International Journal of Audiology 2012; Early Online: 1–8

informa
healthcare

Original Article

Accuracy of distortion-product otoacoustic emissions-based ototoxicity monitoring using various primary frequency step-sizes

Garnett P. McMillan*, Dawn Konrad-Martin*,† & Marilyn F. Dille*,†

*VA RR&D National Center for Rehabilitative Auditory Research, Portland VA Medical Center, Portland, Oregon, USA
†Department of Otolaryngology/HNS, Oregon Health and Science University, Portland, Oregon, USA

Abstract

Objective: A cisplatin ototoxicity monitoring protocol was recently proposed using distortion-product otoacoustic emissions (DPOAEs) measured in 1/48th octave steps over the highest obtainable quarter octave (Dille et al, 2010). This protocol can take up to 40 minutes to complete in both ears among seriously ill patients in a potentially noisy test environment. The goal of the current study was to contrast the diagnostic accuracy of ototoxicity monitoring protocols based on changes in DPOAE levels at wider, more rapidly tested, primary frequency step sizes. *Design:* Measure DPOAE levels in 1/48th octave steps over the highest half-octave of obtainable DPOAEs prior to treatment and at each ototoxicity monitoring session during the course of treatment with cisplatin. *Study sample:* Nineteen cancer patients being treated with cisplatin at the Portland Veterans Affairs Medical Center were observed over 56 monitoring appointments. Hearing thresholds in the sensitive region for ototoxicity (SRO) were measured concurrently with DPOAE levels. *Results:* DPOAE levels measured in 1/24th octave steps provided comparable accuracy, and half the testing time, to the 1/48th octave step protocol previously described. *Conclusions:* DPOAE level shifts measured in 1/24th octave steps may provide a basis for rapid ototoxicity monitoring among adult cancer patients treated with cisplatin.

Key Words: Distortion-product otoacoustic emissions; ototoxicity monitoring; cisplatin; fine structure

Cisplatin is an important chemotherapeutic agent that is dose-limited by hearing loss. Ototoxicity monitoring is conducted to evaluate hearing changes during treatment. Protocols vary, but a simple hearing screen is usually performed during cisplatin administration or hydration, with a full audiometric exam following a poor screening result. In principle, this monitoring program is maintained for each patient at each chemotherapy visit during the course of treatment. Ototoxicity monitoring is challenging, however, because screening is ideally done on the hospital ward and patients are often too ill to complete a reliable behavioral test. A method that accurately screens for ototoxic hearing shifts but that does not require a responsive patient (i.e. is objective) is desirable as part of the armament of screening tests.

Monitoring changes in distortion-product otoacoustic emissions (DPOAEs) is an important clinical option for cisplatin ototoxicity screening. We recently described a method of cisplatin ototoxicity monitoring based on DPOAE levels collected in fine frequency steps (Dille et al, 2010). DPOAEs are low level sounds recorded in the ear canal in response to a set of closely spaced primaries, f1 and f2, where $f1 < f2$. There is increasing support for the hypothesis that two (or more) sources generate DPOAEs. The initial source of the 2f1-f2 DPOAE is a non-linear interaction that occurs in the region of overlap between the two primaries, somewhat closer to the f2 tonotopic peak (Brown & Kemp, 1984). Distortion at 2f1-f2 propagates from this region in two directions. It travels back toward the oval window where it is transmitted to the ear canal (Brown & Kemp, 1984; Martin et al, 1987) and propagates towards its own characteristic frequency place (Goldstein & Kiang, 1967; Kim, 1980). There it is reinforced locally by the cochlear amplifier and combines coherently with backward reflections from randomly spaced discontinuities located near the peak of the 2f1-f2 traveling wave, forming a stimulus-frequency otoacoustic emission (Shera & Zweig, 1993). The two components, with their attendant amplitude and phase responses combine in the ear canal, resulting in fluctuations in DPOAE level, termed DPOAE fine structure. DPOAE fine structure can lead to diagnostic inaccuracies, but this can be mitigated by collecting DPOAEs with small frequency step-sizes (in a fine-structure paradigm) and then smoothing across adjacent frequencies to obtain a DPOAE level measurement (Mauermann & Kollermeier, 2004). Depending on the amount of smoothing, the smoothed DPOAE response can allow for an easier interpretation of the DPOAE, since the averaged level can be thought of as a rough estimate of the initial DPOAE generator source (e.g. Wagner et al, 2008). Moreover, smoothing the DPOAE response reduces the influence of any spurious or missing data, detriments to any clinical DPOAE application.

Correspondence: Garnett P. McMillan, VA RR&D National Center for Rehabilitative Auditory Research, Portland VA Medical Center, 3710 SW US Veterans Hospital Road, Portland, OR 97239, USA. E-mail: garnett.mcmillan@va.gov

(Received 6 July 2011; accepted 9 April 2012)

ISSN 1499-2027 print/ISSN 1708-8186 online
DOI: 10.3109/14992027.2012.688143

Abbreviations

ASHA	American Speech-Language-Hearing Association
AUC	Area under the receiver operating characteristic curve
DPOAE	Distortion-product otoacoustic emissions
ORA	Ototoxicity risk assessment
PLS	Partial least squares
SRO	Sensitive range for ototoxicity

The ototoxicity screening method published by Dille et al (2010) measured DPOAEs using a fixed ratio, primary frequency sweep in 1/48th octave steps over the highest quarter octave of obtainable DPOAEs. In order to determine whether a recorded DPOAE was valid for analysis, DPOAE level measurements were compared to the corresponding noise level recorded in the ear canal and system distortion estimated as the DPOAE level at 2f1-f2 recorded in a standard 2cc cavity (Brüel & Kjær 4153 Coupler). The biological noise was converted to intensity, added to the corresponding mean coupler intensity level and the combined noise and distortion value was transformed to dB SPL. A signal to noise ratio was then defined as the observed DPOAE level in dB SPL minus the back transformed sum of the subject noise and system distortion in dB SPL. For a given stimulus condition, a DPOAE ear canal response was considered valid and present if the SNR was at least 6 dB. If the SNR was less than 6 dB and if the subject noise was less than or equal to mean system distortion plus 2 standard deviations, the DPOAE measure was still considered valid, i.e. the low level emission was considered present, interpretable, and the measurement value was used in the analyses. If the SNR was less than 6 dB and the subject noise was greater than the mean system distortion plus 2 standard deviations, the DPOAE measure was set to missing. DPOAE levels at adjacent frequencies were smoothed using a five-point running average to control variability due to any fine structure or spurious measurements, and to impute DPOAE levels set to missing due to high subject noise levels.

This method, denoted the "ototoxicity risk assessment" (ORA), was reasonably accurate for identifying cisplatin-induced hearing shifts. However, an ototoxicity monitoring test must be rapid, particularly if it is to be used in a population of older, seriously ill patients. DPOAE level shifts measured at larger octave steps may offer comparable accuracy in identifying hearing shifts with considerable time savings.

There are clinical, theoretical, and statistical motives to pursue this possibility. The goal of a screening method is to maximize accuracy while minimizing costs in terms of patient time, discomfort, training, and equipment outlay. The ORA method, which is based on a 1/48th octave primary frequency sweep over the highest quarter octave of obtainable DPOAEs, was as accurate as methods using one half-octave of obtainable DPOAEs while requiring half the patient and clinician time commitment. Even so, the ORA is still time consuming, requiring 12 DPOAE measurements taken on each ear at each monitoring appointment. This can take up to 20 minutes per ear to complete in a noisy hospital ward among older, noisy patients. This motivates searching for a faster test using wider step sizes. Obviously, at some point increasing the step size will penalize the accuracy of the test. Considering the problem as a simple sampling issue, fewer measurements taken at wider primary frequency steps may provide less accurate information about the underlying level data, potentially obscuring DPOAE level shifts that would otherwise indicate cochlear damage.

Finer step measurement allows more locally accurate smoothing over stochastically volatile data and allows imputation of missing data due to high subject noise. Furthermore, finer step measurement permits more locally accurate smoothing over any fine structure, which, left unchecked, potentially degrades accuracy. DPOAE fine structure shows a roughly periodic spacing of about 1/10th octave (e.g. He & Schmiedt, 1993). Following loosely the Nyquist sampling theorem, any DPOAE fine structure needs to be sampled using at least twice this resolution to be accurately represented. Frequency step sizes of the DPOAE measurements close to the typical spacing of DPOAE fine structure can lead to well-known "aliasing" effects. In this case, the obtained local average would differ substantially depending on the choice of the test frequencies relative to dips and valleys in the fine structure. A more reliable local average measurement that still accomplishes the goal of fine structure smoothing, can be obtained for somewhat higher frequency step sizes than the Nyquist frequency.

Finely spaced measurements give a more locally accurate description of the DPOAEs, but DPOAE levels measured at adjacent primary frequencies separated by 1/48th octave are also likely to be highly correlated. This leads one to question whether the sum of the information provided by fine scale measurement significantly exceeds measurements taken at larger, uncorrelated primary frequency step sizes. Furthermore, superfluous measurements at primary frequencies that are largely unaffected by cisplatin will introduce statistical noise into the screening protocol and degrade accuracy.

Ultimately, the decision to use a fine step protocol over a wider primary step protocol will depend on the relative accuracy, which must be empirically determined in exposed ears, as well as time costs. In this paper we contrast the fine step ORA method with alternatives based on larger primary frequency step sizes. Can the ORA method that relies on OAE fine step spacing be replaced by a faster, simpler OAE screening method that uses larger primary frequency step sizes?

Methods

The sample, measurement protocols, and statistical methods are identical to those described previously (Dille et al, 2010). Patients' hearing was tested prior to the start of chemotherapy using test frequencies from 2–20 kHz in order to define a behavioral sensitive range for ototoxicity (SRO). A behavioral high-frequency hearing limit was established as the highest frequency at which a threshold could be obtained using a pure-tone signal of 100 dB SPL or less. Pure-tone thresholds of the six lower adjacent frequencies in 1/6th octave steps plus the high frequency limit constituted the SRO frequencies and the behavioral hearing test employed at each subsequent cisplatin treatment session. Clinically significant hearing shifts were defined using American Speech-Language-Hearing Association (ASHA) criteria in these frequencies (ASHA, 1994) and include: (1) $\geq$ 20 dB change at any one test frequency; (2) $\geq$ 10 dB change at any two consecutive test frequencies; or (3) loss of response at three consecutive test frequencies where responses were previously obtained. Using these criteria, a binary gold standard indicator for presence or absence of hearing change was constructed for each ear at each monitoring appointment. The ototoxicity screening objective is to identify this gold standard indicator at each monitoring session using DPOAE level shifts.

At the same time that behavioral SRO data were obtained, fine resolution DPOAEs were measured near the DPOAE high frequency limit established at baseline. Just as the behavioral SRO test

frequency range is operationally defined based on the highest audible frequency that can be detected, a DPOAE sensitive range is defined using the highest obtainable DPOAE as an upper limit. All DPOAE recordings were obtained using a fixed primary frequency ratio f2/f1 = 1.22 and levels of the f1 and f2 primaries set to L1 = L2 = 65 dB SPL. In order to establish the upper limit, an initial "DP-gram" was constructed at baseline with f2 swept from 1–14 kHz in ½-octave steps. "fine resolution DP-grams" were then measured as a function of f2 over the half-octave range below the upper limit, with frequencies separated by 1/48th octave. To be considered valid for analysis, a DPOAE had to have (1) a + 6 dB signal to noise ratio where noise was the combined subject noise and system distortion at the corresponding frequency/level condition, or (2) low noise such that the noise floor was within two standard deviations of the mean system distortion.

DPOAEs were collected using custom software (Otoacoustic Emission Averager, EMAV; Boys Town National Research Hospital (Neely & Liu, 1993)) run on a personal computer. The software used a CardDeluxe digital signal processing board (Digital Audio Laboratories) to generate stimuli and record responses. Stimuli were outputted through separated channels of the CardDeluxe, passed through a zero gain custom buffer amplifier to two earphones (Etymotic Research, ER-2) and delivered to the sealed ear canal. The ear canal pressure was sampled at a rate of 32 kHz, amplified 20 dB by the ER-10B + pre-amplifier, digitized in 64-ms time windows, and stored in two interleaved buffers, A and B, each averaged in the time domain. DPOAE level at 2f1-f2 was estimated from a fast Fourier transform of the grand average of the two response buffers ([A + B]/2) and the noise level was estimated at the DPOAE frequency from the A-B spectrum. System distortion was below − 20 dB in a Bruel & Kjaer 4157 coupler for the stimulus conditions used in this study.

Let *k* denote a primary frequency step spacing measured in k/48th octave steps. In this paper, we contrast the accuracy of DPOAE fine step protocols that use k/48th octave primary frequency step spacing, k = 1, 2, …, 6, over the highest half- and quarter-octave of obtainable DPOAEs. The goal is to identify the interval-step size combination that most accurately screens for the gold standard behavioral hearing shifts, where 'interval' pertains to the octave width and step size pertains to *k*, as defined above.

Frequency steps considered in this analysis were normalized to the highest obtainable DPOAE as shown in Table 1. Step 1 is the highest primary frequency with an obtainable DPOAE, Step 2 is 1/48th octave below step 1, and so forth. A total of twelve interval-step size combinations are contrasted in this analysis corresponding to primary frequency step sizes, k = 1, 2, …, 6, over two interval widths covering one half or one quarter octaves. Steps included in each interval-step size combination are indicated by an 'x' in Table 1. The ORA protocol is identified by the shaded region in Table 1.

DPOAE levels measured at each session were smoothed using a five-point running average for interval-step size combinations with five or more total steps within the interval width. DPOAEs from interval-step size combinations with fewer than five total steps were smoothed using a three-point moving average. Smoothing was only done over DPOAE levels contributing to the interval-step size combination under consideration. For example, DPOAE levels measured in 4/48th octave steps over the highest quarter octave of obtainable DPOAEs were smoothed using only DPOAE levels measured at those steps and in that interval as indicated in Table 1.

Examples of four sets of DPOAE levels measured at 1/48th octave steps and smoothed over each candidate step size are shown in Figure 1 for four different subjects. Smoothed results (lines) for each step size over the highest half-octave of obtainable DPOAEs are overlayed on the observed DPOAE levels (dots). Smaller step sizes yield smoothed curves that more closely adhere to the underlying fine structure, while smoothing over wider steps yields smoothed levels that are less susceptible to fine structure. Whether or not it is beneficial to collect DPOAEs in fine frequency steps, and to what degree it is most beneficial to do so, depends on the accuracy with which the DPOAE level data identifies behavioral hearing test results.

For any particular patient, DPOAE level shift $\Delta OAE_{s,m}$ at step *s* during monitoring session *m* was defined as $\Delta OAE_{s,m} = OAE_{s,baseline} - OAE_{s,m}$, where OAE_s corresponds to the smoothed DPOAE level at step *s* according to the smoothing methods described above. The $\Delta OAE_{s,m}$ were used as inputs into a partial least squares (PLS) model predicting the gold standard SRO hearing shifts. The PLS model identified three orthogonal, linear combinations of DPOAE level shifts that were then fed into a logistic regression model of the gold standard hearing shift indicator. The logit model also included a dose-ototoxicity factor expressing the underlying risk of a hearing shift as a result of cumulative exposure to cisplatin and pre-treatment hearing levels (Dille et al, 2012). The output from the model is a **risk score** for each ear measured during each ototoxicity monitoring session. The risk score is a linear combination of the $\Delta OAE_{s,m}$ and dose-ototoxicity factor that best predicts the chances that an ear at that monitoring occasion has suffered a hearing shift. The accuracy of each interval-step size combination was measured by the area under the receiver operating characteristic curve (AUC). The standard error of the AUC was estimated with Obuchowski's non-parametric estimator for clustered measurements (Obuchowski, 1997), where clusters correspond to each patient's ear. Each model was trained and tested using leave-one-out cross-validation. The goal is to identify the interval-step size combination requiring the fewest measurements that is at least as accurate as the ORA. This is important evidence for the design of DPOAE-based cisplatin ototoxicity monitoring protocols, with the aim of maximizing drug efficacy while reducing the damage sustained by the auditory end organ, both in responsive and non-responsive patients.

Results

Nineteen cancer patients treated with cisplatin at the Portland Veteran's Affairs Medical Center provided fifty-six ototoxicity monitoring visits. Twenty-three of the 56 monitoring visits (41.1%) showed cisplatin-induced ototoxicity according to ASHA criteria. This analysis evaluates the accuracy with which each interval-step size combination in Table 1 predicts hearing shifts at each of these 56 monitoring appointments.

The possibility of redundant DPOAE level information is confirmed in Figure 2, which plots the correlation between DPOAE levels at pairs of primary frequencies that are k/48th (k = 1, 2, …, 25) octaves apart. The left panel shows results for the unsmoothed DPOAE levels, while the right panel shows results for the smoothed DPOAE levels. Each symbol corresponds to one pair of steps. The left-most position on each graph shows the correlation between all steps that are 1/48th octave apart (e.g. steps 1 and 2 or steps 3 and 4 or steps 19 and 20, etc.), while the right-most region shows the correlation between pairs of steps that approach a half octave apart. A high correlation indicates redundancy in the DPOAE level measurements, which might not necessarily increase accuracy but still induces a clinical cost of time and discomfort. Conversely, correlations near zero indicate a low level of redundancy so that relatively

Table 1. Step size protocols considered in this analysis. Primary frequencies are 'normalized' so that step 1 is the highest frequency with obtainable DPOAEs. 'x' indicates a step used in the interval-step size combination. The shaded region corresponds to DPOAE primary frequency steps used in the ORA.

	Step s																							
	Highest half octave																							
	Highest quarter octave																							
k (1/48th octave steps)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
2	x		x		x		x		x		x		x		x		x		x		x		x	
3	x			x			x			x			x			x			x			x		
4	x				x				x				x				x				x			
5	x					x					x					x					x			
6	x						x						x						x					

more information is gathered in a comparable amount of time. As suspected, the correlation functions indicate that DPOAE levels are highly correlated when the primary frequencies are close together, and the correlation decays the further apart the primary frequencies. The smoothed results are obviously much tighter, since they are based on DPOAE levels that are subjected to five-point running averaging prior to computing the correlations. The suspicion, therefore, is that one might get just as much clinically useful information from measuring DPOAE levels at primary frequencies that are more widely separated than 1/48th octaves.

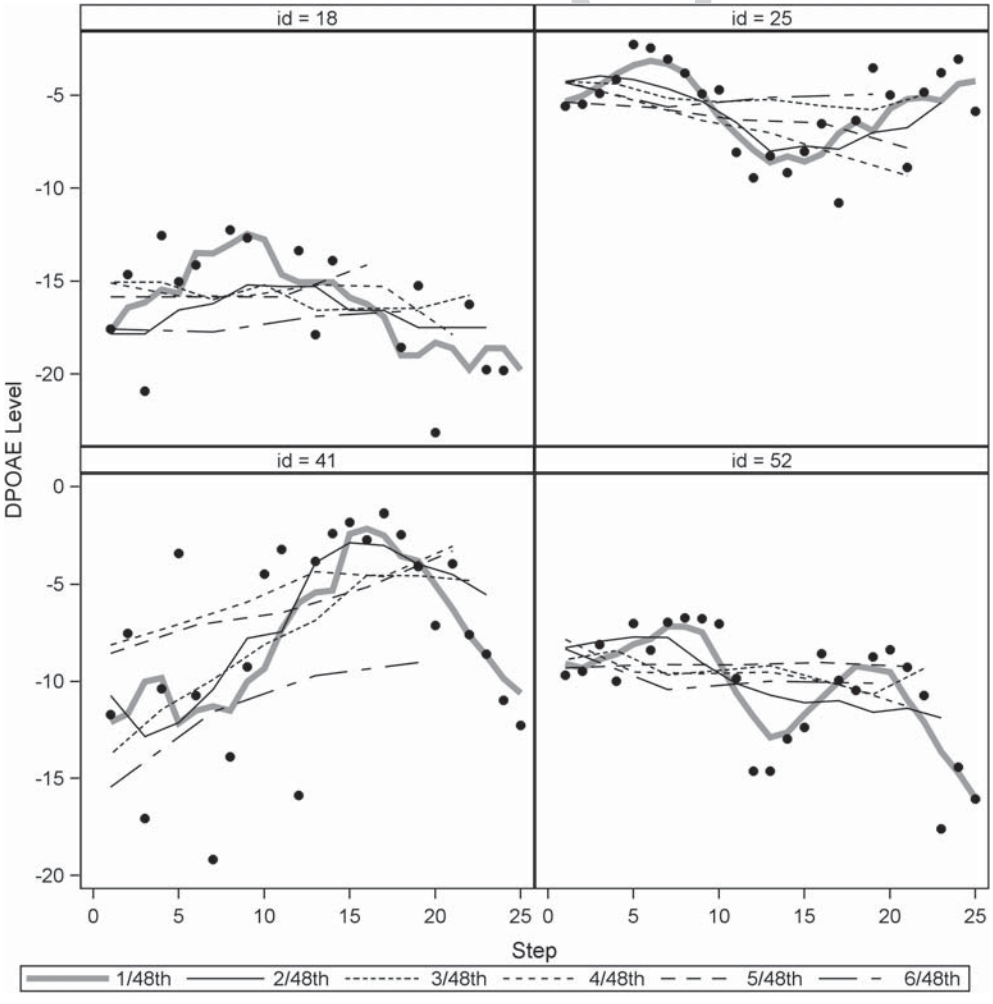

Figure 1. DPOAE levels observed (dots) and smoothed (lines) over one-half octave during four different subjects' ototoxicity monitoring sessions. Steps are ordered from highest (step = 1) to lowest (step = 25) frequency in the half-octave interval.

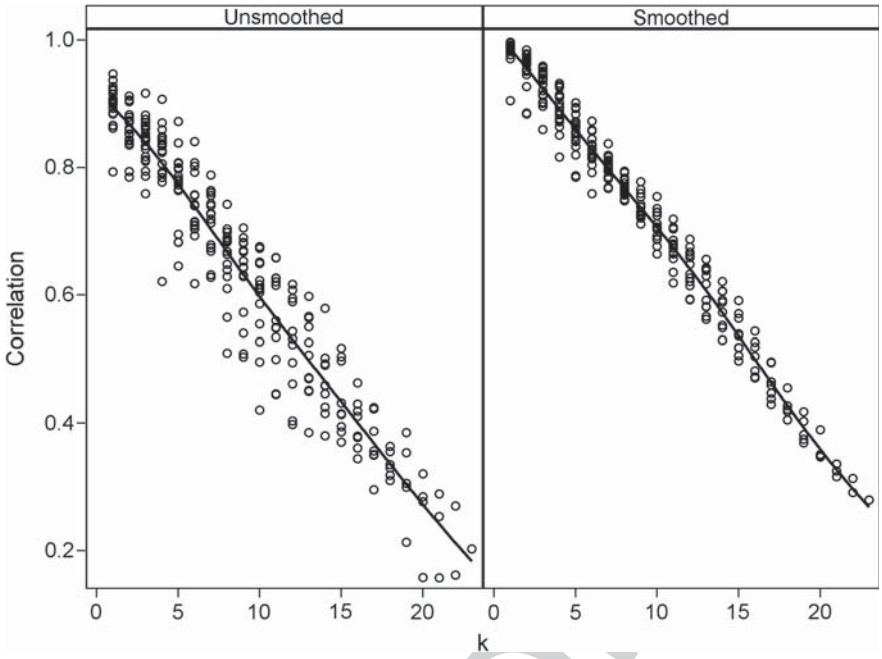

Figure 2. Correlations between DPOAE levels at pairs of primary frequencies separated by k/48th octaves demarcated on the x-axis. Each point on the graphs indicates a pair of primary frequencies. k = 1 denotes pairs of frequencies separated by 1/48th octave, k = 12 denotes a quarter octave separation, and k = 24 denotes one half octave separation. The left panel is based on the unsmoothed results and the right panel shows correlations after five-point moving averaging. The solid line is a fitted loess curve.

Cross-validated AUCs ± one standard error are shown in Figure 3 for each interval-step size combination. The ORA method (circled) is the most accurate method in the quarter-octave interval (left panel), as is the 1/48th octave step size method in the half-octave interval (right panel). Accuracy decreases steadily with increasing step size up to 4/48th (or 1/12th) octave step sizes. Thus, despite the redundancy suggested in Figure 2, smaller step sizes offer more useful information about cochlear damage, as well as perhaps more accurate smoothing over fine structure for detecting ototoxic hearing loss.

A valid contender for the ORA must perform better than the minimum performance threshold used to select the ORA in Dille et al (2010). This threshold was based on the so-called 'One standard error rule' and corresponded to the AUC that was one standard error

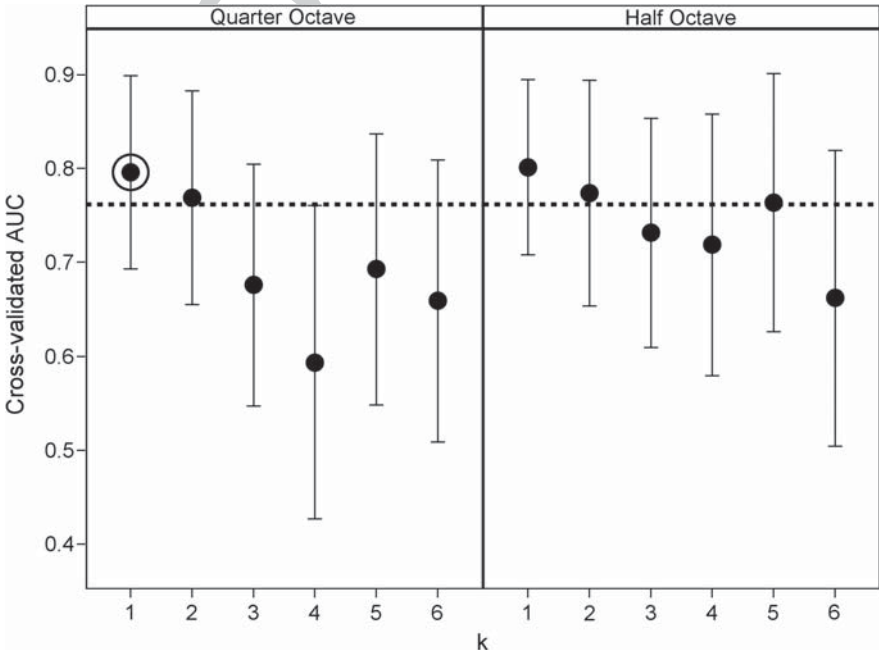

Figure 3. Cross-validated AUC estimates for each interval-step size combination considered. The dashed line indicates the minimum acceptable AUC according the 'one standard error' criteria described in Dille et al (2010). The circled point is the ORA.

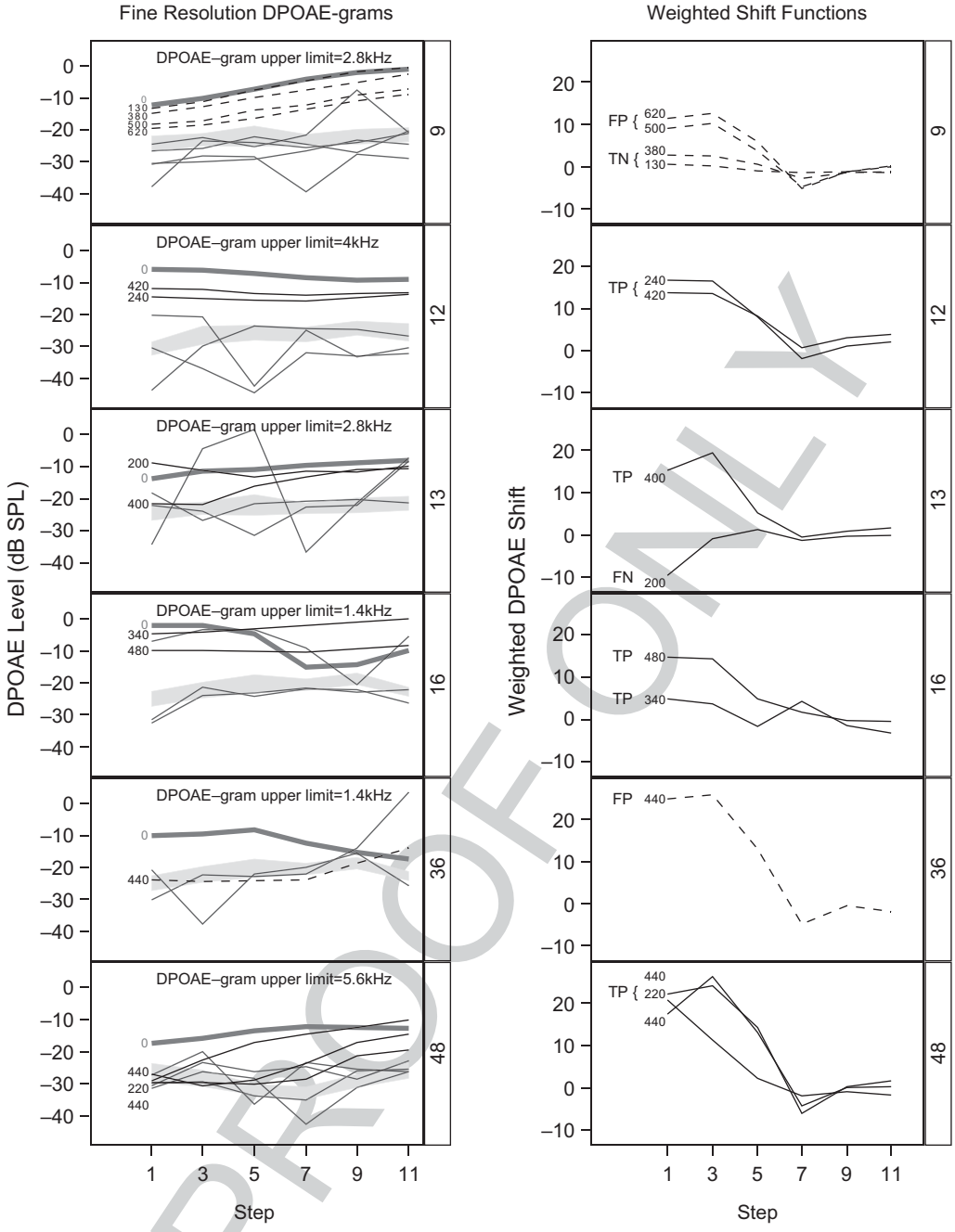

Figure 4. Observed fine resolution DPOAE-grams (left column) and weighted shift functions (right column) at each 48th octave step for each monitoring appointment among six subjects included in the analysis. Each line in the left column corresponds to DPOAE levels or noise (dotted lines). The thick line in the left column is the baseline DPOAE-gram and the shaded region is system distortion. Each line in the right column corresponds to the weighted DPOAE shift functions based on the multivariate 1/24th octave step model selected in Solid lines are monitoring visits for which a hearing shift was observed, and dashed lines indicated no observed hearing shift. Numeric values in each plot indicate the cumulative dose of cisplatin as of that monitoring appointment. Line labels in the right column identify the accuracy of the proposed model: TP = True positive, TN = True negative, FP = False positive, FN = False negative.

below the most accurate, and also most statistically complex, discrimination method. This method, not shown here, was based on a six-component PLS model over one-half octave of 1/48th octave step measurements, which, for statistical reasons, is unlikely to generalize to other populations of cisplatin patients (see Dille et al, 2010 and Hastie et al, 2009 for further details). The dashed line in Figure 3 demarcates the one-standard error threshold, and shows that risk assessments based on 2/48th octave steps measured over one-quarter octave is a valid competitor for the ORA since its AUC lies above the threshold limit. One can therefore achieve roughly the same accuracy as the ORA with a screening protocol that measures DPOAE level shifts in 2/48th octave steps over the highest quarter octave of obtainable DPOAEs. This translates to half the measurement time requirement of the ORA (see Table 1), which is a clinically significant

advantage. Wider step sizes are below the acceptance threshold, and are not further considered for an objective monitoring protocol.

The transformation from monitored DPOAE levels to weighted DPOAE shift functions ($\Delta OAE_{s,m}$) and predictions about behavioral shifts are illustrated in Figure 4. Fine resolution DP-grams are shown in the left column of Figure 4 for six representative subjects (with subject number indicated to the right of each panel). The DP-grams are smoothed functions composed of data taken at 2/48th octave steps over the highest quarter octave of obtainable DPOAEs. Recall that both behavioral SRO and fine resolution DP-gram test frequencies were tailored to each subject according to their auditory function prior to chemotherapy, with the assumption that monitoring is most productive for the highest frequencies that yielded a response at baseline. DPOAE test frequency (f2) is therefore plotted on the x-axis after normalizing to the highest recordable DPOAE frequency, which is indicated in each panel. As in Figure 1, test frequency decreases going from left to right on the x-axis. The baseline DP-gram for each subject is indicated by a thick solid line. DPOAEs collected at monitoring visits are shown by thin solid or dashed lines. Thin solid lines represent visits at which clinically significant hearing shifts were identified using a behavioral hearing test as the gold standard measure of the shift; dashed lines represent visits at which hearing remained stable relative to baseline. Numbers in each panel designate the cumulative cisplatin dose in mg at each visit. Subject noise (dotted lines) as well as system distortion (shaded region) are also included in each plot.

The fine resolution DP-grams shown in Figure 4 illustrate a number of important points. First, they show the advantage of making closely spaced measurements and employing a smoothing algorithm prior to analysing DPOAE data. In particular, some of the curves obtained for subjects 13, 16, and 36 would otherwise have been missing data due to noisy measurements. An examination of the baseline DP-grams reveals that perhaps two of the subjects shown (subjects 16 and 36) show level variations over a frequency span (about 1/10th octave) that is consistent with amplitude fine structure. Fine structure is absent for the other four baseline DP-grams depicted. An examination of the DP-grams associated with monitor visits illustrates some of the challenges of using a DPOAE-based approach to screen for ototoxic hearing shifts in patients who are not able to provide reliable hearing threshold data due to illness. Whereas DP-grams obtained at monitor visits sometimes shift overall downward when hearing also shifts significantly (subject 12, for example), they sometimes shift toward lower DPOAE level values while hearing remains stable (subjects 9 and 36). An even more vexing result is DPOAE levels that fail to decrease when hearing shifts significantly (subject 13, monitor associated with a cumulative cisplatin dose of 200 mg). Along similar lines, for ears with pronounced fine structure, DPOAE levels in fine structure dips frequently *increase* in level, and this can be associated with a hearing change (subject 16).

Application of the multivariate modeling and behavioral test prediction is also illustrated in Figure 4. The right column of the figure shows weighted DPOAE level shifts calculated by multiplying the $\Delta OAE_{s,m}$ by the PLS and logistic regression weighting functions developed for the 1/24th octave step method. The weighted shifts shown in Figure 4 can be summed across test frequency at each monitoring appointment to get a risk score from which behavioral hearing shift predictions are made. The risk score is then compared to a criterion cut-off value to determine whether the DP-gram shift predicts a hearing shift or not. Because both behavioral hearing and objective DPOAE measures were obtained for all subjects at each appointment, it is possible to determine the accuracy of the model-based results using virtually any criterion cut-off values. The general format of the right column in Figure 4 is the same as that of the left column. Predictions for each shift function are based on a criterion cut-off of 0, so that monitoring appointments generating risk scores greater than 0 are expected to show hearing shifts, while scores below zero are predicted to have no hearing shifts. The accuracy of the model is indicated for each of the weighted shift functions in the right column of Figure 4: True positives (TP) are monitoring appointments for which the 1/24th octave step model, in conjunction with the dose-ototoxicity component, correctly predicted a hearing shift. True negatives (TN) are appointments for which no shift was correctly predicted. False positives (FP) are monitoring appointments for which a hearing shift was expected to occur, but did not in actuality materialize. False negatives (FN) are appointments for which an actual hearing shift was incorrectly rejected by the model.

Model performance varied depending on the level shifts and dose at the time of monitoring. Actual hearing shifts at all monitoring appointments for subjects 12, 16, and 48 were correctly identified (true positives), while subject 36 was incorrectly expected to show a shift at that appointment (false positive). Subjects 9 and 13 gave mixed results, with hearing shifts incorrectly expected at the later monitoring appointments (cumulative doses of 500 and 620 mg) due to the observed drop in DPOAE levels. Likewise, subject 13 was not expected to show a hearing shift at 200 mg of cisplatin since little DPOAE shift occurred. Interestingly, the model predicted hearing shifts even for the monitoring appointment DP-gram data for subject 16 that showed an *increase* in DPOAE level at the lower F2. This is because the model was developed using data that sometimes showed clear fine structure at baseline. Fine structure was found to sometimes decrease (i.e. the curve flattened out) following cisplatin exposure, which is consistent with recent findings by Rao and colleagues that fine structure was found to decrease following consumption of high doses of aspirin (Rao et al, 2011). According to these authors the aspirin effectively unmixed DPOAE sources by preferentially acting on reflection source components of the DPOAEs (Rao et al, 2011).

Discussion

The problem addressed in the current report is to devise an objective method incorporating DPOAE level that is detailed enough to capture small changes in cochlear function, yet unencumbered by time consuming, redundant measurement. Closely spaced DPOAE levels are highly correlated (Figure 2), which motivates against a fine step protocol for ototoxicity monitoring. Conversely, more measurements, even apparently highly correlated ones, appear to offer improvement in accuracy either because they provide more information about cochlear damage or offer greater local accuracy when smoothing over fine structure. Interestingly, there is a jump in performance at the 5/48th octave step size (Figure 3). This may be a real phenomenon, a statistical artifact, or may be a byproduct of missing data due to noise but is worthy of further investigation. Because this 5/48th octave step size is roughly the typical period of DPOAE fine structure, we reviewed the fine structure prevalence in the sample according to criteria outlined in Wagner et al (2008). Consistent with our earlier impressions, evidence of fine structure was lacking in most subjects. This is not surprising given the relatively poor hearing among subjects in our study. Mean 4 kHz threshold was about 40 dBHL in our sample, of whom less than 5% are expected to show fine structure according to the estimates given by Wagner et al (2008). Our recommendation of a step size no wider than 1/24th octave is empirically determined in a sample of cisplatin-exposed

clinical patients. This evidence is necessary for the development of accurate, efficient DPOAE-based ototoxicity monitoring protocols.

Patient contact time is always an important consideration in clinical practice. Ototoxicity monitoring with DPOAEs entails about five minutes per ear of instrumentation setup using a protocol in which the audiologist attempts to match the probe fit to the fit achieved at baseline by comparing ear canal transfer functions in each ear between the two visits. During 123 cisplatin ototoxicity test sessions conducted in the Portland Veterans Affairs Medical Center, the 1/48th octave step measurement protocol took on average 3.2 minutes and up to 18 minutes per ear in noisy patients and/or in the presence of higher ambient room noise. The difference in recording times results from our use of measurement-based stopping rules, which permit shorter averaging times when the recorded noise floor is low. The 1/24th octave step measurement protocol took on average 1.7 minutes per ear to complete and up to 10 minutes for noisier recordings. The wider step protocol thus constitutes about 50% time savings on average, and in the noisiest testing situation. These translate to on average about 3 minutes less patient contact time and up to 16 minutes less patient contact time depending on the noise level.

The clinical goal is to have an automated DPOAE-based ototoxicity monitoring protocol that can be used for responsive and non-responsive patients. Prior to the beginning of the treatment course, patients would be given an audiologic exam which includes otoscopy, tympanometry, air conduction pure-tone threshold testing (2–20 kHz), and fine resolution DPOAE level measurements. The DPOAE protocol is applied at each monitoring appointment during the course of treatment, and a risk score based on DPOAE level shifts and dose-ototoxicity criteria is computed (Dille et al, 2010). No further action is taken if, during subsequent measurements, the risk score is within acceptable limits. However, if the risk score exceeds a pre-determined critical threshold, the patient is recommended for a repeat audiological exam. The critical threshold signaling the necessity for a follow-up recommendation depends on the risks of false positives or false negatives one is willing to accept, which is usually established in collaboration with oncology. Communications with the patient and health-care team should include whether the follow up audiogram established a hearing shift at frequencies that often cause speech understanding problems. This finding would lead to rehabilitative solutions during treatment to optimize communication between the patient and his/her family and medical staff. It could also lead to changes in the chemotherapy regimen.

We wish to emphasize that the DPOAE protocols described here were developed in a population of older, mostly male, and seriously ill patients all of whom had at least a mild degree of pre-treatment hearing loss. A younger patient population with better hearing might be most efficiently monitored using a different protocol. On the other hand, the prevalence of fine structure is known to be greatest among young individuals with good hearing (Wagner et al, 2008). A model-based DPOAE method that accurately screens for ototoxic hearing shifts while accounting for changes in fine structure may prove useful for the testing of pediatric patients undergoing chemotherapy. This serves to highlight the importance of diagnostic test development and validation within the patient population for which the monitoring protocol is intended.

As always, these and all other objective ototoxicity screening methods must be validated in a large, independent sample before clinical guidelines can be firmly established. Current efforts, including validating this fine step method, input/output function methods, and development of DPOAE phase change algorithms are underway at the National Center of Rehabilitative Auditory Research of the Portland Veterans Affairs Medical Center.

Conclusion

DPOAE level shifts measured in 1/24th octave steps may provide a basis for rapid ototoxicity monitoring among adult cancer patients treated with cisplatin.

Acknowledgements

Work supported by the Department of Veterans Affairs RR&D Service (grants C4183R and C7113N) and the VA RR&D National Center for Rehabilitative Auditory Research, Portland, Oregon, USA. Helpful comments were provided by an anonymous reviewer to an early draft of this paper.

Declaration of interests: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

American Speech-Language-Hearing Association. 1994. Guidelines for the audiologic management of individuals receiving cochleotoxic drug therapy. *ASHA*, 36, 11–19.

Brown A.M., Kemp D.T. 1984. Suppressibility of the 2f1-f2 stimulated acoustic emissions in gerbil and man. *Hear Res*, 13, 29–37.

Dille M.F., McMillan G.P., Reavis K.M., Jacobs P., Fausti S.A. et al. 2010. Ototoxicity risk assessment (ORA) combining distortion product otoacoustic emissions (DPOAE) with a cisplatin dose model. *J Acoust Soc Am*, 128, 1163–1174.

Dille M.F., McMillan G.P., Wilmington D., Fausti S.A. & Konrad-Martin D. 2012. (In press). Development and validation of a cisplatin dose-ototoxicity model. *J Am Acad Audiol.* 23.

Goldstein J.L. & Kiang N-Y. S. 1967. Neural correlates of the aural combination tone 2f1-f2. Proc. IEEE, 1968; 56, 981–992.

Hastie T., Tibshirani R. & Friedman J. 2009. *The Elements of Statistical Learning*, 2nd ed. Springer Series in Statistics. New York.

He N.J. & Schmiedt R.A. 1993. Fine structure of the 2f1-f2 acoustic distortion product: Changes with primary level. *J Acoust Soc Am*, 94, 2659–2669.

Kim D.O. 1980. Cochlear mechanics: Implications of electrophysiological and acoustical observations. *Hear Res*, 2, 297–317.

Martin G.K., Lonsbury-Marti B.L., Probst R., Scheinin S.A. & Coats A.C. 1987. Acoustic distortion products in rabbit ear canal. II. Sites of origin revealed by suppression contours and pure tone exposures. *Hearing Research*, 28, 191–208.

Mauermann M. & Kollmeier B. 2004. Distortion product otoacoustic emission (DPOAE) input/output functions and the influence of the second DPOAE source. *J Acoust Soc Am*, 116(4 Pt 1), 2199–212.

Neely S.T. & Liu Z. 1993. EMAV: Otoacoustic emission averager. Tech Memo No 17 (Boys Town National Research Hospital, Omaha).

Obuchowski N.A. 1997. Nonparametric analysis of clustered ROC curve data. *Biometrics*, 53, 567–78.

Rao A. & Long G.R. 2011. Effects of aspirin on distortion product fine structure: Interpreted by the two-source model for distortion product otoacoustic emissions generation. *J Acoust Soc Am*, 129, 792–800.

Shera C.A. & Zweig G. 1993. Noninvasive measurement of the cochlear traveling-wave ratio. *J Acoust Soc Am*, 93, 3333–3352.

Wagner W., Plinkert P.K., Vonthein R. & Plontke S.K. 2008. Fine structure of distortion product otoacoustic emissions: Its dependence on age and hearing threshold and clinical implications. *Eur. Arch. Otorhinolaryngol*, 265, 1165–1172.