PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Meniere's Disease Prognosis by Learning from Transient-Evoked Otoacoustic Emission Signals

Sheng-Lun Kao[1]; Han-Wen Lien[1]; Tzu-Chi Liu[1]; Hau-Tieng Wu[2]; Te-Yung Fang[34]; Pa-Chun Wang[34]; Yi-Wen Liu[1]

[1] Department of Electrical Engineering, National Tsing Hua University, Hsinchu, Taiwan

[2] Department of Mathematics and Department of Statistical Science, Duke University, North Carolina, USA

[3] Department of Otolaryngology, Cathay General Hospital, Taipei, Taiwan

[4] School of Medicine, Fu Jen Catholic University, Taipei, Taiwan

ABSTRACT

Accurate prognosis of Meniere's disease (MD) is difficult. The aim of this study is to treat it as a machine-learning problem through the analysis of transient-evoked (TE) otoacoustic emission (OAE) data obtained from MD patients. Thirty-three patients who received treatment were recruited, and their distortion-product (DP) OAE, TEOAE, as well as pure-tone audiograms were taken longitudinally up to 6 months after being diagnosed with MD. By hindsight, the patients were separated into two groups: those whose outer hair cell (OHC) functions eventually recovered, and those that did not. TEOAE signals between 2.5-20 ms were dimension-reduced via principal component analysis, and binary classification was performed via the support vector machine. Through cross-validation, we demonstrate that the accuracy of prognosis can reach >80% based on data obtained at the first visit. Further analysis also shows that the TEOAE group delay at 1k and 2k Hz tend to be longer for the group of ears that eventually recovered their OHC functions. The group delay can further be compared between the MD-affected ear and the opposite ear. The present results suggest that TEOAE signals provide abundant information for the prognosis of MD and the information could be extracted by applying machine-learning techniques.

Keywords: Otoacoustic Emission, Meniere's Disease, Machine Learning, Signal Processing

1. INTRODUCTION

Since the discovery of otoacoustic emission (OAE) in 1978 (8), its presence has been utilized as a non-invasive indicator of the status of the outer hair cells which amplify the inner ear vibration caused by sound waves. The OAE stimulated by short broadband acoustic impulses is referred to as the transient evoked (TE) OAE. According to the characteristics of the cochlea, within 20 ms after the transient click, the high-frequency of OAE is first transmitted then the low-frequency appears (7). TEOAE and distortion product (DP) OAE tests have been widely used in clinical applications (9), such as classifying normal hearing and hearing loss by TEOAE long-latency components (11), classifying normal ears and MD ears by combining TEOAE and a masker tone of 30 Hz in a different phase (13), newborn hearing screening based on TEOAE amplitude (19), and classifying MD-affected ear and the opposite ear by TEOAE phase (12).

Few studies attempted hearing prognosis using TEOAE. Sudden sensorineural hearing loss (SSHL) prognosis by TEOAE was proposed by Amiridavan M et al. (2). Lalaki et al. (10) and Shupak et al. (17) have studied prognosis in idiopathic SSHL patients and concluded with a criterion that TEOAE SNR≥ 3 dB would be an indicator for recovery of hearing. Overall, these studies suggested that TEOAE may have prognostic values for hearing impairment of cochlear origins. To the best of our knowledge, however, MD prognosis based on TEOAE signals has not been attempted.

MD is a cochlear dysfunction possibly due to endolymph swelling, and its symptoms include severe vertigo, tinnitus, and acute hearing loss (23). MD usually affects only one ear and is more likely to happen to adults of age 40 to 60 (23). In 2010, MD is estimated that 615000 cases diagnosed in the United States and 45500 new cases occur each year (23). Accurate prognosis of MD has been known to be difficult (1) and finding a specific TEOAE pattern in MD patients has not been successful (4).

This study focuses on MD prognosis and the aim is to find useful features from TEOAE to build a predicting model using machine learning techniques.

The rest of this paper is organized as follows. Data collection and steps of proposed approach are described in Sec. 2. The machine learning model, results of classification, data visualization, and statistical analysis are shown in Sec. 3. Finally, discussions and conclusions are given in Sec. 4 and 5, respectively.

2. Data collection and signal processing

According to the longitudinal pure-tone audiometry (PTA) (3) data of the MD-affected ears, we manually label whether the patients had recover from hearing loss. Then, a predicting model was built based on the TEOAE signals measured at the first visit. The general flowchart is shown in Figure 1 and details are described in Sec. 2.1 to 2.4.

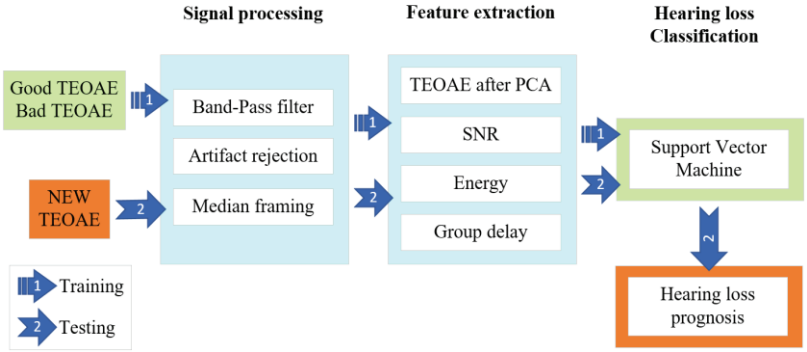

Figure 1 – A block diagram of the proposed approach for machine learning-based prognosis. The definition of the labels “Good” and “Bad” are given in Sec. 2.1.1.

2.1 Subjects

Thirty-three MD patients with unilateral MD-affected ear between age 20 and 65 were recruited and their PTA and TEOAE were longitudinally tracked during the treatment period. For each patient, after being diagnosed with MD, the PTA and TEOAE were measured every week (in week 1-4) in the beginning; after that, the measurements were repeated in the 8th week, the 12th week (the third month), and the sixth month. The PTA and TEOAE were obtained from both ears for every patient during every follow-up visit. The Institutional Review Board of Cathay General Hospital approved this study (Approval Number. CGH-P104045).

The MD affected ears were grouped into two classes by comparing the PTA of the first and last visit according to the following criteria:

- “Good” ears: if an improvement by 15 dB (e.g., the left panel of Figure 2) or more was observed in any of the PTA frequencies, or if the PTA results showed no improvement of greater than 15 dB in any frequency but the overall hearing could be categorized as normal (PTA ≤25 dB) or mild (26-40 dB) hearing loss.
- “Bad” ears: if ear has been diagnosed as MD chronic, or otherwise the PTA results showed no improvement of greater than 15 dB (e.g., the right panel of Figure 2) in any frequency while the hearing loss was categorized as moderate (PTA = 41-70 dB) to severe (>70 dB).

Following the definition above, we found 17 good cases (among which 15 ears had improved hearing, and 2 did not) and 16 bad cases (4 MD Chronic, 12 no improvement in hearing). Data from 32 non-MD affected ears were also collected.

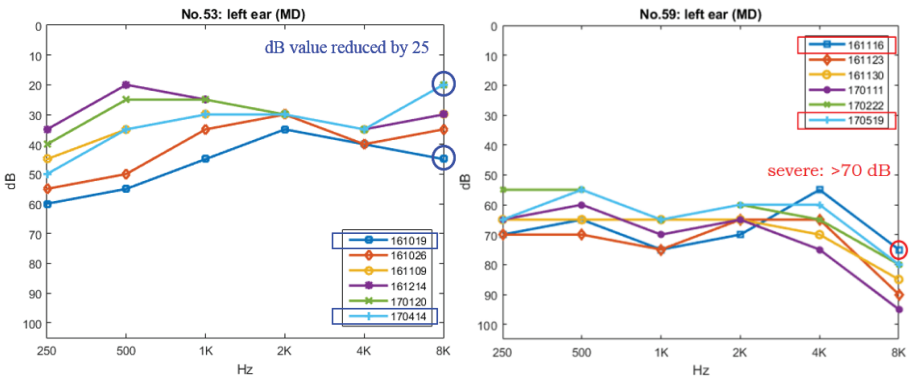

Figure 2 – Examples of pure-tone audiograms obtained longitudinally over 6 months. *Left*: a case of possible OHC function recovery in the end. *Right*: a patient whose OHC function probably did not recover. These two categories are arbitrarily named "Good" and "Bad" to refer to their OHC status.

2.2 Signal acquisition

The click stimulus was the same as described in (22). TEOAE signals were recorded by the ER-10C (Etymotic Research Inc., Elkgrove Village, IL, USA) probe connected to a 24-bit UltraLite-mk3 Hybrid external sound card (MOTU, Cambridge, MA, USA). The transient click was generated by computer and transmitted through the soundcard to the ER-10C earphone. The sampling rate was set to 44.1 kHz. The peak sound pressure level was about 74 dB relative to 20 uPa. The click rate was maintained at 10 per second. The TEOAE was recorded in a quiet office at Cathay General Hospital, Taipei, and the noise floor was approximately 23-27 dB SPL. For each recording session, 3000 clicks were played. Both the MD-affected ear and the opposite ear were measured.

2.3 Signal processing

This section describes the procedure of extracting TEOAE from the raw click responses. The process consists of the following steps.

2.3.1 Band-Pass filtering

Because the signals were not acquired in a soundbooth, power-line contamination at multiples of 60 Hz was observed with a peak at 120 Hz. To mend this problem, an 1.0-8.0 kHz bandpass filter was designed using a Kaiser window (6) of length = 90 taps. The raw signals were convolved with the impulse response of this filter in the time domain.

2.3.2 Artifact rejection

Large-amplitude acoustic noise (e.g., the sound of swallowing or the sound of friction due to body movement) exists and it must be removed before analysis of TEOAE. Each recording session is cut into 3000 frames so each frame contains one click. A frame is abandoned if any of the following criteria is met:

- The amplitude of transient click is not between 1±0.1 times is the average amplitude of clicks among all frames.
- The instantaneous sound pressure of the frame exceeds 2 mPa at any moment. Note that, empirically, human TEOAE amplitude in 2.5 to 20 ms after the click should not exceed 0.2 mPa or 20 dB SPL. Therefore, any noise at greater than 2 mPa is regarded as large deviation and should be discarded.

Also, as much as an entire file is concerned, if the total number of preserved frames is less than 700, the session is regarded as a failure and the file would not be included in the dataset.

2.3.3 Median frame computing and noise-floor estimation

The TEOAE signal is estimated by computing the median among those frames that are preserved after artifact rejection *for every sample in the time domain*. A schematic view of the procedure is shown in Figure 3, and the result is hereafter called the *median frame*. Also, to estimate the signal to noise ratio (SNR) by TEOAE repeatability (16), a noise-floor signal is first defined as follows,

$$f_{noise}(i) = \frac{f_1(i) - f_2(i) + \cdots + (-1)^{n-1} f_n(i)}{n}, \tag{1}$$

where $f_1, f_2, \ldots, f_n$ denote individual frames, i denotes the time index, and n denotes the total number of preserved frames after artifact rejection. An example of the estimated TEOAE signal (i.e., the median frame) and the corresponding $f_{noise}(i)$ for that session are shown in Figure 4. The power spectrum of this noise-floor signal enables calculation of SNR at every frequency, as shown in Figure 5.

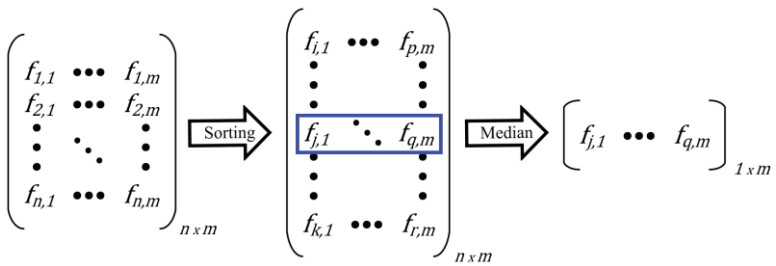

Figure 3 – A schematic illustration of computing the median frame; n denotes the number of

preserved frames after artifact rejection, and m is the length of each frame.

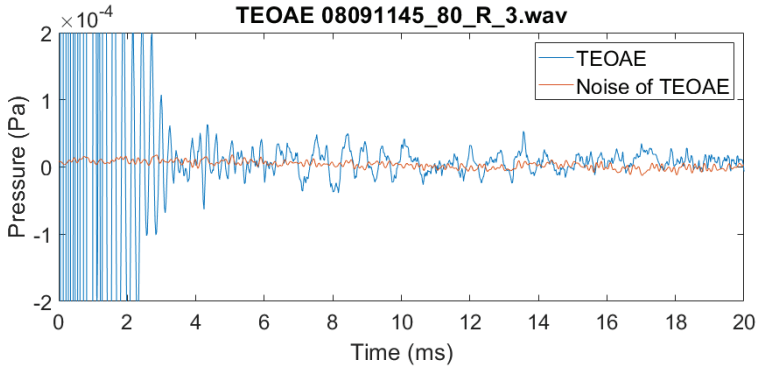

Figure 4 – An Example of the estimated TEOAE signal (blue) and the noise-floor signal (red).

2.4 Feature extraction

This section describes the procedure of transferring the TEOAE frame (2.3.4) to features. The process consists of four steps listed as follows.

2.4.1 TEOAE after principal component analysis (PCA)

After denoising the signal via taking median, an application of PCA for dimension reduction is carried out (21). The TEOAE signals in 2.5-20 ms after the click are analyzed and dimension equals to the number of samples, which is 772. In order to avoid overfitting in the subsequent supervised learning procedures, the dimensionality is reduced to three by PCA (14) first.

PCA is a statistical method of unsupervised learning which simplifies the dataset at hands. PCA is often used to reduce the dimensionality of data while maintaining the features that contribute the most in the data set in terms of the variance.

In this research, the dataset is denoted as a matrix $\mathrm{X}_{M\times N}$, where M is the total number of ears, and $N = 772$ is the total number of samples in each TEOAE signal. First, the covariance matrix $\mathbf{C}(\mathrm{X})_{N\times N}$ is calculated as follows:

$$\mathbf{C}(\mathrm{X}) = E[(X - E[X])(X - E[X])^T], \tag{2}$$

where the expectation is taken along the ear dimension. Then, the eigenvalue (λ) and the corresponding eigenvector (w) of are calculated by solving the following system of equations,

$$(\mathbf{C}(\mathrm{X}) - \lambda I)w = \mathbf{0}. \tag{3}$$

Finally, the eigenvalues are sorted in the descending order, and the eigenvector corresponding to the first $K = 3$ eigenvalues are selected as the eigenvector set $\mathrm{Y}_{NxK}$for projection. The features after dimension reduction, X'_{MxK}, is defined as follows,

$$X'_{M\times K} = X_{M\times N} \times Y_{N\times K}. \tag{4}$$

Each column in the matrix X'_{MxK} is considered a feature and the values are hereafter called *TEOAE-after-PCA*.

2.4.2 SNR

The signal-to-noise ratio (SNR) captures the relative strength of the TEOAE signal with the background noise. In this study, we attempt to evaluate its usage as a feature for machine prognosis purposes. The SNR is defined as follows,

$$\mathrm{SNR} = 10\log_{10}\left(\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{u}\mathbf{u}^T}\right), \tag{5}$$

where $\mathbf{v}$ and $\mathbf{u}$ are row-vector representations for the TEOAE signal and the noise-floor signal, respectively for each recording session, respectively.

2.4.3 Energy

Although the information might be somewhat redundant, to evaluate its usefulness for machine prognosis purposes, the absolute energy is also calculated,

$$E_s = \mathbf{v}\mathbf{v}^T. \tag{6}$$

2.4.4 Group delay

The group delay $\tau(\omega)$ is the energy delay time of the signals in different frequencies. It can be calculated as follows:

$$\tau(\omega) = -\frac{d}{d\omega}[\angle X(e^{j\omega})], \tag{7}$$

where $\angle X(e^{j\omega})$ denotes the phase spectrum of the TEOAE signal. An example of $\tau(\omega)$ of an ear that was not affected by MD is shown in Figure 5. Empirically, since MD is known to affect the low-frequency hearing (5) more than the high frequency hearing, we selected the group delay at 1 and 2 kHz as potentially useful features for machine prognosis.

In practice, the phase spectrum of a signal would be less reliable at frequencies with low SNR; for example, in the left panel of Figure 5, noise magnitude is close or even larger than signal magnitude at some points. To avoid large errors in subsequent calculation of the group delay, a point pick procedure was adopted. If any frequency has an SNR of less than 2 dB, that frequency bin would be abandoned. In order to get continuous phase, the lowess smoothing with length 150 points, about 3.2 kHz, was applied, and the phase curve after smoothing was shown in Figure 5 (right panel).

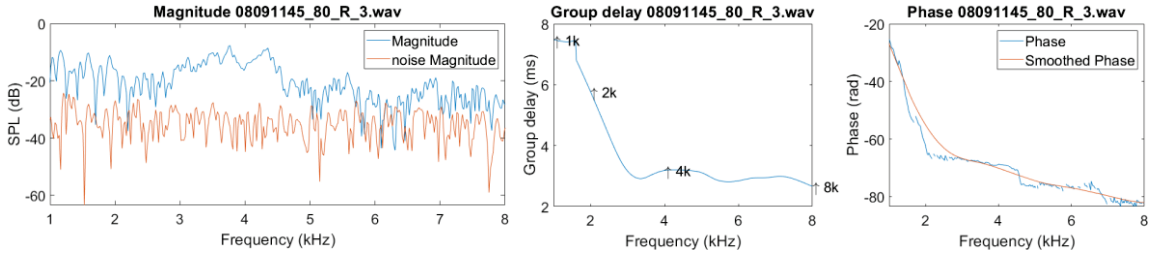

Figure 5 – An Example of the TEOAE magnitude spectrum (left), the corresponding group delay (middle) and the phase spectrum (right). Note that both the phase spectrum before and after smoothing are shown for comparison.

3. Machine learning results of prognosis

A support vector machine (SVM) (20) classifier was trained to predict whether an MD-affected ear would turn out to be a "good" case or a "bad" case (See Sec. 2.1 for the definition). The SVM classifier takes the features listed in Sec. 2 (e.g., TEOAE after PCA, SNR, group delay at 1 and 2 kHz) as the input, acquired during the patient's first visit to the clinic after the onset of initial MD symptoms. The machine learning methods and evaluations are given in 3.1, the features of TEOAE are visualized in 3.2, and some additional statistical analysis is given in 3.3.

3.1 Classification by SVM

First, the value of each feature is normalized so the range becomes between 0 (the minimum among all ears) and 1 (the maximum). This normalization step is necessary so the SVM training would not favor one feature over the other simply because of the choice of unit. The adopted SVM library is from Scikit-Learn (15), and the best penalty parameter and kernel coefficient were determined by cross validation.

Empirically, we chose the sigmoid function as the kernel function for SVM. The prognosis accuracies of different feature combinations with 3, 4, 5, or 16-fold cross validation are shown in Table 1. In particular, the confusion matrix of 4 cross-validation using TEOAE-after-PCA and group delay is shown in Table 2.

Observing the accuracy in Table 1, note that the combination of TEOAE-after-PCA and group delay produces the best results; the results do not differ by more than 10% between 3, 4, 5, and 16-fold cross validation.

Table 1 – Prognosis accuracy (%) based on different feature combination

Features	3 cross-validation	4 cross-validation	5 cross-validation	16 cross-validation
TEOAE-after-PCA	81.81	78.125	80.00	**81.25**
SNR	60.60	65.625	73.33	62.50
Energy	54.54	71.875	66.66	62.50
Group delay	75.75	78.125	76.66	78.13
TEOAE-after-PCA & Group delay	**84.84**	**90.625**	**86.66**	**81.25**

Table 2 – The confusion matrix of 4 cross-validations by combining group delay at two frequencies and TEOAE projections onto three principal components

M=32	Predicted "Good"	Predicted "Bad"
Actual: "Good"	13	3
Actual: "Bad"	0	16

3.2 PCA data visualization

After PCA, the dimension of TEOAE data is reduced from 772 to 3 and it can be visualized by 3-D plots as shown in Figure 6.

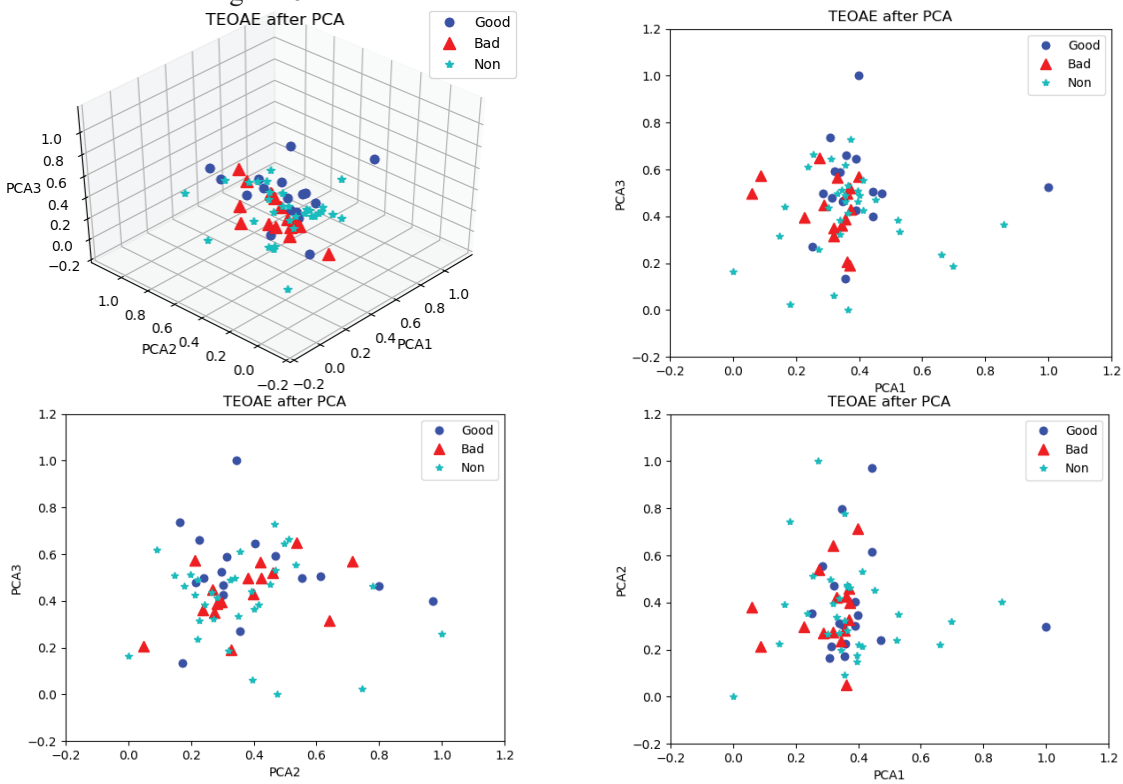

Figure 6 – (3 cases) TEOAE data reduced to three dimensions after PCA

The eigenvectors corresponding to three largest principal components are shown in Figure 7. By inspection, note that the first two components seem to capture the middle-ear ringing more than the third component, and the third component looks more reminiscent of an actual TEOAE signal. However, in Figure 6 it is hard to say whether the "good" and "bad" ears tend to segregate better along one of the PCA axes. Further analysis on the distribution of each feature is given next.

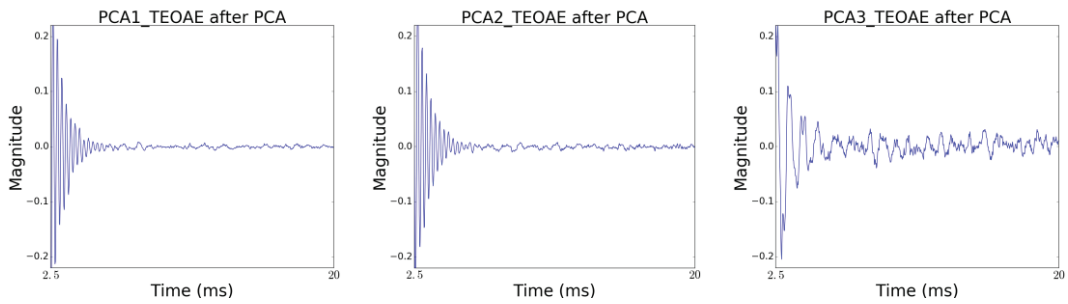

Figure 7 – (Left to Right) Eigenvectors of the largest three principal components

3.3 Statistical analysis for each feature

The difference in the statistical distribution between "good" and "bad" ears along each feature dimension is quantified by the Welch's t-test (18), and the results are shown in Table 3. Note that the

difference in group delay was the most significant (with the smallest p values). Also note that no significant difference was found in SNR, and energy features. This argues against the traditional wisdom of treating the magnitude of TEOAE per se as an indicator for inferring the cochlear status.

Table 3 – Statistical analysis by Welch's t-test based on different feature

Features	Test Statistic	p-value
TEOAE after PCA1	2.3089	0.0291
TEOAE after PCA2	-0.2732	0.7869
TEOAE after PCA3	1.7692	0.0926
SNR	-0.0926	0.9268
Energy	0.8160	0.4212
Group delay_1k	2.9409	0.0066
Group delay_2k	3.9127	0.0008

4. Discussion

This study focuses on finding useful features for MD prognosis purposes. By extracting features from TEOAE waveforms, we found that the combination of TEOAE-after-PCA and TEOAE group delay at 1 and 2 kHz give the best prediction accuracy. The present results suggest that TEOAE waveforms might contain useful information for inferring the cochlear conditions in MD affected ears.

For the first and second eigenvectors shown in Figure 7, the instantaneous frequency at the beginning seems to be higher than toward the end, which is consistent with the characteristics of TEOAE (7). In the future, if more data are available, more principal components could be visualized and perhaps useful information could be unraveled.

Presently, we encounter low SNR at 4 and 8 kHz due to relatively strong middle-ear ringing at the time course when high frequency (>4 kHz) components are supposed to be emitted from the cochlea. Therefore, though not shown in this article, the group delay calculated at 4 and 8 kHz turned out to be not helpful for SVM-based classification, because the group delay value was mostly close to 3 milliseconds in each case. To extract reliable signal at high frequencies, a nonlinear paradigm for signal acquisition might be considered. Nevertheless, since MD is known mostly to affect low frequency hearing rather than high frequency (5), chances are that 1 and 2 kHz components are still more reliable than higher frequencies for inferring cochlear conditions in MD-affected ears. For the prognosis of other acute hearing loss of cochlear origins, however, acquiring features from a broader frequency range might be preferred. In the future, frequency components below 1.0 kHz should also be analyzed if SNR could be improved and the power-line noise encountered in the present study can be physically blocked.

5. CONCLUSIONS

Useful features in TEOAE have been discovered for MD prognosis purposes. Combining TEOAE waveforms in 2.5~20ms and TEOAE group delay at 1 and 2 kHz gives the best prediction accuracy. Results of this study support our research motivation that the doctor could know if an MD patient's hearing had any chance to improve based upon examination during the first visit. The present results suggest that TEOAE signals provide abundant information for the prognosis of MD and the information could be extracted via machine-learning techniques.

ACKNOWLEDGEMENTS

This research is supported by the Ministry of Science and Technology, Taiwan, under Grant No. 107-2221-E-007-093-MY2.

REFERENCES

1. Li JC, Lorenzo N. Ménière disease (idiopathic endolymphatic hydrops). Medscape. 2016.
2. Amiridavan M, Nemati S, Hashemi SM, Jamshidi M, Saberi A, Asadi M. Otoacoustic emissions and

auditory brainstem responses in patients with sudden sensorineural hearing loss. Do otoacoustic emissions have prognostic value?. Journal of Research in Medical Sciences. 2006;11(4):263-269.
3. Carhart R, Jerger JF. Preferred method for clinical determination of pure-tone thresholds. Journal of Speech and Hearing Disorders. 1959;24(4):330-345.
4. Ciorba A, Skarżyński PH, Corazzi V, Bianchini C, Aimoni C, Hatzopoulos S. Assessment tools for use in patients with Ménière Disease: an update. Medical Science Monitor: International Medical Journal of Experimental and Clinical Research. 2017;23:6144-6149.
5. Dauman R, Aran JM, Portmann M. Summating potential and water balance in Meniere's disease. Annals of Otology, Rhinology & Laryngology. 1986;95(4):389-395.
6. Kaiser J, Schafer R. On the use of the I0-sinh window for spectrum analysis. IEEE Transactions on Acoustics, Speech, and Signal Processing. 1980;28(1):105-107.
7. Keefe DH, Feeney MP, Hunter LL, Fitzpatrick, DF. Comparisons of transient evoked otoacoustic emissions using chirp and click stimuli. The Journal of the Acoustical Society of America. 2016;140(3):1949-1973.
8. Kemp DT. Stimulated acoustic emissions from within the human auditory system. The Journal of the Acoustical Society of America. 1978;64(5):1386-1391.
9. Kuroda T. Clinical investigation on spontaneous otoacoustic emission (SOAE) in 447 ears. Auris Nasus Larynx. 2007;34(1):29-38.
10. Lalaki P, Markou K, Tsalighopoulos MG, Daniilidis I. Transiently evoked otoacoustic emissions as a prognostic indicator in idiopathic sudden hearing loss. Scandinavian Audiology. 2001;30(1):141-145.
11. Mertes IB, Goodman SS. Short-latency transient-evoked otoacoustic emissions as predictors of hearing status and thresholds. The Journal of the Acoustical Society of America. 2013;134(3):2127-2135.
12. Mom T, Montalban A, Bascoul A, Gilain L, Avan P. Acoustic phase shift: objective evidence for intralabyrinthine pressure disturbance in Meniere's disease provided by otoacoustic emissions. European Annals of Otorhinolaryngology, Head and Neck Diseases. 2012;129(1):17-21.
13. Nubel K, Kabudwand E, Scholz G, Mrowinski D. Diagnosis of endolymphatic hydrops with low tone masked otoacoustic emissions. Laryngo-Rhino-Otologie. 1995;74(11):651-656.
14. Pearson K. LIII. On lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science. 1901;2(11):559-572.
15. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B. Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É. Scikit-learn: machine learning in python. Journal of Machine Learning Research. 2011;12:2825-2830.
16. Prieve BA, Gorga MP, Schmidt A, Neely S, Peters J, Schultes L, Jesteadt W. Analysis of transient-evoked otoacoustic emissions in normal-hearing and hearing-impaired ears. The Journal of the Acoustical Society of America. 1993;93(6):3308-3319.
17. Shupak A, Zeidan R, Shemesh R. Otoacoustic emissions in the prediction of sudden sensorineural hearing loss outcome. Otology & Neurotology. 2014;35(10):1691-1697.
18. Welch BL. The generalization of student's problem when several different population variances are involved. Biometrika. 1947;34(1/2):28-35.
19. White KR, Vohr BR, Maxon AB, Behrens TR, McPherson MG, Mauk GW. Screening all newborns for hearing loss using transient evoked otoacoustic emissions. International Journal of Pediatric Otorhinolaryngology. 1994;29(3):203-217.
20. Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. In Proceedings of the 5th Annual Workshop on Computational Learning Theory; July 1992; Pittsburgh, PA, USA 1992. p. 144-152.
21. Johnstone IM. High dimensional statistical inference and random matrices. Proceedings of The International Congress of Mathematicians; August 2006; Madrid, Spain 2006, p. 307–333.
22. Liu TC, Wu HT, Chen YH, Chen YH, Fang TY, Wang PC, Liu YW. Analysis of click-evoked otoacoustic emissions by concentration of frequency and time: Preliminary results from normal hearing and Ménière's disease ears. AIP Conference Proceedings. 2018;1965(1):170005.
23. Ménière's Disease. National institutes on deafness and other communication disorders; NIH Publication No. 10–3404; July 2010; Bethesda, USA. From https://www.nidcd.nih.gov/health/menieres-disease.