ ).

### Analysis of the Measurement Error

This paper attempts to identify and assess the values of the sources of error in determining the hearing threshold of Web-based examinations. The literature data on standard deviation of the hearing threshold determined in test-retest examinations carried out in clinical settings are presented in Table 3. On the basis of these data, the value of the standard deviation of the hearing threshold determined in test-retest examinations was adopted for the further calculations at the level of  $6\text{dB}$ .

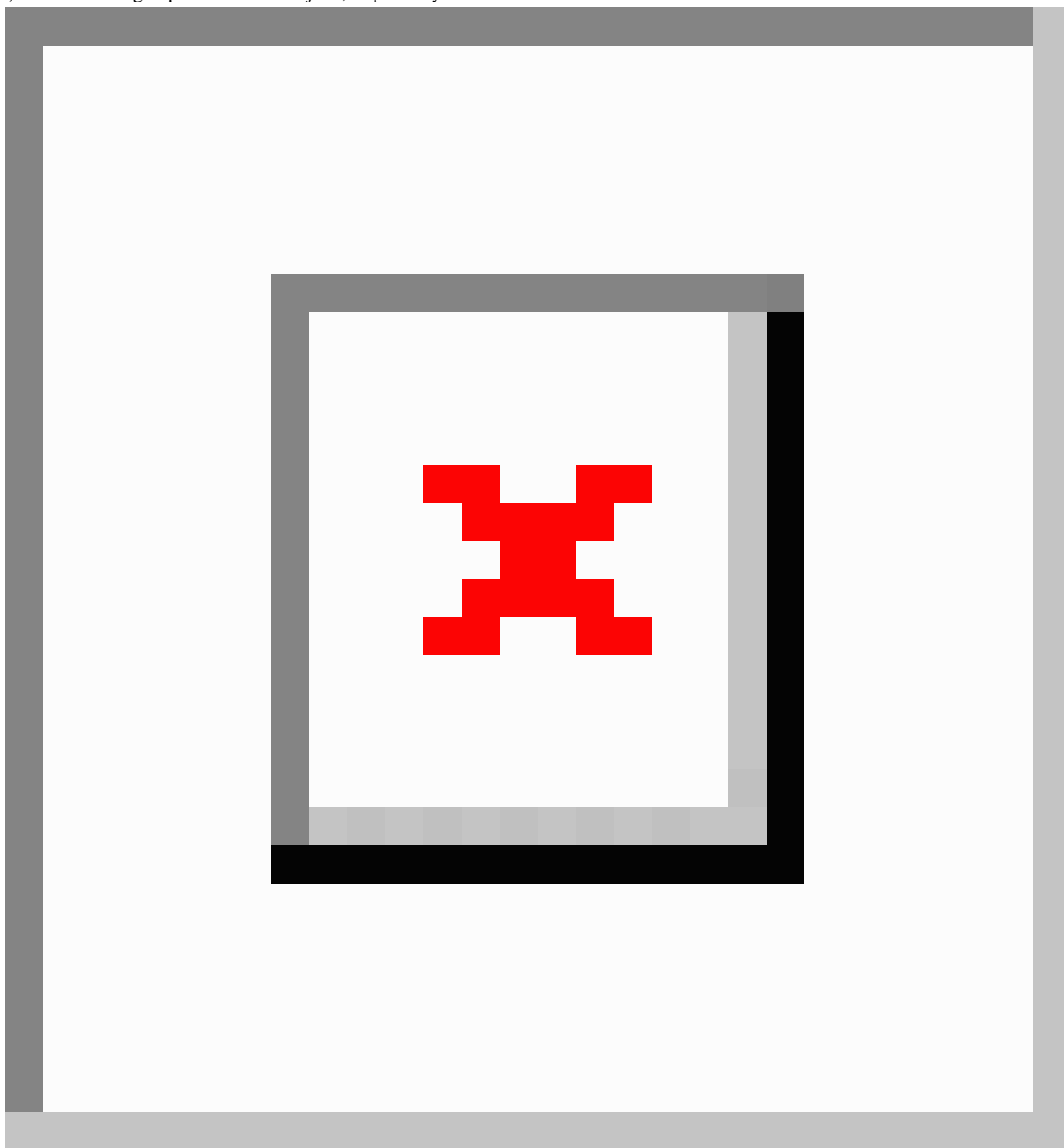
**Table 1.** The mean difference  $m$  in the hearing threshold value between the series calculated on the basis of  $n$  data points collected on a group of  $N$  subjects, its standard deviation  $\sigma$ , Pearson's correlation coefficient  $r$  and linear estimators  $a$ ,  $b$  of Deming's regression  $y=ax+b$  and corresponding confidence intervals CI at  $P=.05$ .

	f	n	$m$	$\sigma$	$r$ (CI)	$a$ (CI)	$b$ (CI)
<b>Series 1 vs 2, N=51</b>							
	250Hz	100	-3.70	7.30	.88 (0.82, 0.92)	1.08 (1.06, 1.09)	-5.49 (-5.84, -5.13)
	500Hz	101	0.52	7.20	.88 (0.83, 0.92)	0.93 (0.92, 0.94)	2.09 (1.83, 2.34)
	1kHz	101	-4.59	6.91	.92 (0.88, 0.94)	0.85 (0.84, 0.86)	-1.16 (-1.39, -0.92)
	2kHz	101	-2.42	7.22	.93 (0.90, 0.95)	0.95 (0.94, 0.96)	-1.18 (-1.44, -0.91)
	4kHz	100	2.15	7.68	.93 (0.90, 0.95)	0.97 (0.96, 0.98)	2.88 (2.64, 3.12)
	8kHz	100	-1.20	8.88	.91 (0.86, 0.94)	0.93 (0.92, 0.93)	0.71 (0.47, 0.95)
	Total	603	-1.54	7.88	.90 (0.89, 0.92)	0.96 (0.96, 0.96)	-0.52 (-0.57, -0.48)
	Mean	99	-1.52	5.42	.94 (0.91, 0.96)	0.94 (0.93, 0.96)	-0.23 (-0.46, -0.01)
<b>Series 1 vs 3, N=37</b>							
	250Hz	71	-5.07	12.03	.69 (0.55, 0.80)	1.23 (1.19, 1.26)	-10.42 (-11.14, -9.70)
	500Hz	71	-0.35	10.43	.79 (0.68, 0.86)	1.05 (1.03, 1.08)	-1.57 (-2.08, -1.05)
	1kHz	72	-1.24	9.28	.85 (0.76, 0.90)	1.01 (0.99, 1.03)	-1.49 (-1.96, -1.02)
	2kHz	72	-1.58	10.46	.88 (0.81, 0.92)	0.90 (0.87, 0.92)	1.04 (0.48, 1.59)
	4kHz	72	-1.42	10.77	.87 (0.80, 0.92)	0.98 (0.96, 1.00)	-0.83 (-1.33, -0.33)
	8kHz	71	1.63	10.05	.88 (0.82, 0.93)	1.05 (1.04, 1.07)	0.31 (-0.11, 0.73)
	Total	429	-1.34	10.66	.84 (0.81, 0.86)	1.02 (1.01, 1.02)	-1.71 (-1.80, -1.62)
	Mean	68	-1.44	7.59	.89 (0.83, 0.93)	1.01 (1.00, 1.03)	-1.73 (-2.14, -1.33)

**Table 2.** The mean difference  $m$  in the hearing threshold  $t$  between the series calculated on the basis of  $n$  data points collected on a group of  $N$  subjects, its standard deviation  $\sigma$ , Pearson's correlation coefficient  $r$  and linear estimators  $a$ ,  $b$  of Deming's regression  $y=ax+b$  and corresponding confidence intervals CI at  $P=.05$  for the hearing threshold below 40dBHL as well as greater than or equal to 40dBHL.

Group	n	$m$	$\sigma$	$r$ (CI)	$a$ (CI)	$b$ (CI)
<b>Series 1 vs 2, N=51</b>						
$t < 40\text{dBHL}$	498	-1.45	7.37	.74 (0.70, 0.78)	0.84 (0.84, 0.84)	1.24 (1.19, 1.30)
$t \geq 40\text{dBHL}$	105	-1.99	9.88	.64 (0.51, 0.74)	1.07 (1.04, 1.09)	-5.67 (-7.04, -4.30)
<b>Series 1 vs 3, N=37</b>						
$t < 40\text{dBHL}$	339	-1.52	9.72	.58 (0.51, 0.65)	1.11 (1.10, 1.12)	-3.38 (-3.53, -3.22)
$t \geq 40\text{dBHL}$	90	-0.64	13.67	.44 (0.25, 0.59)	0.66 (0.62, 0.69)	17.59 (15.56, 19.62)

**Figure 1.** The relationship between hearing thresholds in series 1 and 2 (first two columns on left) and between series 1 and 3 (last two columns on right) calculated in a group of 51 and 37 subjects, respectively.



**Table 3.** Standard deviations of the hearing threshold obtained in test-retest examination (literature data) (separate measurements were carried out for supra-aural headphones (series I, II, III) and in-ear headphones (IV, V, VI) and for the group of young (series I and IV), older (II, V) and the oldest (III, VI) subjects; N=number of subjects) [16-18].

Investigator	Series	N	250	500	1kHz	2kHz	4kHz	8kHz	Mean
Brown 1948 [16]		30	5.3	4.4	5.8	4.5	5.4	7.1	5.4
Erlandsson et al 1979 [17] <sup>a</sup>		10		6.3	6.2	4.7	5.3	10.4	6.6
Landry et al 1999 [18]	I	20	5.0	4.6	5.2	4.1	4.3	6.0	5.9
	II	10	8.5	7.9	5.2	6.4	9.5	6.9	
	III	10	5.9	5.4	6.4	2.1	6.7	7.1	
	IV	20	8.8	7.7	4.6	5.5	2.9	3.6	
	V	10	7.1	4.4	7.4	5.5	2.6	5.4	
	VI	10	9.3	4.6	3.3	5.3	9.0	9.4	
Mean			7.1	5.7	5.5	4.8	5.7	7.0	6.0

<sup>a</sup>Estimated by comparison with Békésy audiometry performed 5 times on a group of 10 subjects.

Knowing that the variance of the sum of the two independent random variables  $X$  and  $Y$  is the sum of their variance (see [Multimedia Appendix 1](#)), the standard deviation for a single measurement of the hearing threshold in clinical settings  $\sigma_{clin}$  was calculated at the level of 4.25dB by dividing the  $\sigma_{2clin}$  by the square root of two (see [Multimedia Appendix 2](#)).

Therefore, assuming the variability of the hearing threshold measurement in series 1  $\sigma_{(i)}$  at the level of literature data  $\sigma_{clin}$ , the standard deviation of the measurement error in Web-based hearing tests for series 2  $\sigma_{(ii)}$  was calculated at 6.64dB based on the value  $\sigma_{(i)}$  and the value  $\sigma_{(i)(ii)}$ , which is the standard deviation of the hearing threshold difference between the values determined in series 1 and 2. Similarly, for series 3, we obtain the standard deviation  $\sigma_{(iii)}$  equal to 9.78dB (see [Multimedia Appendix 3](#)).

In series 2, the standard deviation of the measurement error depends mainly on the standard deviation of the error connected with the procedure of determining threshold value  $\sigma_{proc}$ . In series 3, apart from the error connected with the procedure of determining the threshold value, we can distinguish other sources of error influencing the standard deviation: the standard deviation of the calibration error  $\sigma_{cal}$ , which is the reference sound level evaluation error at the calibration frequency, the frequency nonlinearity error  $\sigma_{nonlin}$ , which is the difference between the actual reference sound level and that set by the model, the gain error  $\sigma_{gain}$ , and the error connected with background noise  $\sigma_{noise}$  (see [Multimedia Appendix 4](#)).

Of course in series 2, the calibration error, frequency nonlinearity error, and gain error were the same in all tests and did not influence the value of standard deviation, while the error connected with background noise can be omitted since the tests were conducted in an acoustic cabin.

[Figure 2](#) shows the standard deviation of threshold differences between series in relation to the hearing threshold for series 1 and the frequency. The lowest values of standard deviations are observed in small hearing losses, ie, for those measurements that should be most affected by the background noise. Therefore, we can assume that the error connected with background noise is significantly smaller than other components of the measurement error. In the following considerations, the standard deviation of error connected with background noise was omitted ( $\sigma_{noise} \approx 0\text{dB}$ ), both for series 2 and for series 3.

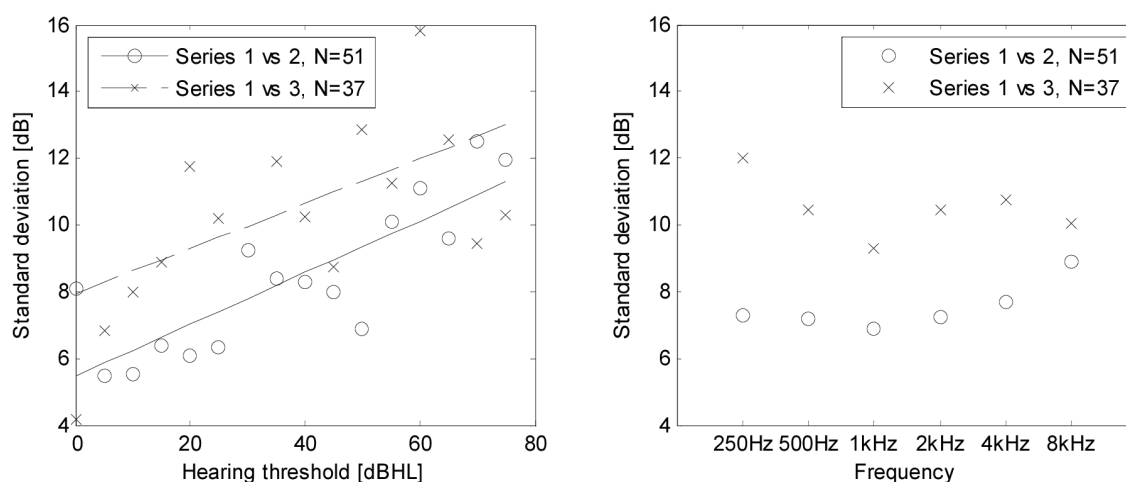
With the increase in intensity, increment of the standard deviation of measurement error in series 3 is slightly smaller than in series 2. In series 2, the standard deviation of the gain error equals zero since all tests are done on the same set. The lack of increase in standard deviation in series 3 above the values observed in series 2 indicates a negligible effect of the gain error on the measurement values ( $\sigma_{gain} \approx 0\text{dB}$ ).

The standard deviation of the calibration error  $\sigma_{cal}$  can be estimated by comparing the standard deviations for series 2 and 3 at the calibration frequency, ie, 1kHz. Since the calibration is performed on the same set as the test, the error connected with the nonlinearity of the frequency response at the calibration frequency is zero ( $_{1\text{kHz}}\sigma_{nonlin} = 0\text{dB}$ ). Taking into account the above, we obtain the calibration error  $\sigma_{cal}$  equal to 6.19dB (see [Multimedia Appendix 5](#)). Finally, knowing the calibration error  $\sigma_{cal}$  and assuming the insignificance of the error connected with background noise as well as the gain error, the standard deviations of the frequency nonlinearity error were calculated. The highest value  $_{250\text{Hz}}\sigma_{nonlin}$  was observed at the frequency of 250Hz at the level of 7.28dB (see [Multimedia Appendix 5](#)).

No asymmetry of measurement error between the left and right sides was detected in any of the tests, which could raise suspicion concerning incorrect audio balance.



**Figure 2.** Standard deviation of the hearing threshold differences between the series calculated for the group of N subjects in relation to the hearing threshold evaluated in series 1 (left) and the frequency (right).



## Discussion

The aim of the research was to identify the measurement error connected with determining the hearing threshold conducted by means of self-administered Web-based pure tone audiometry, as well as identify and analyze factors affecting its value.

The average difference between the value of the hearing threshold determined in series 1 and in series 2 was  $-1.54\text{dB}$  with standard deviation of  $7.88\text{dB}$  and Pearson's correlation coefficient of  $.90$ . Between series 1 and series 3, these values were  $-1.35\text{dB} \pm 10.66\text{dB}$  and  $.84$  respectively. For the average hearing threshold, these values were, for series 2  $-1.52\text{dB} \pm 5.42\text{dB}$  and  $.94$  respectively, and for series 3,  $-1.44\text{dB} \pm 7.59\text{dB}$  and  $.89$ . The results from series 2 are consistent with the results presented by Honeth et al, who determined the correlation coefficient of the average hearing threshold for the group of 72 people at the level of  $.94$  and  $.93$  respectively for the right and left ears [10]. 64 out of 72 people (89%) were examined on the same test set calibrated by the same person.

In series 3, the standard deviation values were most influenced by the error connected with the procedure of determining the hearing threshold, the same as the error in series 2 ( $\sigma_{proc}=6.64\text{dB}$ ), the calibration error ( $\sigma_{cal}=6.19\text{dB}$ ), and additionally at the frequency of 250Hz, the frequency response nonlinearity error ( $_{250\text{Hz}}\sigma_{nonlin}=7.28\text{dB}$ ).

The error of determining the hearing threshold during examination, which consisted of self-adjustment of the position of the volume slider in 5dB steps, was estimated at the level of  $\sigma_{proc}=6.64\text{dB}$ . The reason for such high error value compared to the error value of the ascending method ( $\sigma_{clin}=4.25\text{dB}$ ) may be attributed to the difficulty connected with self-assessment of the hearing threshold. Standard deviations between series 1 and 2 as well as between series 1 and 3 increase together with the increase in the hearing loss (Figure 2). This suggests that the task of self-evaluation of the hearing threshold proves more difficult to perform for subjects with moderate and profound

hearing impairment compared to normal hearing subjects. Replacement of this method with the ascending method will contribute to improving the accuracy of the examination. The method of self-evaluation of the hearing threshold was originally chosen as simpler, faster and more flexible. Consequently, it was less monotonous and more attractive for the patient.

The purpose in the calibration was to approximate the 0dBHL level at the examined frequencies. The approximation can be conducted in a number of ways. The reference sound level can be determined independently for each frequency in relation to the hearing threshold of the reference person [10]. In this case, the calibration error at each frequency is burdened with measurement error of the hearing threshold and the difference between the hearing threshold of the reference person and 0dBHL. Another solution, which was adopted in this paper, is to measure the hearing threshold at one frequency, at which differences between the values of the hearing threshold among the population are the smallest and then determine the reference sound level at other frequencies on the basis of the model. Alternatively, calibration can be performed at a number of frequencies and the correctness of the calibration may be assessed based on its accordance with the model. At the same time, in the case of a large discrepancy between the measurement results, the calibration can be repeated by another person or on different headphones. The choice of the optimal method requires further research.

The calibration also involves the problem of selecting a procedure for determining the hearing threshold of the reference person. Calibration is usually conducted by young persons with normal hearing. Therefore, in order to achieve more accurate calibration results, one can use the procedure requiring greater hearing efficiency compared to the procedure used in examination. This study applies calibration consisting in changing the intensity of two sounds with the stable difference of 5dB. This method requires considerable concentration and optimal hearing ability. However, in contrast to the ascending method, it is not burdened with discretization errors.

Calibration error ( $\sigma_{cal}=6.19\text{dB}$ ) is connected with the error of hearing threshold assessment  $\sigma_{proc\_cal}$  and the standard deviation in the hearing threshold of the reference persons  $\sigma_{pop}$  at the calibration frequency, ie, 1kHz. The standard deviation of the hearing threshold determined by means of the ascending method among the population of young subjects without prior hearing problems,  $\sigma_{pop\_asc}$ , can be estimated on the basis of literature data [19]. The distribution of the hearing threshold at 1kHz determined for the population of 2490 subjects was approximated with normal distribution. The standard deviation was calculated based on the values of 1st and 3rd quartile at the level of  $\sigma_{pop\_asc}=5.6\text{dB}$ . Assuming the standard deviation of the ascending method  $\sigma_{clin}$  is at the level of 4.25dB, we obtain  $\sigma_{pop}$  at 3.65dB and then  $\sigma_{proc\_cal}$  at 5.00dB (see [Multimedia Appendix 6](#)).

The method of the hearing threshold evaluation used during calibration ( $\sigma_{proc\_cal}=5.00\text{dB}$ ), despite the possibility of setting the threshold with the accuracy of 1dB step, is characterized by a slightly higher error than in the ascending method  $\sigma_{clin}=4.25\text{dB}$ , whose step equals 5dB. Therefore, accuracy improvement of the Web-based examination can be achieved by modifying the method. An interesting solution is offered by, eg, the Békésy's method, as it is characterized by lower standard deviation in the test-retest examination compared to the ascending method [17], and additionally it is not burdened with discretization errors.

The frequency nonlinearity error is the difference between the actual values of the reference sound level and the values determined by the model. The greatest value of  $250\text{Hz}\sigma_{nonlin}=7.28\text{dB}$  was observed at 250Hz with the mean value calculated for all frequencies, excluding the calibration frequency  $mean\sigma_{nonlin}=4.05\text{dB}$ . The frequency nonlinearity error is connected with different values of Reference Equivalent Threshold Sound Pressure Level (RETSPL) of the headphones used in the examination and the frequency nonlinearity of the sound card. The biggest differences between the RETSPL values occur at low frequencies [20], also in the low frequency range, more often than at mid and high frequencies, frequency distortion of the sound card can be observed, eg, when bass

boost option or equalizer settings are enabled. The improvement in accuracy, especially at low frequencies, can be acquired through selection of headphones for which RETSPL values are known. However, in practice this will be hard to achieve for tests carried out at home. An alternative solution would be control determination of the calibration coefficient at low frequencies and verification against the value determined by the model.

Taking into account the significant value of the calibration error and frequency nonlinearity error, it seems interesting to conduct pure-tone audiometry examination on generally available appliances with known electro-acoustic parameters, eg, smartphones. On determining frequency characteristics of the selected smartphone model with bundled headphones, it seems possible to omit the calibration stage and thus obtain more precise examination results.

Web-based pure-tone audiometry can be used without previous training in its conducting. However, it requires the knowledge of basic terms such as the hearing threshold or frequency. On the other hand, an attractive, intuitive and user-friendly interface can largely replace training. Prior to the tests in series 3, patients were instructed on how to perform the test and had performed similar tests before in series 2 under the supervision of an audiologist. The knowledge of the application could lead to reduction in the value of the measurement error.

The obtained results of sensitivity and specificity confirm the possibility of applying Web-based pure-tone audiometry in screening tests. Moreover, Web-based pure-tone audiometry may be used for self-monitoring of hearing, especially if tests are to be conducted under the same conditions. If the same equipment is used, the relative error between subsequent examinations will be reduced by frequency nonlinearity error, and in the case of the same calibration coefficients, relative error will be reduced by the calibration error. Self-monitoring of hearing may become applicable in hearing disorders, such as fluctuating hearing loss, tinnitus, sudden hearing loss, otosclerosis, Ménière's disease, as well as during treatment with ototoxic drugs. In the future, modifications of the method leading to the decrease in measurement error can broaden the scope of Web-based pure-tone audiometry application.

---

## Acknowledgments

The trial was financed by Wroclaw Medical University as part of Research Grant for Young Scientists. Grant No. 38/Pbmn.

---

## Conflicts of Interest

The author of this article is the owner of the Internet portal, e-audiologia.pl, on the basis of which the tests above were conducted.

---

## Multimedia Appendix 1

Variance of the sum of the two independent random variables.

[\[PNG File, 4KB-Multimedia Appendix 1\]](#)

---

## Multimedia Appendix 2

The standard deviation for a single measurement of the hearing threshold determined by means of the ascending method.

[\[PNG File, 69KB-Multimedia Appendix 2\]](#)



### Multimedia Appendix 3

The standard deviation for a single measurement of the hearing threshold determined in series 2 and series 3.

[\[PNG File, 101KB-Multimedia Appendix 3\]](#)

---

### Multimedia Appendix 4

Factors influencing the standard deviation of the measurement error of the hearing threshold in the series 2 and 3.

[\[PNG File, 145KB-Multimedia Appendix 4\]](#)

---

### Multimedia Appendix 5

The standard deviation of the calibration error and frequency nonlinearity error.

[\[PNG File, 217KB-Multimedia Appendix 5\]](#)

---

### Multimedia Appendix 6

The standard deviation of the hearing evaluation method used during calibration.

[\[PNG File, 136KB-Multimedia Appendix 6\]](#)

---

### Multimedia Appendix 7

CONSORT-EHEALTH checklist V1.6.2 [21].

[\[PDF File \(Adobe PDF File\), 994KB-Multimedia Appendix 7\]](#)

---

## References

1. Swanepoel de W, Hall JW. A systematic review of telehealth applications in audiology. *Telemed J E Health* 2010 Mar;16(2):181-200. [doi: [10.1089/tmj.2009.0111](https://doi.org/10.1089/tmj.2009.0111)] [Medline: [20187743](https://pubmed.ncbi.nlm.nih.gov/20187743/)]
2. Chung JH, Des Roches CM, Meunier J, Eavey RD. Evaluation of noise-induced hearing loss in young people using a Web-based survey technique. *Pediatrics* 2005 Apr;115(4):861-867 [FREE Full text] [doi: [10.1542/peds.2004-0173](https://doi.org/10.1542/peds.2004-0173)] [Medline: [15805356](https://pubmed.ncbi.nlm.nih.gov/15805356/)]
3. Smits C, Kramer SE, Houtgast T. Speech reception thresholds in noise and self-reported hearing disability in a general adult population. *Ear Hear* 2006 Oct;27(5):538-549. [doi: [10.1097/01.aud.0000233917.72551.cf](https://doi.org/10.1097/01.aud.0000233917.72551.cf)] [Medline: [16957503](https://pubmed.ncbi.nlm.nih.gov/16957503/)]
4. Bexelius C, Honeth L, Ekman A, Eriksson M, Sandin S, Bagger-Sjöbäck D, et al. Evaluation of an internet-based hearing test--comparison with established methods for detection of hearing loss. *J Med Internet Res* 2008;10(4):e32 [FREE Full text] [doi: [10.2196/jmir.1065](https://doi.org/10.2196/jmir.1065)] [Medline: [18940783](https://pubmed.ncbi.nlm.nih.gov/18940783/)]
5. Givens GD, Blanarovich A, Murphy T, Simmons S, Blach D, Elangovan S. Internet-based tele-audiometry system for the assessment of hearing: a pilot study. *Telemed J E Health* 2003;9(4):375-378. [doi: [10.1089/153056203772744707](https://doi.org/10.1089/153056203772744707)] [Medline: [14980095](https://pubmed.ncbi.nlm.nih.gov/14980095/)]
6. Swanepoel de W, Mngemane S, Molemong S, Mkwanzazi H, Tutshini S. Hearing assessment-reliability, accuracy, and efficiency of automated audiometry. *Telemed J E Health* 2010 Jun;16(5):557-563. [doi: [10.1089/tmj.2009.0143](https://doi.org/10.1089/tmj.2009.0143)] [Medline: [20575723](https://pubmed.ncbi.nlm.nih.gov/20575723/)]
7. Smits C, Merkus P, Houtgast T. How we do it: The Dutch functional hearing-screening tests by telephone and internet. *Clin Otolaryngol* 2006 Oct;31(5):436-440. [doi: [10.1111/j.1749-4486.2006.01195.x](https://doi.org/10.1111/j.1749-4486.2006.01195.x)] [Medline: [17014457](https://pubmed.ncbi.nlm.nih.gov/17014457/)]
8. Choi JM, Lee HB, Park CS, Oh SH, Park KS. PC-based tele-audiometry. *Telemed J E Health* 2007 Oct;13(5):501-508. [doi: [10.1089/tmj.2007.0085](https://doi.org/10.1089/tmj.2007.0085)] [Medline: [17999612](https://pubmed.ncbi.nlm.nih.gov/17999612/)]
9. Kimball SH. Inquiry into online hearing test raises doubts about its validity. *Hearing Journal* 2008 Mar;61(3):38-46. [doi: [10.1097/01.HJ.0000314717.42321.4a](https://doi.org/10.1097/01.HJ.0000314717.42321.4a)]
10. Honeth L, Bexelius C, Eriksson M, Sandin S, Litton JE, Rosenhall U, et al. An internet-based hearing test for simple audiometry in nonclinical settings: preliminary validation and proof of principle. *Otol Neurotol* 2010 Jul;31(5):708-714. [doi: [10.1097/MAO.0b013e3181de467a](https://doi.org/10.1097/MAO.0b013e3181de467a)] [Medline: [20458255](https://pubmed.ncbi.nlm.nih.gov/20458255/)]
11. Leensen MC, de Laat JA, Dreschler WA. Speech-in-noise screening tests by internet, part 1: test evaluation for noise-induced hearing loss identification. *Int J Audiol* 2011 Nov;50(11):823-834. [doi: [10.3109/14992027.2011.595016](https://doi.org/10.3109/14992027.2011.595016)] [Medline: [21988504](https://pubmed.ncbi.nlm.nih.gov/21988504/)]
12. Leensen MC, de Laat JA, Snik AF, Dreschler WA. Speech-in-noise screening tests by internet, part 2: improving test sensitivity for noise-induced hearing loss. *Int J Audiol* 2011 Nov;50(11):835-848. [doi: [10.3109/14992027.2011.595017](https://doi.org/10.3109/14992027.2011.595017)] [Medline: [21970351](https://pubmed.ncbi.nlm.nih.gov/21970351/)]

13. British Society of Audiology. Pure-tone air-conduction and bone-conduction threshold audiometry with and without masking. 2011. Recommended Procedure URL: [http://www.thebsa.org.uk/docs/Guidelines/BSA\\_RP\\_PTA\\_FINAL\\_24Sept11.pdf](http://www.thebsa.org.uk/docs/Guidelines/BSA_RP_PTA_FINAL_24Sept11.pdf) [accessed 2012-05-27] [WebCite Cache ID 67yJEIAP1]
14. International Electrotechnical Commission. Electroacoustics - Sound level meters - Part 2: Pattern evaluation tests. In: International Standard IEC 61672-2:2003(E). Geneva: IEC; 2003.
15. Klockhoff I, Drettner B, Hagelin KW, Lindholm L. A method for computerized classification of pure tone screening audiometry results in noise-exposed groups. *Acta Otolaryngol* 1973 Apr;75(4):339-340. [Medline: [4702637](#)]
16. Brown RE. Experimental studies on the reliability of audiometry. *J Laryngol Otol* 1948 Aug;62(8):487-524. [Medline: [18876519](#)]
17. Erlandsson B, Håkanson H, Ivarsson A, Nilsson P. Comparison of the hearing threshold measured by manual pure-tone and by self-recording (Békésy) audiometry. *Audiology* 1979;18(5):414-429. [Medline: [496724](#)]
18. Landry JA, Green WB. Pure-Tone Audiometric Threshold Test-Retest Variability in Young and Elderly Adults. *J Speech Lang Pathol Audiol* 1999;23(2):74-80 [FREE Full text]
19. Engdahl B, Tambs K, Borchgrevink HM, Hoffman HJ. Screened and unscreened hearing threshold levels for the adult population: results from the Nord-Trøndelag Hearing Loss Study. *Int J Audiol* 2005 Apr;44(4):213-230. [Medline: [16011050](#)]
20. Yost YA, Killion MC. Hearing thresholds. In: Crocker MJ, editor. *Encyclopedia of acoustics*. New York: John Wiley; 1997:1545-1553.
21. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. *J Med Internet Res* 2011;13(4):e126 [FREE Full text] [doi: [10.2196/jmir.1923](#)] [Medline: [22209829](#)]

## Abbreviations

**dBHL:** decibel hearing level

**RETSPL:** Reference Equivalent Threshold Sound Pressure Level

*Edited by G Eysenbach; submitted 29.06.12; peer-reviewed by M Eriksson; comments to author 26.09.12; revised version received 04.11.12; accepted 09.02.13; published 12.04.13*

*Please cite as:*

*Masalski M, Kręcicki T*

*Self-Test Web-Based Pure-Tone Audiometry: Validity Evaluation and Measurement Error Analysis*

*J Med Internet Res* 2013;15(4):e71

URL: <http://www.jmir.org/2013/4/e71/>

doi: [10.2196/jmir.2222](#)

PMID: [23583917](#)

©Marcin Masalski, Tomasz Kręcicki. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 12.04.2013. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.