

Estimating Hearing Thresholds From Stimulus-Frequency Otoacoustic Emissions

Trends in Hearing
Volume 24: 1–15
© The Author(s) 2020
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2331216520960053
journals.sagepub.com/home/tia



Qin Gong^{1,2} , Yin Liu¹, and Zewen Peng¹

Abstract

It is of clinical interest to estimate pure-tone thresholds from potentially available objective measures, such as stimulus-frequency otoacoustic emissions (SFOAEs). SFOAEs can determine hearing status (normal hearing vs. hearing loss), but few studies have explored their further potential in predicting audiometric thresholds. The current study investigates the ability of SFOAEs to predict hearing thresholds at octave frequencies from 0.5 to 8 kHz. SFOAE input/output functions and pure-tone thresholds were measured from 230 ears with normal hearing and 737 ears with sensorineural hearing loss. Two methods were used to predict hearing thresholds. Method 1 is a linear regression model; Method 2 proposed in this study is a back propagation (BP) network predictor built on the bases of a BP neural network and principal component analysis. In addition, a BP network classifier was built to identify hearing status. Both Methods 1 and 2 were able to predict hearing thresholds from 0.5 to 8 kHz, but Method 2 achieved better performance than Method 1. The BP network classifiers achieved excellent performance in determining the presence or absence of hearing loss at all test frequencies. The results show that SFOAEs are not only able to identify hearing status with great accuracy at all test frequencies but, more importantly, can predict hearing thresholds at octave frequencies from 0.5 to 8 kHz, with best performance at 0.5 to 4 kHz. The BP network predictor is a potential tool for quantitatively predicting hearing thresholds, at least at 0.5 to 4 kHz.

Keywords

objective estimate of hearing threshold, hearing loss, back propagation neural network, principal component analysis

Received 1 January 2020; Revised 13 March 2020; accepted 8 April 2020

Audiometric thresholds are the gold standard for quantitatively evaluating hearing status. However, because pure-tone audiometry requires responses from subjects, its reliability depends on subject attention and cooperation, which may be difficult to obtain in certain populations. Hence, objective estimates of pure-tone threshold are clinically desirable.

Hearing thresholds can be determined objectively using electrophysiological measurements, such as the auditory brainstem response (Gorga et al., 2006; Johnson & Brown, 2005), the auditory steady-state response (Yeung & Wong, 2007), and the cortical auditory-evoked potentials (Lightfoot & Kennedy, 2006). Electrophysiological methods, however, are time-consuming (e.g., approximately 10.5 min were needed for a single frequency; Van Dun et al., 2015). Therefore, it is worthwhile to explore alternative objective methods, such as otoacoustic emissions (OAEs).

Many studies indicate that distortion-product otoacoustic emissions (DPOAEs) can distinguish between normal and impaired ears from 2 to 4 kHz (Gorga et al., 1993a, 1993b, 1997, 2000; Kim et al., 1996; Musiek & Baran, 1997; Norton et al., 2000; Stover et al., 1996). DPOAE thresholds derived from DPOAE input/output (I/O) functions (Boege & Janssen, 2002; Gorga et al., 2003; Johnson et al., 2007, 2010; Oswald & Janssen, 2003) are significantly correlated with audiometric thresholds (e.g., $r = .65$, Boege & Janssen, 2002;

¹Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing, China

²School of Medicine, Shanghai University, Shanghai, China

Corresponding Author:

Qin Gong, Office C212, School of Medicine, Tsinghua University, Beijing 100084, China.

Email: gongqin@mail.tsinghua.edu.cn



$r = .83$, Gorga et al., 2003). Cochlear status at the f_2 place is typically predicted from DPOAE response at a lower frequency, $2f_1-f_2$. Because DPOAEs must be resolved in background noise, whose energy increases as frequency decreases, $2f_1-f_2$ DPOAE responses are more susceptible to the higher noise levels below the f_2 frequency, resulting in poor performance for 0.5 kHz and other low frequencies (Dorn et al., 1999; Gorga et al., 1993a, 1993b, 1997; Kim et al., 1996; Stover et al., 1996). Also, the interpretation of DPOAE thresholds is complicated by the certitude that DPOAEs contain multiple contributions from both distortion and linear coherent reflection (Dhar & Shaffer, 2004; Johnson et al., 2007; Shera, 2004; Shera & Guinan, 1999; Stover et al., 1999; Talmadge et al., 1999; Vetešník et al., 2009).

Transient-evoked otoacoustic emissions (TEOAEs) are able to identify hearing status at 2 and 4 kHz (Gorga et al., 1993a; Hurley & Musiek, 1994; Hussain et al., 1998; Lichtenstein & Stapells, 1996; Prieve et al., 1993) but not at 0.5 kHz (Gorga et al., 1993a; Prieve et al., 1993). Several previous studies (Gorga et al., 1993a; Hurley & Musiek, 1994; Hussain et al., 1998; Lichtenstein & Stapells, 1996; Prieve et al., 1993) failed to measure TEOAEs greater than 4 kHz due to their analysis methods described by Bray and Kemp (1987) and Kemp et al. (1990), in which the first 2.5 ms of TEOAEs were set zero, and an onset ramp was applied from 2.5 to 5.0 ms to reduce stimulus artifact. Because TEOAE latencies decreased with increasing frequency, elimination of TEOAEs' first 5 ms reduced high-frequency (>4 kHz) TEOAEs. Later studies (Goodman et al., 2009; Keefe et al., 2011) adopted a new technique based on the double-evoked paradigm to measure TEOAEs up to 16 kHz, suggesting the clinical potential of TEOAEs in predicting hearing status from at least 1 to 10 kHz.

Stimulus-frequency otoacoustic emissions (SFOAEs) are measured at the same frequency as the probe tone within the cochlea, providing frequency-specific responses. However, SFOAEs have received less attention in clinical applications than DPOAEs and TEOAEs. Avan et al. (1991) found that audiometric thresholds at 1.5 and 2 kHz were significantly correlated with SFOAE thresholds at 0.75 and 1 kHz, respectively. Ellison and Keefe (2005) showed that SFOAEs can distinguish between normal and impaired ears from 0.5 to 8 kHz. Although clinical decision theory (Swets, 1988) has been widely used to identify the presence or absence of hearing loss (Ellison & Keefe, 2005; Go et al., 2019; Gorga et al., 1997; Stover et al., 1996), there is still much to learn of the ability of SFOAEs to quantitatively predict pure-tone thresholds.

Artificial neural networks (ANN) are mathematical models comprising many nodes ("neurons") arranged

in layers connected to each other. Each neuron sums weighted inputs and then applies a certain function to the sum to reach the output. The ANN that has been received most attention is the back propagation (BP) neural network (Rumelhart et al., 1986). A BP neural network is constructed with at least three layers: An input layer receives and distributes the input pattern, one or more hidden layers capture the nonlinearities of input-output relationship, and one layer, the output layer, produces the output pattern. It uses a supervised learning technique called BP for training with the advantages of being able to approximate any nonlinear function with satisfactory precision and capture useful information from patterns. Furthermore, BP neural network is widely used due to its strong generalization ability, which refers to the ability of applying the trained model to new samples. Upon "training" with many trials under supervision, the BP neural network "learns" the input-output relations, and then the model can be applied to other (different) samples. Here, we use a BP neural network to systematically assess the ability of SFOAEs to predict pure-tone thresholds.

Materials and Methods

Subjects

Data were collected monaurally from 230 ears of 123 subjects (62 females) with normal hearing (NH) and 737 ears of 538 subjects (256 females) with sensorineural hearing loss (SNHL) due to cochlear lesions (i.e., a loss of hair cell function). Normal-hearing subjects had air-conduction (AC) thresholds for both ears equal to or less than 25 dB HL between 0.25 and 8 kHz, with age ranging from 18 to 42 years (mean = 23.78 years, standard deviation [SD] = 4.13 years). For subjects with SNHL, AC thresholds were greater than 25 dB HL and less than or equal to 75 dB HL for at least one octave frequency between 0.5 and 8 kHz. Their ages ranged from 12 to 75 years (mean = 47.25 years, SD = 14.37 years). All participants had normal middle-ear function. The SNHL group was divided into three subcategories on a frequency-by-frequency basis, which were mild (i.e., >25 dB HL and \leq 40 dB HL), moderate (i.e., \geq 45 dB HL and \leq 60 dB HL), and severe (i.e., \geq 65 dB HL), respectively. Thus, it was possible for an individual ear to be classified as having both moderate and severe hearing loss at different separate frequencies. Table 1 lists the total number of normal and impaired ears for each test frequency (the number of NH/SNHL: 218/227, 198/244, 206/238, 218/250, and 214/244 at 0.5, 1, 2, 4, and 8 kHz, respectively). During the SFOAE test, all subjects sat comfortably on the recliner in the sound-attenuating chamber and were instructed to sleep or watch silent films with subtitles, avoiding gnashing,

chewing, and swallowing to reduce noises. All subjects were informed of all experimental procedures and objectives and provided written, informed consent. They were given appropriate compensation. All procedures were approved by the institutional review board at Tsinghua University (IRB00008273).

Stimulus Generation and SFOAE Recording

Stimulus generation and SFOAE recording were performed using a custom software program. Digital-to-analog conversions and analog-to-digital conversions were accomplished with a 24-bit sound card (Fireface 800, RME, Haimhausen, German) using a sampling rate of 48 kHz. Stimuli were presented to the ear via an insert earphone (ER-2, Etymotic Research, Elk Grove Village, IL, USA), and responses were recorded using a low-noise microphone (ER-10B+, Etymotic Research, Elk Grove Village, IL, USA) with an amplification of 20 dB. Prior to data collection, stimuli were calibrated in a Brüel & Kjær ear simulator (type 4157; IEC 711 standard) at half-octave frequencies from 0.125 to 8 kHz.

SFOAEs were recorded using a procedure based on the two-tone suppression method (Brass & Kemp, 1993). Figure 1 shows how the probe and suppressor tones are

presented for a single SFOAE acquisition. Interval M and N were added to the traditional four-interval paradigm to eliminate the effects of system delay and SFOAE latency. There was an interval of $2T_d$ followed by five intervals of T_w in duration (~ 50 ms). T_d was the system delay of 14.5 ms in duration measured in advance. The stimuli comprising the probe and suppressor tones were delivered by two earphones. The probe tone was continuous pure tone with the same polarity in Intervals A, B, C, D, and N. The suppressor was a tone burst, with rise and decay time windowed by a 5-ms cosine window. Interval D of the suppressor was inverted in phase relative to Interval C. Given the pressure responses measured in Intervals A, B, C, and D (p_1, p_2, p_3, p_4), the SFOAE residual was $[p_1 + p_2 - (p_3 + p_4)]/2$, and the stimulus pressure response in the ear canal in the absence of SFOAEs equaled $(p_3 + p_4)/2$. A real-time high-pass filter (cutoff frequency of 500 Hz for test frequencies from 1 to 8 kHz and cutoff frequency of 350 Hz for test frequency of 0.5 kHz) was used to reduce low-frequency noises. The noise floor at the probe frequency was calculated from the spectrum of the subtraction between time-domain averages stored in two separate buffers, with one buffer containing odd-numbered recordings and the other buffer containing even-numbered recordings. In the present work, SFOAE transfer function (T_{sf}) magnitude was SFOAE amplitude normalized to the ear-canal sound pressure level of stimulus.

Table 1. Summary of the Number of Ears in Each Category for Each Test Frequency.

Frequency (kHz)	Normal	Impaired			Total
		Mild	Moderate	Severe	
0.5	218	80	102	45	227
1	198	81	103	60	244
2	206	81	96	61	238
4	218	80	107	63	250
8	214	81	103	60	244

Procedure

All subjects underwent an external auditory canal examination prior to the test, and cerumen (if present) was removed from the ear canal. Pure-tone AC and bone-conduction threshold from 0.25 to 8 kHz were measured in 5-dB steps on a clinical diagnostic audiometer (Otometrics, Denmark Inc., Astera). Tympanometry

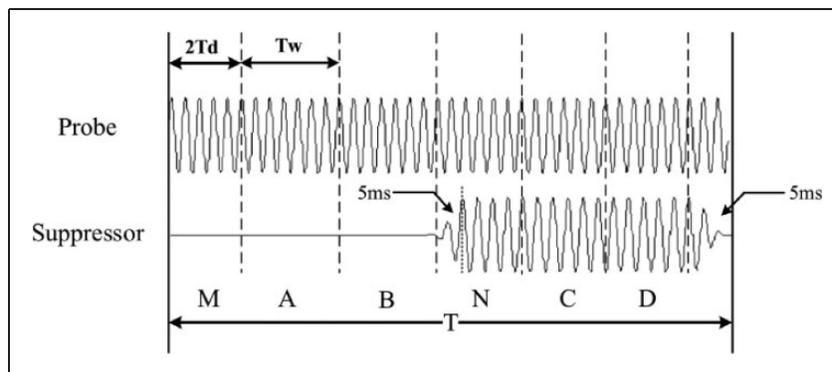


Figure 1. Presentation of Probe and Suppressor Tones for a Single SFOAE Acquisition. Top line shows the presentation of probe tones, and the bottom line shows the presentation of suppressor tones for a single SFOAE acquisition. The tones are presented in six consecutive Intervals M, A, B, N, C, and D, and the duration of the first interval is $2T_d$, followed by five intervals of T_w in duration. The probe tone has the same polarity in Intervals A, B, C, D, and N. The suppressor is presented inverted in Intervals C and D. The SFOAE residual is the subtraction of the sound pressure in intervals (A + B) and (C + D) (Gong et al., 2014).

was performed using a 0.226-kHz probe via a clinical middle-ear analyzer (Grason-Stadler Inc., TympanStar). Normal 0.226-kHz tympanometry and air-bone gaps of 10 dB or less altogether ensured that all participants had normal middle-ear function. Normal tympanometry required peak pressure between -83 and 0 daPa, peak-compensated admittance between 0.3 and 1.4 mmhos, and equivalent ear-canal volume between 0.6 and 1.5 ml. To avoid interference from spontaneous otoacoustic emissions (SOAEs), the test frequencies with strong SOAEs (i.e., peak amplitudes > 3 dB) ± 300 Hz around the center frequencies of 1 , 2 , 4 , 8 kHz and ± 150 Hz around the center frequency of 0.5 kHz were excluded. SFOAEs were not measured in 1.5% , 5.5% , 3.6% , 2.0% , and 1.2% of ears due to the presence of SOAEs at 0.5 , 1 , 2 , 4 , and 8 kHz, respectively. SFOAE I/O functions from 0.5 to 8 kHz were measured by fixing the probe frequency f_p and the suppressor frequency ($f_s = f_p - 47$ Hz), with the probe level L_p increasing in 5 -dB increments from 5 to 70 dB sound pressure level (SPL) at 0.5 , 1 , 2 , 8 kHz and from 5 to 60 dB SPL at 4 kHz. To obtain total suppression, suppressor level L_s was fixed at 70 dB SPL for the probe levels from 5 to 55 dB SPL and was 15 dB SPL above L_p for the probe levels of 60 , 65 , and 70 dB SPL. Because SFOAEs in response to low probe levels were difficult to elicit, we typically employed 96 averages at 5 to 10 dB SPL, 64 averages at 15 to 20 dB SPL, and 32 averages at the probe levels of 25 dB SPL or greater. A single test of I/O function lasted ~ 6 min per frequency except 4 kHz (~ 5.4 min).

Data Analyses

Part 1: SFOAEs as Predictors of Hearing Thresholds

Here, we proposed a new method based on a BP neural network and principal component analysis (PCA) to predict hearing thresholds. To test the effectiveness of this method, we compared it with the method of Boege and Janssen (2002) and Gorga et al. (2003), who did a correlation analysis between hearing thresholds and DPOAE thresholds.

Method 1

SFOAE thresholds were estimated with the approach of Boege and Janssen (2002) and Gorga et al. (2003). There were four inclusion subcriteria (collectively identical to the inclusion criterion of Method 1) for subsequent analyses. First, at least three points of the SFOAE I/O functions must have signal-to-noise ratio (SNR) ≥ 6 dB. Figure 2 shows SFOAE level (upper panel) and SFOAE pressure (lower panel) as a function of probe level at 1 kHz for Subject #6 (left panel) and Subject #12

(right panel). SFOAE I/O functions (upper panel) were converted into SFOAE pressure (μ Pa) against the probe level (lower panel; Figure 2). Linear regression analysis was performed to find linear dependencies between the SFOAE pressure and the probe level: One or two line segments were fitted to the data so as to account for the greatest variance (see Figure 2 bottom panel) (this procedure differed slightly from that of Boege and Janssen, 2002, which used only one segment). Data were included only if the slopes of the individual linear regressions were ≥ 0.2 μ Pa/dB, the variance accounted for (r^2) was ≥ 0.8 , and the standard error was ≤ 10 dB (when two linear regressions were fitted, only the linear regression for low probe levels was used). Otherwise, the data were excluded from further analyses. SFOAE threshold was taken as the probe level (in dB SPL) at which the SFOAE pressure equaled to 0 μ Pa. Linear regression analysis was performed again to determine the significance of SFOAE threshold as a predictor of hearing threshold. The mean absolute error (MAE) of the linear regression model, calculated as the mean of the absolute differences between the estimated and the measured hearing thresholds, was used to quantify the performance of the prediction of hearing thresholds.

Method 2

Inclusion Criterion. Figure 3 shows the process of extracting SFOAE threshold for three normal ears (left panel) and three ears with SNHL (right panel), respectively. An inclusion criterion different from that used for Method 1 was used to determine if thresholds could be predicted accurately in more ears. The probe level was raised in 5 -dB increments from 5 dB SPL until SFOAE SNR ≥ 9 dB (point in the dark gray-shaded area of Figure 3). This level was regarded as the SFOAE threshold if at least $N-1$ stimulus point(s) of the following N stimulus points had SNR ≥ 9 dB (N equaled to 3 if there were 3 or more stimulus points after the candidate; otherwise, N equaled the total number of stimulus points after the candidate; see top right and bottom right panel). Generally, SFOAE thresholds in impaired ears were larger than those in normal ears. If an SFOAE threshold could not be determined, this ear was excluded in further analyses.

Feature Extraction. In previous studies (Ellison & Keefe, 2005; Go et al., 2019), SFOAE level or SNR at a certain probe level was typically used as an independent predictor to predict hearing status or thresholds. In the present Method 2, feature extraction was not limited to SFOAE measurements at only one probe level. Rather, we captured as much information related to pure-tone thresholds as possible from SFOAEs measured at all probe levels. Given the likelihood of highly correlated

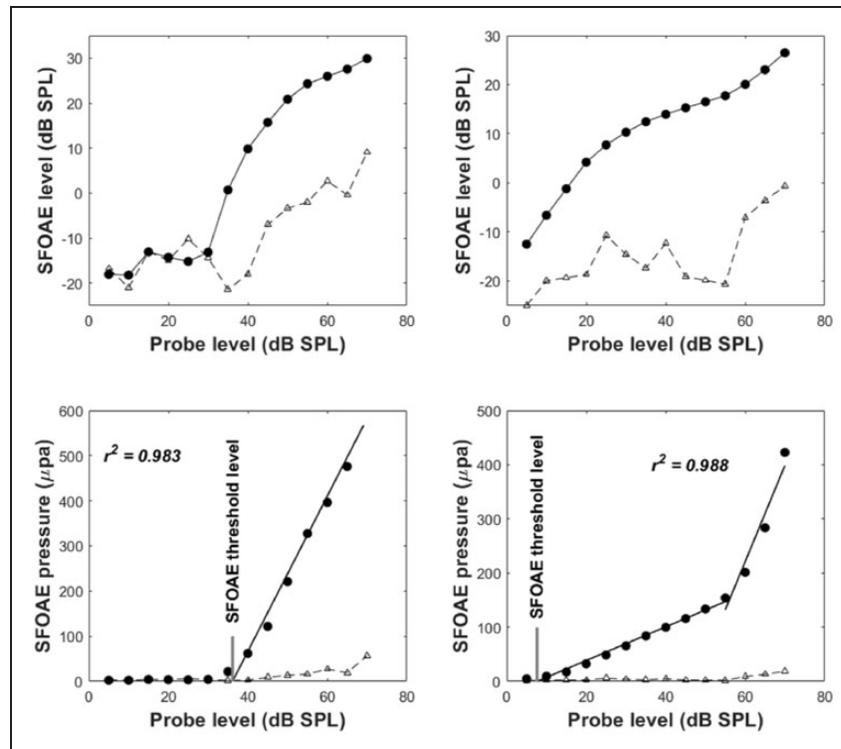


Figure 2. SFOAE I/O Function at 1 kHz for Subject #6 (Right Ear) and #12 (Left Ear). In the upper panel, the SFOAE level (dB SPL) is plotted as a function of probe level (log-log scale), and in the lower panel, the same data are plotted as SFOAE pressure as a function of probe level. In the lower panel, the solid line shows the fitted linear function (left panel: $r^2 = .983$; right panel: $r^2 = .988$) together with the estimated SFOAE threshold level (vertical line). Filled circles and open triangles represent the SFOAE levels and the noise floors, respectively.

SFOAE = stimulus-frequency otoacoustic emission; SPL = sound pressure level.

SFOAE parameters across probe levels, PCA was performed on each of the three data sets—that is, SFOAE levels (X_1), SFOAE SNRs (X_2), and T_{sf} magnitudes (X_3) measured at all probe levels—to reduce data dimension and minimize associations between parameters at different probe levels. In Method 2, the following variables were used as the input to the BP network predictor: SFOAE threshold, principal component (PC) of SFOAE levels across all probe levels (hereinafter referred to as “PC of SFOAE level”), PC of SFOAE SNRs across all probe levels (hereinafter referred to as “PC of SFOAE SNR”), and PC of T_{sf} magnitudes across all probe levels (hereinafter referred to as “PC of T_{sf} ”). Briefly, PCA was accomplished by transforming a set of associated original variables into an equal number of uncorrelated ones called principal components (PCs) by orthogonal transformation. Each PC is a linear combination of original variables and is assigned an eigenvalue. As such, the first PC (PC1) associated with the highest eigenvalue explains the most variance in the data set, the second PC (PC2) with the second highest eigenvalue accounts for the second most variance, and so on. The PCA was performed using MATLAB R2017b routine (The Mathworks INC., Natick, USA). A score

matrix was obtained by transforming each raw data matrix X_i ($i=1, 2, 3$) into the principal component space. Thus, the principal component scores were the representations of X_i in the principal component space and used in the present work as PCs. In PCA, each variable in the original data set was centered with respect to its average so that the obtained PCs had similar magnitude ranges.

Figure 4 shows the percentage of variances (i.e., information) in the original data set explained by each PC for the three data sets, as well as correlation coefficient (r) between each PC and the measured pure-tone threshold. As shown in Figure 4, the first PCs accounted for the most information in the original data set (typically more than 70% of variances). And the majority of them were more relevant to the measured pure-tone threshold than other subsequent PCs. The exception was that the relation between the second PC and the pure-tone threshold was strongest when PCA was separately performed across SFOAE levels and T_{sf} magnitudes for 0.5 kHz. There is no objective way to determine how many PCs are sufficient to explain original data set as it depends on specific application and needs. For the sake of simplicity of model on the premise of extracting useful information

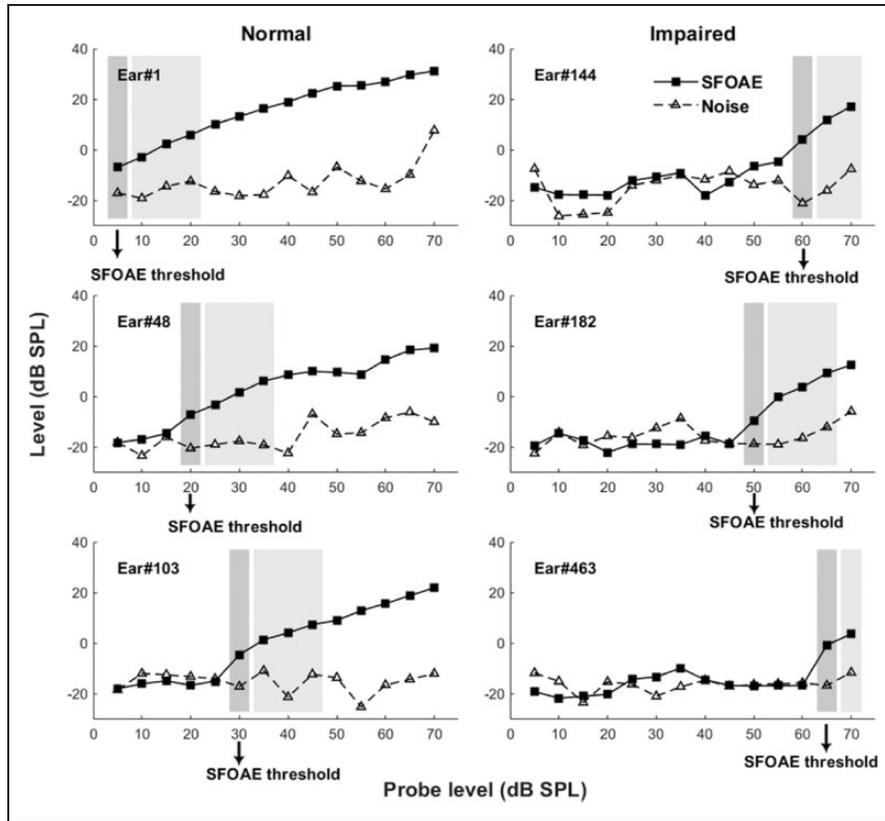


Figure 3. The Process of Extracting SFOAE Threshold for I/O Functions Meeting the Inclusion Criterion of Method 2. SFOAE = stimulus-frequency otoacoustic emission; SPL = sound pressure level.

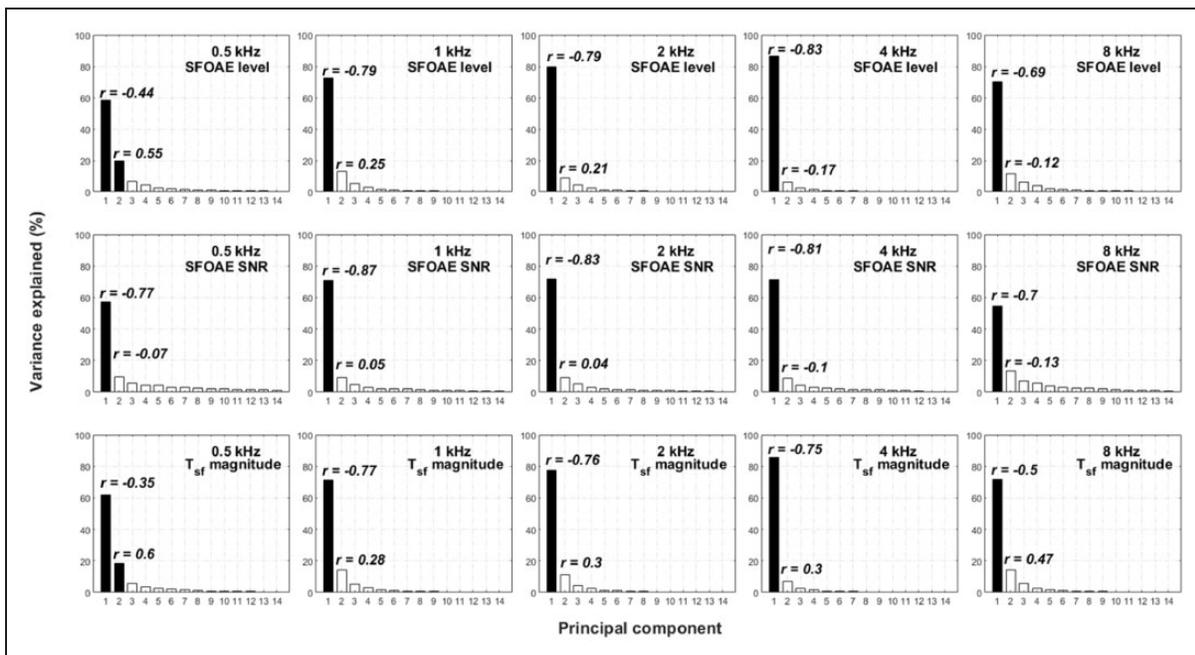


Figure 4. The Percentage of Variance in the Original Data Set Explained by Each Principal Component (PC) for SFOAE Level, SFOAE SNR, and T_{sf} Magnitude. Also shown is the Pearson correlation coefficient (r) between the first or second PCs and the measured pure-tone threshold (see the r value on the top of bar). Each column represents a frequency. In each panel, the abscissa indicates the k^{th} principal component, and the ordinate represents the percentage of variance that the corresponding PC can explain. Black bar in each panel represents the PC that explains more than 15% of variance and is selected as the input variable of the model. SFOAE = stimulus-frequency otoacoustic emission; SNR = signal-to-noise ratio.

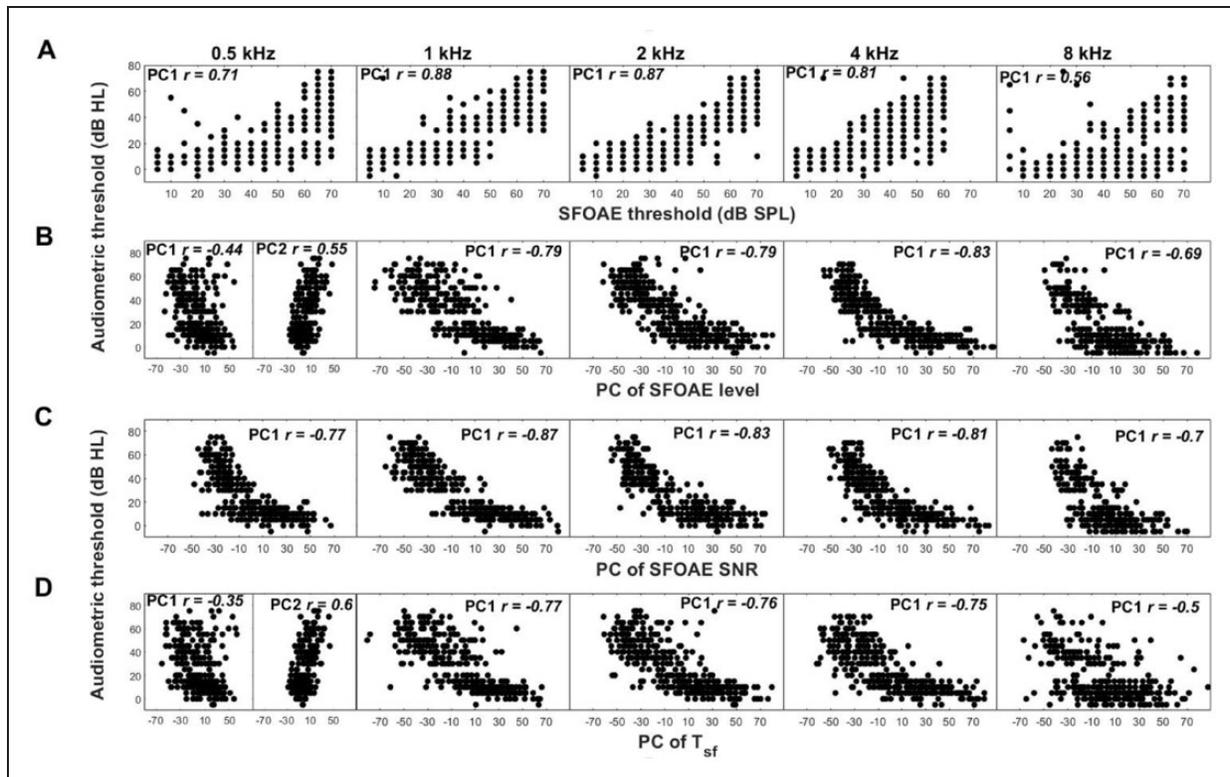


Figure 5. Audiometric Threshold (dB HL) as a Function of Each Input Variable in Method 2 From 0.5 to 8 kHz. A: Audiometric threshold (dB HL) plotted versus SFOAE threshold in Method 2. B: Audiometric threshold (dB HL) plotted versus principal component (PC) of SFOAE level. C: Audiometric threshold (dB HL) plotted versus principal component (PC) of SFOAE SNR. D: Audiometric threshold (dB HL) plotted versus principal component (PC) of T_{sf} magnitude. In each panel, also shown is the Pearson correlation coefficient (r) between audiometric threshold and each input variable. Note for 0.5 kHz, both the first and second PCs of SFOAE level and T_{sf} magnitude are plotted in this figure.

SFOAE = stimulus-frequency otoacoustic emission; SPL = sound pressure level; PC = principal component; SNR = signal-to-noise ratio.

regarding pure-tone thresholds, we retained PCs explaining more than 15% of variances as indicated by the black bars in Figure 4. Therefore, the first PCs of SFOAE levels, SFOAE SNRs, and T_{sf} magnitudes were separately selected as “PC of SFOAE level,” “PC of SFOAE SNR,” and “PC of T_{sf} ” except that the first and second PCs were jointly adopted at 0.5 kHz for SFOAE levels and T_{sf} magnitudes.

Pearson correlation analysis was performed to determine the significance of each input variable of Method 2 as a predictor of hearing threshold. Figure 5(A–D) plots the measured audiometric threshold as a function of SFOAE threshold, principal component (PC) of SFOAE level, PC of SFOAE SNR, and PC of T_{sf} magnitude from 0.5 to 8 kHz (note that for 0.5 kHz, both the first and second PC of SFOAE level and T_{sf} magnitude are plotted in Figure 5). The SFOAE threshold obtained with Method 2 was closely related to audiometric threshold for all test frequencies ($r = .71$ for 0.5 kHz, $r = .88$ for 1 kHz, $r = .87$ for 2 kHz, $r = .81$ for 4 kHz, $r = .56$ for 8 kHz, $p < .001$ for all frequencies, see Figure 5A). The PC of SFOAE level, of SFOAE SNR, and of T_{sf}

magnitude were separately significantly correlated with pure-tone threshold for each frequency ($p < .001$ for all frequencies). The correlation coefficients between the PCs of SFOAE level and pure-tone thresholds were .55, $-.79$, $-.79$, $-.83$, and $-.69$ for 0.5 to 8 kHz, respectively (see Figure 5B). The strong relationship between PC of SFOAE SNR and pure-tone threshold (see Figure 5C) was clearly demonstrated by the correlation analysis with a correlation coefficient $r = -.77$, $-.87$, $-.83$, $-.81$, and $-.7$ ($p < .001$ for all frequencies) from 0.5 to 8 kHz. Also, PC of T_{sf} magnitude was significantly correlated with pure-tone threshold (see Figure 5D, $r = .6$, $-.77$, $-.76$, $-.75$, $-.5$, respectively). Thus, all variables were suitable as the input to the BP network predictor of Method 2, providing useful information for predicting hearing thresholds.

Model Construction and Evaluation. BP neural network was used to predict hearing thresholds. As shown in Figure 6A, the structure of BP network predictors contained three layers: the input layer, the hidden layer, and the output layer. The number of nodes in the input layer

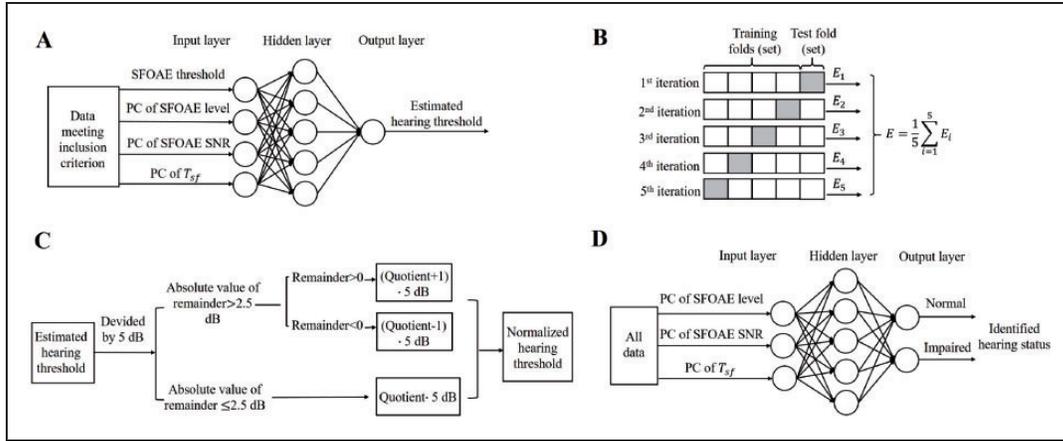


Figure 6. BP Neural Network Model Construction and Evaluation. A: The architecture of BP network predictor for estimating hearing thresholds in Method 2. B: The schematic diagram of fivefold cross-validation. C: The process of normalizing the pure-tone thresholds estimated by BP network predictor at 5-dB intervals. D: The architecture of BP network classifier for identifying hearing status. SFOAE = stimulus-frequency otoacoustic emission; PC = principal component; SNR = signal-to-noise ratio.

and hidden layer was 4 (or 6 for 0.5 kHz, i.e., the number of input variables) and 5, respectively. Only one node in the output was the estimate of hearing thresholds. Each frequency was analyzed separately and thus had its own neural network. Five experimental runs (or iterations) were conducted through fivefold cross-validation to avoid overfitting. As shown in Figure 6B, each data set was divided into five approximately equal-sized disjoint folds. Each fold is in turn a test set to validate accuracy of the model trained by the other four folds (i.e., training set). The process of network training and prediction is shown as follows:

Step 1: During the k th run, take four folds as a training set and the remaining fold as a test set (see Figure 6B). The initial connection weights among the nodes are randomly assigned first.

Step 2: The operating signal of the training set is propagated from the input layer, via the hidden layer, to the output layer. During the forward propagating process, the weights are constant, and each neuron's status only influences the next layer.

Step 3: If the expected output cannot be obtained in the output layer, it then turns to the BP of error signal (i.e., the difference between the real output and expected output of the network). In the BP of error signal, the error signal is back propagated from the output end to the input layer of the network for updating weights.

Step 4: Repeat Step 3, the weight value of network is continuously updated to make the output closer to the expected one, until the error is reduced to a set minimum value or reaching the steps of training, the weights are fixed, and network training has been completed.

Step 5: Take samples of the training set as input of the trained network, it can get prediction results of the training set. Likewise, take being predicted samples of the test

set as input of the trained network, predicted hearing thresholds can be obtained. Predicted hearing thresholds of the training set and the test set are then normalized at intervals of 5 dB according to Figure 6C.

Step 6: MAE of the training set and the test set was calculated separately for each run. After all five runs were completed, the final performance was the average of the five MAEs resulting from these five runs. Thus, mean MAE of five runs was calculated for the training set and the test set to monitor whether the model was overfitting. MAE is adopted to evaluate the performance for estimating hearing threshold, defined as Equation 1.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |Y_i - X_i| \quad (1)$$

where Y_i and X_i are the predicted and measured hearing threshold, and n represents the number of samples.

Part II: SFOAEs as Identifiers of Hearing Status

Feature Extraction. All data were included in the BP network classifiers to identify hearing status (the inclusion criterion was irrelevant, as all data were meaningful in terms of identifying hearing status). For the BP network classifier, three features (PC of SFOAE level, PC of SFOAE SNR, PC of T_{sf} magnitude) were used as input variables. The same procedure as Method 2 was performed to extract the PC of each SFOAE parameter.

Model Construction and Evaluation. The structure of the BP network classifier is shown in Figure 6D. It consists of an input layer, a hidden layer, and an output layer. The input and hidden layers were constructed with 3 (or 5 for 0.5 kHz, i.e., the number of input variables) and 5 nodes,

respectively. The two nodes in the output layer represented two classes of hearing status (normal vs. impaired). Each frequency was analyzed separately and thus had its own neural network. Fivefold cross-validation was conducted. The process of network training and prediction for the BP network classifier are in common with the aforementioned BP network predictor in Method 2. Receiver operating characteristic (ROC) curves (plots of true positive rate, which is the proportion of ears with hearing loss that were correctly identified, versus false positive rate, which is the proportion of ears with NH incorrectly classified as hearing loss) were constructed. The area under the ROC curve (AUC) and classification accuracy (i.e., the percentage of ears that were correctly identified) were used to assess performance of the BP network classifier for each frequency.

Results

Part I: SFOAEs as Predictors of Hearing Thresholds

Evaluation of the Audiometric Thresholds of Ears Not Meeting the Inclusion Criterion. Data analyzed with Method 1 were derived from SFOAE I/O functions in which at least three points had SNRs ≥ 6 dB and for which the linear fits between SFOAE pressure and probe level had slopes ≥ 0.2 $\mu\text{Pa}/\text{dB}$, $r^2 \geq 0.8$, and standard errors ≤ 10 dB. Table 2 lists the percentage of cases failing to meet the inclusion criterion of Methods 1 and 2. After applying these inclusion criteria, nearly half of data were excluded from further analyses (62.7%, 49.3%, 46.6%, 54.3%, 62.0% for 0.5 to 8 kHz, respectively). Of these excluded cases, more than 70% had audiometric thresholds exceeding 25 dB HL, with a mean threshold for these conditions of 39.3 to 46.7 dB HL for 0.5 to 8 kHz ($SD = 21.2$ – 23.3 dB). Compared with Method 1, a higher proportion of data were selected to predict hearing thresholds in Method 2 (only 23.2%, 16.2%, 17.6%, 21.2%, 34.7% of ears did not meet the inclusion

criterion for 0.5 to 8 kHz, respectively). More than 90% of these excluded ears had hearing loss (i.e., audiometric threshold > 25 dB HL), with a mean threshold for these conditions of 51.6 to 58.3 dB HL for 0.5 to 8 kHz ($SD = 11.7$ – 17.2 dB).

Performance of Method 1 in Predicting Hearing Thresholds.

The measured audiometric thresholds were plotted as a function of estimated SFOAE threshold for each of test frequencies in Figure 7. Linear regression analyses revealed that SFOAE threshold was significantly correlated with audiometric threshold for all frequencies—correlation coefficients r of .6, .86, .79, .68, .49 ($p < .001$ for all frequencies) from 0.5 to 8 kHz, respectively. MAE was used to quantify the prediction performance of the linear regression model. As shown in Figure 7, the best performance was achieved at 1 kHz, with a MAE of 6.34 dB ($SD = 5.85$ dB). MAEs were 8.47 ($SD = 7.82$ dB), 7.72 ($SD = 6.32$ dB), and 8.28 dB ($SD = 7.88$ dB) for 0.5, 2, and 4 kHz, respectively. The poorest performance occurred at 8 kHz (MAE of 9.96 dB; $SD = 7.57$ dB).

Performance of Method 2 in Predicting Hearing Threshold. The BP network predictor in Method 2 can predict hearing thresholds. The prediction performance of the predictor was quantified by the MAE. Table 3 lists mean MAE values of fivefold cross-validation for each test frequency. There was no overfitting in the model as the mean MAE of the training set was quite close to that of the test set, with the difference of mean MAE between the training set and the test set not exceeding 0.05 dB at each f_p . The peak performance was achieved at 1 kHz, where the mean MAE of the test set (6.25 dB) was lowest, and the percentage of ears meeting the inclusion criterion of Method 2 was highest. Also, excellent performance in predicting pure-tone thresholds was observed at 0.5, 2, and 4 kHz, with the mean MAE of the test set ranging

Table 2. Percentage of Ears Failing to Meet the Inclusion Criterion of Method 1 and Method 2, the Percentage of These Ears That Have Audiometric Thresholds Greater Than 25 dB HL, and the Mean and Standard Deviations for These Ears at Each Frequency.

Method	Frequency (kHz)	% of conditions failing inclusion criterion	% of thresholds > 25 dB HL	Mean thresholds (dB HL)	Standard deviation (dB HL)
Method 1	0.5	62.70	71.33	39.28	21.23
	1	49.32	80.28	45.21	21.35
	2	46.62	82.13	46.74	20.61
	4	54.27	74.02	43.50	21.63
	8	62.01	74.30	40.62	23.31
Method 2	0.5	23.15	94.17	51.89	15.78
	1	16.17	98.65	56.49	14.96
	2	17.57	97.44	56.92	14.31
	4	21.15	97.98	58.33	11.72
	8	34.72	91.82	51.64	17.23

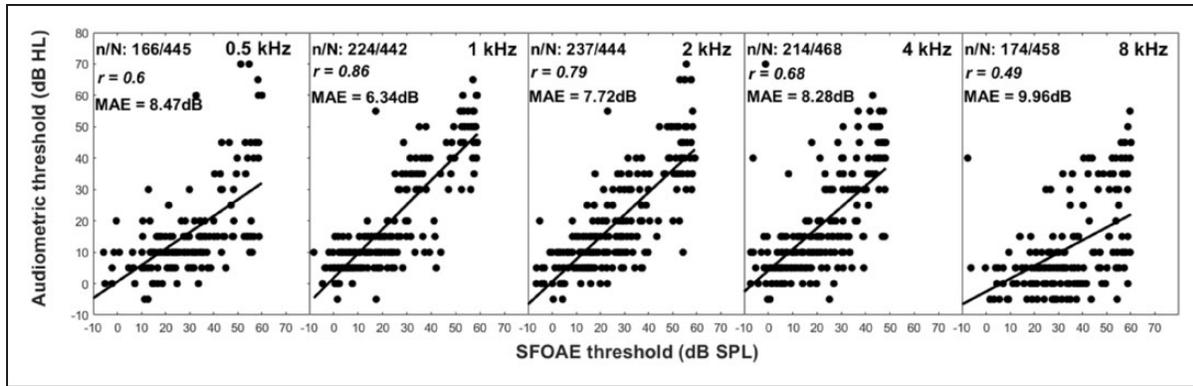


Figure 7. Audiometric Threshold (dB HL) as a Function of SFOAE Threshold (dB SPL) Estimated With Method 1 From 0.5 to 8 kHz. The solid lines represent the best-fit line to the data in each panel. Also indicated is the Pearson correlation coefficient (r), the number of threshold comparisons (n) (i.e., the number of I/O functions meeting all inclusion subcriteria of Method 1), the total number of I/O functions (N), and the MAE of the linear regression model. MAE = mean absolute error; SFOAE = stimulus-frequency otoacoustic emission; SPL = sound pressure level.

Table 3. Mean MAE of Fivefold Cross-Validation When Using Method 2's BP Network Predictor to Estimate Hearing Thresholds.

Performance metric	Frequency (kHz)	0.5	1	2	4	8
Mean MAE (dB)	N	342	368	366	369	299
	Training set	7.628	6.243	6.896	6.446	9.173
	Test set	7.634	6.246	6.912	6.464	9.185

Note. N represents the number of I/O functions meeting the inclusion criterion of Method 2. MAE = mean absolute error.

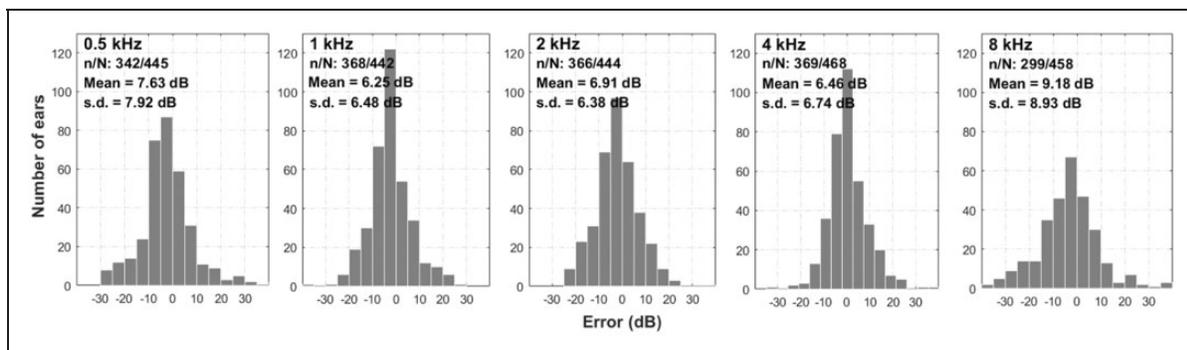


Figure 8. The Distribution of Hearing Threshold Prediction Error (i.e., the Difference Between the Estimated and Measured Hearing Thresholds) for Each Frequency When Using Method 2. Also shown is the number of I/O functions meeting inclusion criterion of Method 2 (n), and the total number of I/O functions (N), the mean of absolute error (mean), and the standard deviation (SD).

from 6.46 to 7.63 dB. Compared with lower frequencies, hearing thresholds were predicted with poorer performance at 8 kHz, where the mean MAE of the test set was 9.19 dB and 34.7% of I/O function data were excluded due to a failure to meet the inclusion criterion.

Fivefold cross-validation was conducted for modeling and evaluation. Each fold was in turn a test fold (or a test set of each run, see Figure 6B) so that each of the five folds was used exactly once as the test samples to test the model without repeating. Figure 8 shows that all test

folds of five runs were collectively used to plot the distribution of prediction error (i.e., the difference between the estimated and measured hearing thresholds) for each frequency. Each panel shows the result for a different frequency, going from 0.5 (left panel) to 8 kHz (right panel). In most ears, predictions with Method 2 were within ± 10 dB of measured pure-tone thresholds across all test frequencies, but a (very) few samples had large errors (≥ 20 dB), especially for 0.5 and 8 kHz. Also shown is the mean and SD of absolute error

(averaged across all test samples at once), which are slightly different from the mean MAE of the test set in Table 4. Both ways of error calculation for fivefold cross-validation in the present work was reasonable, just for the purpose of interpreting the results from different perspectives.

Table 4 compares the percentage of cases meeting the inclusion criterion and MAE for Methods 1 and 2. It can be seen that Method 2 performed better than Method 1 in predicting hearing thresholds at all test frequencies as a larger number of ears met the inclusion criterion for Method 2 and a lower MAE was observed with Method 2. An additional analysis that Method 2 was applied to the same data set as Method 1 (i.e., the data meeting the inclusion criteria for Method 1 instead of Method 2's own inclusion criterion) was performed for each frequency, which resulted in test MAE of 6.70, 5.02, 6.56, 5.63, and 8.19 dB for 0.5, 1, 2, 4, and 8 kHz, respectively. Based on the same inclusion criterion, the lower MAE for the additional analysis than Method 1 further verified the advantage of Method 2.

Table 4. Performance Comparison Between Method 1 and Method 2.

Method	Frequency (kHz)	% of conditions meeting inclusion criterion	MAE (dB)	Standard deviation of absolute error (dB)
Method 1	0.5	37.30	8.47	7.82
	1	50.68	6.34	5.85
	2	53.38	7.72	6.32
	4	45.73	8.28	7.88
	8	37.99	9.96	7.57
Method 2	0.5	76.85	7.63	7.92
	1	83.83	6.25	6.48
	2	82.43	6.91	6.38
	4	78.85	6.46	6.74
	8	65.28	9.18	8.93

Note. MAE = mean absolute error.

To compare how well thresholds in ears with different degrees of hearing loss were correctly estimated with Method 2, the mean and SD of absolute error (i.e., the absolute difference between the estimated and measured hearing thresholds) for each category were calculated as shown in Table 5. Hearing thresholds in ears with severe hearing loss were correctly predicted least often, resulting in the largest MAEs compared with other categories at each frequency. Small errors were observed in ears with NH and moderate hearing loss. Compared with ears with NH and moderate hearing loss, mild loss group exhibited larger MAEs.

Part II: SFOAEs as Identifiers of Hearing Status

A BP network classifier was built to identify hearing status for all tested ears using an NH criterion of 25 dB HL. Fivefold cross-validation was conducted. The mean ROC curve and average AUC value of fivefold cross-validation for each frequency are shown in Figure 9A. Figure 9B compares the AUC in the present study with that of Ellison and Keefe (2005). The performance of the BP network classifier was also evaluated according to its accuracy (i.e., the percentage of ears that were correctly identified; see Figure 9C). It can be reasonably assumed that the models at all frequencies did not overfit the data as the accuracies of the training set and the test set were nearly the same. These results showed that the BP network classifier exhibited excellent performance at all test frequencies. The mean AUC exceeding 0.97 and accuracy of more than 92.1% were observed at frequencies from 0.5 to 4 kHz. The best performance was achieved at 1 kHz, which resulted in a largest AUC of 0.99 ± 0.009 and highest accuracy of 94.1%. Ears were less often correctly identified at 8 kHz than other frequencies, with classification accuracy of 88.0% and AUC of 0.94 ± 0.02 .

Discussion

Two methods were used here to predict hearing thresholds from SFOAEs. Method 1 used a linear regression

Table 5. Mean (M) and Standard Deviation (SD) of Absolute Error in Method 2 for Each Category: Normal, Mild, Moderate, and Severe Hearing Loss.

	Frequency (kHz)	0.5	1	2	4	8
Normal	N	212	197	204	216	201
	M ± SD (dB)	6.20 ± 7.72	3.96 ± 5.06	5.49 ± 5.85	5.81 ± 6.85	7.24 ± 7.71
Mild	N	58	70	69	68	54
	M ± SD (dB)	8.88 ± 7.26	10.6 ± 6.96	9.86 ± 6.64	8.31 ± 5.83	14.3 ± 8.87
Moderate	N	57	75	67	71	33
	M ± SD (dB)	8.77 ± 7.09	5.53 ± 5.30	5.07 ± 4.48	4.51 ± 4.32	8.03 ± 8.10
Severe	N	15	26	26	14	11
	M ± SD (dB)	18.7 ± 6.40	13.8 ± 6.53	15 ± 5.10	17.5 ± 7.53	23.2 ± 11.2

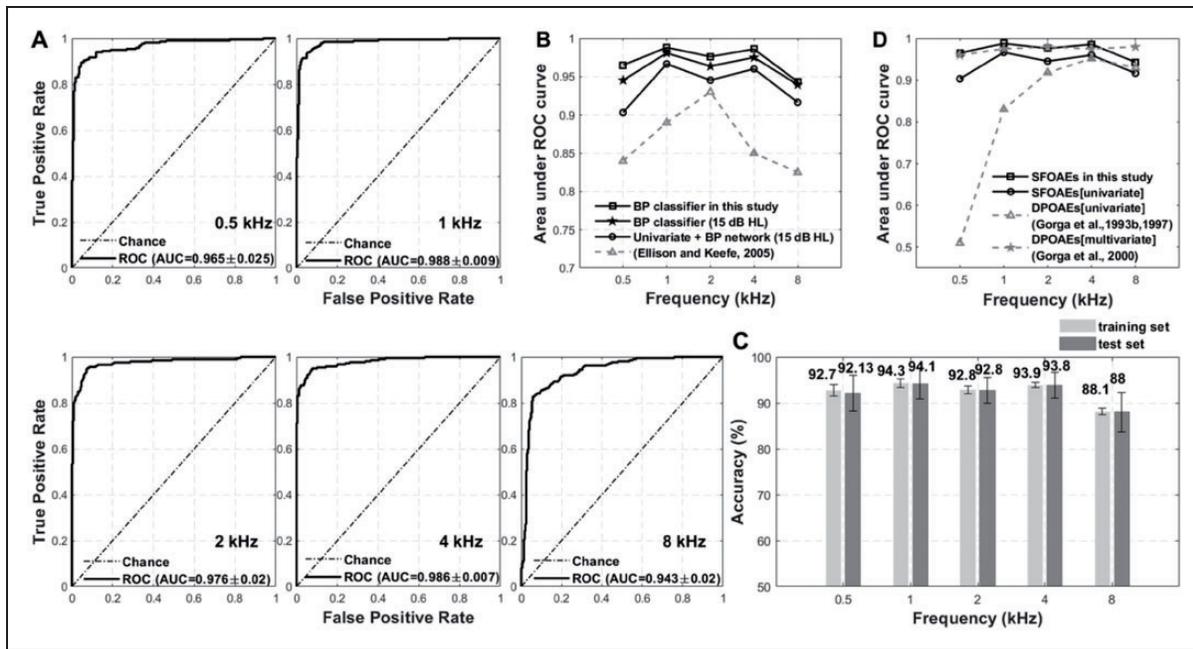


Figure 9. Performance of the BP Network Classifier for Identifying Hearing Status. Training and testing were conducted with fivefold cross-validation. A: The ROC curve for the classifier. B: The mean AUC of fivefold cross-validation for the BP network classifiers at all test frequencies (squares), the BP network classifier in this study using a normal-hearing criterion of 15 dB HL (stars), a BP network using univariate SFOAE as the input based on 15 dB HL criterion (circles), and the best AUC from a previous SFOAE study (Ellison & Keefe, 2005; triangles). C: Classification accuracy (%) of the training set and the test set for each frequency. The error bars represent the standard deviations of the fivefold cross-validation. D: The AUC obtained with SFOAEs in this study was compared with univariate (Gorga et al., 1993b, 1997) and multivariate DPOAE models (Gorga et al., 2000). Note the AUCs obtained in Gorga et al. (2000) are approximations based on the published plots.

ROC = receiver operating characteristic; AUC = area under the ROC curve; BP = back propagation; SFOAEs = stimulus-frequency otoacoustic emissions; DPOAEs = distortion-product otoacoustic emissions.

model with estimated SFOAE threshold as an independent variable according to the approach of Boege and Janssen (2002) and Gorga et al. (2003). Method 2, based on a BP neural network, performed better than Method 1 at all frequencies, as revealed by lower MAEs and higher percentage of ears meeting the inclusion criterion. The better performance of Method 2 may result from the use of PCA, which contributed to maximize the extraction of pure-tone threshold information from SFOAEs, and of multiple variables (SFOAE threshold, principal component of SFOAE level, SFOAE SNR, and T_{sf} magnitude) rather than a single variable. Moreover, the BP neural network algorithm was superior to linear regression because it could approximate any function (linear or nonlinear) with satisfactory precision and captured useful information from patterns.

The BP network predictors of Method 2 performed well in estimating hearing thresholds at all test frequencies, but prediction performance differed across frequency: Better performance was observed at 1 to 4 kHz than at 0.5 kHz, probably due to increased noise levels with frequency decreasing during SFOAE measurement at 0.5 kHz (much as previously found in studies involving

DPOAEs and TEOAEs; Gorga et al., 1993b; Prieve et al., 1993). SFOAEs were weaker at 8 kHz than at lower frequencies (as found by Dewey & Dhar, 2017; Dhar & Shaffer, 2004) and hence were difficult to separate from noise, causing a larger proportion of ears not meeting the inclusion criterion and larger MAEs.

Using Method 2, large errors in predicting high hearing thresholds (with severe hearing loss) probably resulted from too small or unreliable SFOAEs. It is well known that OAEs are generated as a by-product of the normal function of outer hair cells in the cochlea, and outer hair cell-related SNHL generally accounts for hearing loss no more than 60 dB HL. It may be also the reason why ears with severe hearing loss were almost unable to be correctly predicted. Poor prediction performance also occurred in ears with mild hearing loss, consistent with previous studies (Ellison & Keefe, 2005; Gorga et al., 1997).

The BP network classifiers achieved excellent performance in determining the presence or absence of hearing loss across all test frequencies, with performance better than in the SFOAE study of Ellison and Keefe (2005) regardless of using a normal audiometric criterion of 15

or 25 dB HL (see Figure 9B). In another set of tests with the 15 dB HL (i.e., the NH criterion used by Ellison and Keefe, 2005), SFOAE level or SNR at moderate probe level (50- or 60-dB SPL) were taken as the univariate input to a BP neural network. As shown in Figure 9B, the best AUC obtained with univariate analysis in this study was larger than the AUC in Ellison and Keefe (2005) but lower than that of the BP network classifier in the present study for each frequency. Thus, it seems, the improved performance in the present study compared with that of Ellison and Keefe (2005) reflects the advantage of multivariate models over univariate models and the use of BP neural network, as well as PCA. In addition, the present study excluded a small number of ears with strong SOAEs while these were included in Ellison and Keefe (2005).

Several investigations have shown that DPOAEs can be used to predict hearing status. Figure 9D compares the AUCs of the present SFOAE study with those of univariate and multivariate DPOAE models (Gorga et al., 1993b, 1997, 2000). The performance of SFOAEs and DPOAEs in identifying hearing status was generally similar except that SFOAEs were slightly poorer than DPOAEs for 8 kHz. Also, univariate SFOAE models were superior to univariate DPOAE models for 0.5 and 1 kHz. Standard error was also calculated for Method 2 to make comparison between the performance of SFOAEs and DPOAEs in predicting thresholds, as shown in Table 6. SFOAEs performed better in predicting hearing thresholds than DPOAEs for 1, 2, and 8 kHz (Gorga et al., 2003; Johnson et al., 2007), as evidenced by a much higher percentage of ears meeting the inclusion criterion and lower standard error for SFOAEs (see Table 6). For 0.5 kHz, despite a larger standard error for SFOAEs than DPOAEs, SFOAEs appeared to be superior to DPOAEs as the result for DPOAEs (Gorga et al., 2003) at this frequency was obtained from only 17% of ears meeting the inclusion criterion (27 ears), while a significantly larger proportion (76.9%) of ears met the inclusion criterion (342 ears) in

the present SFOAE study. Similar performance of SFOAEs and DPOAEs in threshold prediction was observed for 4 kHz (Johnson et al., 2007). Thus, SFOAEs have similar potential to DPOAEs in the identification of SNHL and improve upon the prediction of hearing thresholds at some frequencies. A more complete comparison of the prediction performance of SFOAEs and DPOAEs would be to carry out all these tests on the same subjects. It is also clear that recording times should be shortened and signal extraction simplified prior to clinical applications.

The present study has two limitations. One is that ears failing to meet the inclusion criterion were excluded for analysis of the regression model. Another is that SFOAEs were measured at a set of standard audiometric frequencies (i.e., 0.5, 1, 2, 4, 8 kHz). SFOAE spectra are plagued by deep notches whose frequencies differ across ears. These notches in the SFOAE spectra might shift to the recording standard frequency as level increases, and thus, level-dependent notches would be observed in some SFOAE I/O functions. The preexisting notches may have reduced SFOAE level and SNR, thus leading to the overestimation of the hearing thresholds. The choice of the input variables in this study (extracting the principal components from SFOAEs at all probe levels, instead of relying solely on a single level at which a notch might happen to occur) probably minimized the effects of notches. However, a better methodology in future study would be to measure SFOAE I/O functions and corresponding audiometric thresholds at frequencies chosen individually for each ear to avoid these notches, even if they differ somewhat from the standard values.

In conclusion, SFOAEs can quantitatively predict hearing thresholds at octave frequencies from 0.5 to 8 kHz, with best performance at 0.5 to 4 kHz. In addition, SFOAEs can identify hearing status with great accuracy at all test frequencies. Further work is needed to improve prediction accuracy in ears with mild hearing

Table 6. Comparison of Threshold Prediction Performance Between SFOAEs in the Present Study and DPOAEs in the Studies of Gorga et al. (2003) and Johnson et al. (2007).

Study	% of cases meeting the inclusion criterion					Standard error (dB)				
	Frequency (kHz)					Frequency (kHz)				
	0.5	1	2	4	8	0.5	1	2	4	8
SFOAEs for Method 2 in this study	76.9	83.8	82.4	78.9	65.3	11.0	9.0	9.4	9.3	12.8
DPOAEs in Gorga et al. (2003)	17.1	32.8	40.3	54.8	30.0	9.0	11.6	10.6	11.2	19.2
DPOAEs in Johnson et al. (2007)	—	—	57.1	80.0	—	—	—	9.9	10.3	—

Note. Standard error is also calculated for Method 2 in this study to compare with the results for DPOAEs, also shown is the percentage of cases meeting the inclusion criterion. Dashes indicated that predictions were not reported at that frequency. SFOAEs = stimulus-frequency otoacoustic emissions; DPOAEs = distortion-product otoacoustic emissions.

loss and reduce the test time to improve the clinical potential of SFOAEs.

Acknowledgments

We thank Professor Mario Ruggero (Northwestern University) for helping with text editing and Fei Ji (The General Hospital of the People's Liberation Army) for helping with data collection.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Foundation of China (grant numbers 61871252 and 61271133).

ORCID iD

Qin Gong  <https://orcid.org/0000-0002-5664-8477>

References

- Avan, P., Bonfils, P., Loth, D., Narcy, P., & Trotoux, J. (1991). Quantitative assessment of human cochlear function by evoked otoacoustic emissions. *Hearing Research*, *52*(1), 99–112. [https://doi.org/10.1016/0378-5955\(91\)90191-B](https://doi.org/10.1016/0378-5955(91)90191-B)
- Boege, P., & Janssen, T. (2002). Pure-tone threshold estimation from extrapolated distortion product otoacoustic emission I/O-functions in normal and cochlear hearing loss ears. *The Journal of the Acoustical Society of America*, *111*(4), 1810–1818. <https://doi.org/10.1121/1.1460923>
- Brass, D., & Kemp, D. T. (1993). Suppression of stimulus frequency otoacoustic emissions. *The Journal of the Acoustical Society of America*, *93*(2), 920–939. <https://doi.org/10.1121/1.405453>
- Bray, P., & Kemp, D. (1987). An advanced cochlear echo technique suitable for infant screening. *British Journal of Audiology*, *21*(3), 191–204. <https://doi.org/10.3109/03005368709076405>
- Dewey, J. B., & Dhar, S. (2017). Profiles of stimulus-frequency otoacoustic emissions from 0.5 to 20 kHz in humans. *Journal of the Association for Research in Otolaryngology*, *18*(1), 89–110. <https://doi.org/10.1007/s10162-016-0588-2>
- Dhar, S., & Shaffer, L. A. (2004). Effects of a suppressor tone on distortion product otoacoustic emissions fine structure: Why a universal suppressor level is not a practical solution to obtaining single-generator DP-grams. *Ear and Hearing*, *25*(6), 573–585. <https://doi.org/10.1097/00003446-200412000-00006>
- Dorn, P. A., Piskorski, P., Gorga, M. P., Neely, S. T., & Keefe, D. H. (1999). Predicting audiometric status from distortion product otoacoustic emissions using multivariate analyses. *Ear and Hearing*, *20*(2), 149–163. <https://doi.org/10.1097/00003446-199904000-00006>
- Ellison, J. C., & Keefe, D. H. (2005). Audiometric predictions using SFOAE and middle-ear measurements. *Ear and Hearing*, *26*(5), 487–503. <https://doi.org/10.1097/01.aud.0000179692.81851.3b>
- Go, N. A., Stamper, G. C., & Johnson, T. A. (2019). Cochlear mechanisms and otoacoustic emission test performance. *Ear and Hearing*, *40*(2), 401–417. <https://doi.org/10.1097/AUD.0000000000000625>
- Gong, Q., Wang, Y., & Xian, M. (2014). An objective assessment method for frequency selectivity of the human auditory system. *Biomedical Engineering Online*, *13*(1), 171. <https://doi.org/10.1186/1475-925X-13-171>
- Goodman, S. S., Fitzpatrick, D. F., Ellison, J. C., Jesteadt, W., & Keefe, D. H. (2009). High-frequency click-evoked otoacoustic emissions and behavioral thresholds in humans. *The Journal of the Acoustical Society of America*, *125*(2), 1014–1032. <https://doi.org/10.1121/1.3056566>
- Gorga, M. P., Johnson, T. A., Kaminski, J. K., Beauchaine, K. L., Garner, C. A., & Neely, S. T. (2006). Using a combination of click- and toneburst-evoked auditory brainstem response measurements to estimate pure-tone thresholds. *Ear and Hearing*, *27*(1), 60–74. <https://doi.org/10.1097/01.aud.0000194511.14740.9c>
- Gorga, M. P., Neely, S. T., Bergman, B., Beauchaine, K. L., Kaminski, J. R., Peters, J., Schulte, L., & Jesteadt, W. (1993a). A comparison of transient-evoked and distortion product otoacoustic emissions in normal-hearing and hearing-impaired subjects. *The Journal of the Acoustical Society of America*, *94*(5), 2639–2648. <https://doi.org/10.1121/1.407348>
- Gorga, M. P., Neely, S. T., Bergman, B., Beauchaine, K. L., Kaminski, J. R., Peters, J., & Jesteadt, W. (1993b). Otoacoustic emissions from normal-hearing and hearing-impaired subjects: Distortion product responses. *The Journal of the Acoustical Society of America*, *93*(41), 2050–2060. <https://doi.org/10.1121/1.406691>
- Gorga, M. P., Neely, S. T., Dorn, P. A., & Hoover, B. M. (2003). Further efforts to predict pure-tone thresholds from distortion product otoacoustic emission input/output functions. *The Journal of the Acoustical Society of America*, *113*(6), 3275. <https://doi.org/10.1121/1.1570433>
- Gorga, M. P., Neely, S. T., Ohlrich, B., Hoover, B., Redner, J., & Peters, J. (1997). From laboratory to clinic: A large scale study of distortion product otoacoustic emissions in ears with normal hearing and ears with hearing loss. *Ear and Hearing*, *18*(6), 440–455. <https://doi.org/10.1097/00003446-199712000-00003>
- Gorga, M. P., Nelson, K., Davis, T., Dorn, P. A., & Neely, S. T. (2000). Distortion product otoacoustic emission test performance when both 2 f₁–f₂ and 2 f₂–f₁ are used to predict auditory status. *The Journal of the Acoustical Society of America*, *107*(4), 2128–2135. <https://doi.org/10.1121/1.428494>
- Hurley, R. M., & Musiek, F. E. (1994). Effectiveness of transient-evoked otoacoustic emissions (TEOAEs) in predicting hearing level. *Journal of the American Academy of Audiology*, *5*(3), 195.
- Hussain, D. M., Gorga, M. P., Neely, S. T., Keefe, D. H., & Peters, J. (1998). Transient evoked otoacoustic emissions in

- patients with normal hearing and in patients with hearing loss. *Ear and Hearing*, 19(6), 434–449. <https://doi.org/10.1097/00003446-199812000-00005>
- Johnson, T. A., & Brown, C. J. (2005). Threshold prediction using the auditory steady-state response and the tone burst auditory brain stem response: A within-subject comparison. *Ear and Hearing*, 26(6), 559–576. <https://doi.org/10.1097/01.aud.0000188105.75872.a3>
- Johnson, T. A., Neely, S. T., Kopun, J. G., Dierking, D. M., Tan, H., Converse, C., Kennedy, E., & Gorga, M. P. (2007). Distortion product otoacoustic emissions: Cochlear-source contributions and clinical test performance. *The Journal of the Acoustical Society of America*, 122(6), 3539–3553. <https://doi.org/10.1121/1.2799474>
- Johnson, T. A., Neely, S. T., Kopun, J. G., Dierking, D. M., Tan, H., & Gorga, M. P. (2010). Clinical test performance of distortion-product otoacoustic emissions using new stimulus conditions. *Ear and Hearing*, 31(1), 74–83. <https://doi.org/10.1097/AUD.0b013e3181b71924>
- Keefe, D. H., Goodman, S. S., Ellison, J. C., Fitzpatrick, D. F., & Gorga, M. P. (2011). Detecting high-frequency hearing loss with click-evoked otoacoustic emissions. *The Journal of the Acoustical Society of America*, 129(1), 245–261. <https://doi.org/10.1121/1.3514527>
- Kemp, D. T., Ryan, S., & Bray, P. (1990). A guide to the effective use of otoacoustic emissions. *Ear and Hearing*, 11(2), 93–105. <https://doi.org/10.1097/00003446-199004000-00004>
- Kim, D. O., Paparello, J., Jung, M. D., Smurzynski, J., & Sun, X. (1996). Distortion product otoacoustic emission test of sensorineural hearing loss: Performance regarding sensitivity, specificity and receiver operating characteristics. *Acta Oto-Laryngologica*, 116(1), 3–11. <https://doi.org/10.3109/00016489609137705>
- Lichtenstein, V., & Stapells, D. R. (1996). Frequency-specific identification of hearing loss using transient-evoked otoacoustic emissions to clicks and tones. *Hearing Research*, 98(1–2), 125–136. [https://doi.org/10.1016/0378-5955\(96\)00084-6](https://doi.org/10.1016/0378-5955(96)00084-6)
- Lightfoot, G., & Kennedy, V. (2006). Cortical electric response audiometry hearing threshold estimation: Accuracy, speed, and the effects of stimulus presentation features. *Ear and Hearing*, 27(5), 443–456. <https://doi.org/10.1097/01.aud.0000233902.53432.48>
- Musiek, F. E., & Baran, J. A. (1997). Distortion product otoacoustic emissions: Hit and false positive rates in normal-hearing and hearing-impaired subjects. *American Journal of Otology*, 18(4), 454–461.
- Norton, S. J., Gorga, M. P., Widen, J. E., Folsom, R. C., Sininger, Y., Cone-Wesson, B., Vohr, B. R., Mascher, K., & Fletcher, K. (2000). Identification of neonatal hearing impairment: Evaluation of transient evoked otoacoustic emission, distortion product otoacoustic emission, and auditory brain stem response test performance. *Ear and Hearing*, 21(5), 508–528. <https://doi.org/10.1097/00003446-200010000-00013>
- Oswald, J. A., & Janssen, T. (2003). Weighted DPOAE input/output-functions: A tool for automatic assessment of hearing loss in clinical application. *Zeitschrift für Medizinische Physik*, 13(2), 93–98. <https://doi.org/10.1078/0939-3889-00148>
- Prieve, B. A., Gorga, M. P., Schmidt, A., Neely, S., Peters, J., Schultes, L., & Jesteadt, W. (1993). Analysis of transient-evoked otoacoustic emissions in normal-hearing and hearing-impaired ears. *The Journal of the Acoustical Society of America*, 93(6), 3308–3319. <https://doi.org/10.1121/1.405715>
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating. *Nature*, 323(6088), 533–536. <https://doi.org/10.1038/323533a0>
- Shera, C. A. (2004). Mechanisms of mammalian otoacoustic emission and their implications for the clinical utility of otoacoustic emissions. *Ear and Hearing*, 25(2), 86–97. <https://doi.org/10.1097/01.AUD.0000121200.90211.83>
- Shera, C. A., & Guinan, J. J. (1999). Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs. *The Journal of the Acoustical Society of America*, 105(21), 782–798. <https://doi.org/10.1121/1.426948>
- Stover, L., Gorga, M. P., Neely, S. T., & Montoya, D. (1996). Toward optimizing the clinical utility of distortion product otoacoustic emission measurements. *The Journal of the Acoustical Society of America*, 100(2), 956–967. <https://doi.org/10.1121/1.416207>
- Stover, L. J., Neely, S. T., & Gorga, M. P. (1999). Cochlear generation of intermodulation distortion revealed by DPOAE frequency functions in normal and impaired ears. *The Journal of the Acoustical Society of America*, 106(5), 2669–2678. <https://doi.org/10.1121/1.428097>
- Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240(4857), 1285–1293. <https://doi.org/10.1126/science.3287615>
- Talmadge, C. L., Long, G. R., Tubis, A., & Dhar, S. (1999). Experimental confirmation of the two-source interference model for the fine structure of distortion product otoacoustic emissions. *The Journal of the Acoustical Society of America*, 105(1), 275–292. <https://doi.org/10.1121/1.424584>
- Van Dun, B., Dillon, H., & Seeto, M. (2015). Estimating hearing thresholds in hearing-impaired adults through objective detection of cortical auditory evoked potentials. *The Journal of the Acoustical Society of America*, 26(4), 370–383. <https://doi.org/10.3766/jaaa.26.4.5>
- Vetešník, A., Turcanu, D., Dalhoff, E., & Gummer, A. W. (2009). Extraction of sources of distortion product otoacoustic emissions by onset-decomposition. *Hearing Research*, 256(1–2), 21–38. <https://doi.org/10.1016/j.heares.2009.06.002>
- Yeung, K. N., & Wong, L. L. (2007). Prediction of hearing thresholds: Comparison of cortical evoked response audiometry and auditory steady state response audiometry techniques. *International Journal of Audiology*, 46(1), 17–25. <https://doi.org/10.1080/14992020601102238>