

**LOUDNESS DISCOMFORT LEVELS: A
RETROSPECTIVE STUDY COMPARING
DATA FROM PASCOE (1988) AND
WASHINGTON UNIVERSITY SCHOOL OF
MEDICINE**

by

Jennifer N. Keller

**A Capstone Project
submitted in partial fulfillment of the
requirements for the degree of:**

Doctor of Audiology

**Washington University School of Medicine
Program in Audiology and Communication Sciences**

May 19, 2006

**Approved by:
Michael Valente, Ph.D., Capstone Project Advisor**

Abstract: Loudness discomfort levels (LDLs) were gathered from three Washington University School of Medicine sites, for a total of 325 subjects (total ears = 454). These levels were compared to mean LDLs reported by Pascoe (1988). The results revealed that the mean LDL measured at WUSM (ie., the IHAF procedure) is significantly different than the LDL reported by Pascoe (1988).

INTRODUCTION

History

Currently, approximately 28 million people living throughout America have a hearing impairment. This steadily increasing prevalence is a common problem in our modern society due to the combined effects of noise, aging, disease, and heredity. No age, ethnic group, or socio-economical class is immune to the consequences of these life-altering conditions. For every 1,000 children under the age of 18 years, 17 already have a hearing loss, and the rate of incidence increases with age. By the age of 75 years, 40 to 50% of the populace will report a hearing loss (NIDCD, 2001). Hearing loss can vary by type (ie., conductive, sensorineural, mixed, or retrocochlear) and degree (ie., mild to profound). The most prevalent type of hearing loss is sensorineural hearing loss (SNHL).

People with SNHL often experience auditory recruitment. Recruitment means a faster than normal growth of loudness between the elevated threshold and high sound levels, where loudness typically returns to normal values (Heinz et al, 2005). As early as 1944, loudness discomfort measures were being utilized as a method to identify auditory recruitment in SNHL (Watson, 1944). Recruitment can cause significant difficulty during a hearing aid fitting, because even though patients with SNHL have impaired auditory thresholds, the intensity level where sound is perceived to be uncomfortably loud, or the loudness discomfort level (LDL), does not increase (Stach, 2003). Therefore, the dynamic range (DR) of the input signal must be fit within a reduced dynamic range for the hearing impaired individual. Hearing aids must overcome this reduced dynamic range; soft sounds require amplification to become audible, whereas loud sounds must not be amplified to avoid painfully loud sounds (Heinz et al, 2005).

For many years, audiologists have recognized the need to address a patient's LDL during the hearing aid fitting. The most recent guideline, which has been submitted for publication by the American Academy of Audiology (AAA), reports that, at a minimum, the Output (Saturation) Sound Pressure Level (OSPL90) of a hearing aid should not exceed the listener's LDL in order to ensure patient comfort and to reduce exposure to potentially damaging input levels. The OSPL90 allows the audiologist to see the maximum output the hearing aid can generate and gives a quick estimate of aid's output.

Studies have shown that if a patient's OSPL90 in the hearing aid is programmed higher than an individually measured LDL, real-world loudness discomfort can occur (Munro and Patel, 1998). Dillon et al (1984) suggested that those with SSPL90 values that were less than or equal to LDL measurements reported no discomfort. However, if the OSPL90 surpasses a patient's LDL, loudness discomfort might be perceived in everyday environments. If incoming signals are amplified beyond the patient's LDL, Hawkins (1984) suggested the following could occur (Mueller and Hornsby, 2002):

- The hearing aid user continually changes the volume control to adjust for different input levels.
- To avoid the above, the user may use a low volume control setting, thereby sacrificing the audibility of lower level inputs.
- The hearing aid is only worn in quiet environments.
- The hearing aid is rejected because the disadvantages outweigh the benefits.

Dissatisfaction/rejection of hearing aids

It has been repeatedly reported that one of the primary reasons for dissatisfaction with hearing aids is loudness discomfort. A survey by Franks and Beckmann (1985) reported that

88% of the surveyed retirees that had rejected their hearing aids did so because they believed amplified sounds were too loud. Of those that accepted amplification, 32% still reported environmental sounds as “too loud.” More recent reports of hearing aid dissatisfaction due to loudness discomfort include results from a Kochkin (2000) survey, which found that the overamplification of environmental sounds is the third most common reason for hearing aid dissatisfaction. Additionally, 58% of those surveyed reported they would like to see “loud sounds less painful” as a “highly desirable” improvement in hearing aid technology. In a later survey of over 3000 hearing aid users, Kochkin (2002) reported that only 59% of those tested were satisfied with their hearing aids (re: loudness).

There are four predominant reasons hearing aids could be considered too loud by patients:

- (a) correction factors for binaural summation are not automatically applied in prescriptive targets and hearing aid manufacture software algorithms;
- (b) correction factors for channel summation are not automatically applied,
- (c) many audiologists use predicted measures instead of measuring individual LDLs, and
- (d) most hearing aid manufacturer modules for NOAH and manufacturers of real ear equipment employ predicted LDL data from Pascoe (1988).

Binaural Summation

First, not all fitting strategies adjust for binaural summation, and this potential oversight is a factor that may contribute to loudness discomfort. Binaural summation is the collective effect of sound reaching both ears, which results in an enhancement in hearing with both ears as opposed to only hearing monaurally. This is characterized by a binaural improvement in hearing sensitivity at the threshold level of approximately 3 dB over monaural sensitivity (Haggard and

Hall, 1982). While NAL-NL1 (Bryne et al, 2001) and the Independent Hearing Aid Fitting Forum (IHAF) protocol (Valente and Vliet, 1997) allow the audiologist to adjust for binaural summation, DSL [i/o] (Cornelisse et al, 1995) and FIG6 (Gitles and Niquette, 1995) do not.

Nearly every published article varies on recommendations for precise adjustments that can or should be made to account for binaural summation at a suprathreshold level. Older articles include suggestions for reduction as great as 7-10 dB for binaural summation (Hawkins et al, 1987). More recently, Bentler and Nelson (2001) studied the effects of binaural summation in 40 individuals (20 normal hearing and 20 hearing impaired) using pure tones, multitone complex (similar to speech) and continuous discourse. They reported a mean binaural summation of 6 dB independent of stimulus type. Looking specifically at 500, 1000, 2000 and 4000 Hz, Mueller and Bentler (2005) found binaural summation varied from a mean of 2.3 dB (4000 Hz) to 7.3 dB (at 500 Hz). To account for binaural summation, the overall gain at WUSM is decreased 3-6 dB HL, as recommended by the ASHA Ad Hoc Committee on hearing aid selection and fitting (ASHA, 1998; Valente and Vliet, 1997).

Channel Summation

Another major contributing factor to loudness dissatisfaction may include channel summation. Channel summation, also referred to as power summation, results when the channels in the hearing aid combine, resulting in a wider bandwidth and increase in output. As more channels are added, further reductions are needed. Currently, hearing aids can range from 2 to 64 channels. Obviously, the hearing aid with 64 channels (ie., Interton) has a greater chance of power summation compared to a two- channel hearing aid and thus needs to be adjusted accordingly. However, fitting strategies vary depending on the freedom the user is allowed regarding channel summation. For instance, FIG6 and IHAF assume two channels are being

used. In NAL-NL1 (Bryne et al, 2001), the user can choose one to four channels, and in DSL [i/o] information up to nine channels can be provided (Palmer and Lindley, 2002). However, NAL-NL1 is the only fitting strategy that corrects for channel summation (Bryne et al, 2001).

Predicting LDL from Threshold Measures

The third reason patients may experience loudness discomfort could be due to LDL measurements being prescribed, instead of measured, by audiologists. It has been reported that far too many audiologists fail to perform LDL measures regularly during hearing aid fittings. Martin et al (1998), who did not seek information regarding the type of stimuli used, reported that 90% of audiologists report they measure LDLs within their clinical setting. However, of these, only 60% perform LDL measurements on hearing aid candidates. Another survey, by Mueller (2003), also found only 60% of audiologists perform LDL measurements for hearing aid candidates. Out of those surveyed, only 27% of audiologists performed pure tone or narrow-band LDL measurements, while approximately 70% choose speech stimuli (Mueller, 2003).

Reasons for this small percent of audiologists that measure LDLs could include time constraints. Although using LDL predictions may save valuable clinical time, but studies are divided over the debate of whether or not these predictions are a valid indicator of real-world loudness discomfort. There are two studies that have researched the real-world validity of the LDL measures (Munro and Patel, 1998; Filion and Margolis, 1992). Filion and Margolis (1992) determined that the LDL is not an accurate predictor of loudness discomfort. In their study, seven young (mean = 25 years) individuals with normal hearing from a “happy hour” setting and six slightly older (mean = 42.5 years), normal to moderately hearing impaired individuals from a manufacturing plant were recruited. LDLs were measured for FM tones, speech and a sample of the noise from the corresponding environment. A questionnaire had subjects recall their

impressions of loudness comfort and discomfort from their noisy, real world, environment. This questionnaire reported large discrepancies between LDLs and judgments of loudness discomfort in real-life environments. One drawback of the Filion and Margolis (1992) study is the methodology, in which only a small sample size (N = 13 subjects) was collected. Comparisons within this study are difficult, since one group had normal hearing, and four out of the six individuals from the manufacturing plant had a mild to moderate hearing loss. Second, the recruitment setting could affect the subjects' percepts of loudness. People are more likely to tolerate loudness that is pleasing, such as music or loud talking in a bar, versus unpleasant sounds, such as a steady white noise or construction noise. Also, both groups could experience acclimatization to their particular setting, which would, in turn, cause them to rate the real world noise softer than normal on the questionnaire. Also, information on whether or not the subjects at the "happy hour" setting were imbibing alcohol, which could strongly effect loudness perceptions, was not included. Finally, Filion and Margolis (1992) reported there were faults within the questionnaire, which was used to relate the subjects' real-world experiences to the clinical measurements. The reliability of the questionnaire format was only "found to be adequate." Finally, subjects were required to base loudness judgments from memory since the questionnaire was given anywhere from five days to four weeks following exposure.

In an alternate study, Munro and Patel (1998) disagreed with the results of Filion and Margolis (1992). Twenty subjects (mean 68 years) were fit monaurally with a hearing aid. Individual LDLs were measured using a probe-tube microphone close to the eardrum. Individual real-ear to coupler differences (RECD) was added to the OSPL90 to predict the maximum power output (MPO) of the hearing aid. Next, the subjects completed a questionnaire at the time of the measurements asking them to rate the loudness of four types of environmental sounds within a

clinical setting. These results revealed a significant correlation between sounds of longer duration and the LDL measurements. Munro and Patel (1998) concluded that subjects are more likely to experience loudness tolerance problems when the output of the hearing aid exceeds their clinical measurement of loudness discomfort. Subjects did not express real-world auditory discomfort when the MPO of the hearing aid matched or was below the LDL.

Instead of reviewing real-world discomfort, other researchers studied whether or not measuring LDLs improves the accuracy of a hearing aid fitting. Preminger et al (2001) attempted to validate the NAL-R (Byrne and Dillon, 1986) maximum OSPL90 selection procedure for multi-channel hearing aids (Dillon and Storey, 1998). They reported that using individually measured LDLs rather than the predicted LDLs did little to improve the accuracy of the fitting. These predictions are based on their findings that 85% of the test subjects fell within the acceptable range for the low-frequency channel, but only 65% (18 out of 29 subjects) fell within the acceptable range in the high-frequency channel with a crossover frequency at 1500 Hz. The acceptable ranges were defined as the OSPL90 not being set too high, as subjectively reported by the subjects. These findings correspond with Kochkin's (2002) findings, which reported that only 59% of subjects were satisfied with their hearing aids. The large number of patients that fall outside the acceptable norms, for low and high frequency LDLs respectively, constitutes an ample cause to measure the individual, frequency-specific LDL.

Loudness Judgments Based from Pascoe (1988) Predictions

Finally, and the most important aspect to this study, most hearing aid manufacturers and manufacturers of real ear equipment predict patient LDL from Pascoe's (1988) study. This study measured the width of the DR – hearing threshold level to LDL - as well as the most comfortable level of hearing (MCL) for 508 hearing-impaired ears. To measure the LDL, subjects responded

to stimuli by choosing a descriptive anchor, ranging from “zero” to “nine,” that subjectively matched their perception of the loudness of the stimuli. This nine point scale includes:

- 9 “too loud”
- 8 “very loud”
- 7 “loud”
- 6 “OK (louder)”
- 5 “OK”
- 4 “OK (softer)”
- 3 “soft”
- 2 “very soft”
- 1 “too soft”
- 0 “nothing”

The LDL was defined as the level at which a subject reported on average that the input signal became “too loud.” This descriptive anchor was further explained to patients as the point at which sounds became “so loud that they would not want to hear anything stronger” (page 132). Data was then averaged to derive a mean LDL for each frequency at each hearing threshold in dB HL.

Since the time of Pascoe’s (1988) study, these averaged LDLs have been incorporated into many hearing aid manufacturer modules for NOAH and manufacturers of real ear equipment. Hearing aid manufacturer modules utilize a prescription, which is calculated from threshold and supra-threshold measurements, to create a target (or targets) of gain as a function of frequency. These targets are then used to select, set, or verify the hearing aid fitting (Palmer and Lindley, 2002). When fitting to a prescriptive target, some NOAH hearing aid modules

allow the audiologist to enter individual LDL information whereas others employ predicted measurements. A survey of 600 hearing professionals reported that 78% of audiologists “usually” or “always” use a prescriptive fitting technique (Kirkwood, 2003; Mueller, 2003). These audiologists, when asked to choose their preferred prescriptive method, strongly favored the National Acoustics Laboratory procedures (NAL) (e.g. 38% for NAL-NL1; 28% for NAL-R) or DSL [i/o] (27%). With DSL [i/o], LDLs can be either measured or predicted (Seewald et al, 1997). If predicted measurements are not obtained, DSL [i/o] predicts the LDL from Pascoe (1988). However, these LDLs are not used to determine the hearing aid’s maximum power output (MPO). Instead, the upper limit of comfort is predicted at one standard deviation below the mean data from Pascoe (1988) (Palmer and Lindley, 1997). NAL NL-1 does not allow frequency-specific LDL measurements to be entered.

Since the 1908’s the prescriptive method was changed to incorporate real-ear measures (REM), which uses a probe tube microphone to measure the sound pressure level (SPL) near the tympanic membrane. Not only does REM allow the audiologist to measure gain (output minus gain) and output, but instead of just trying to match measured gain to a prescribed gain target, targets measuring output can become measurements taken from the individual’s DR. Similar to the prescriptive method, the DR is determined from threshold and supra-threshold measurements measured or predicted in dB SPL. One benefit of using REM measurements in this manner is that specific resonances in the external auditory canal (EAC) are accounted for in the hearing aid fitting.

The three, primary real-ear manufacturers are Frye, AudioScan Verifit, and MedRX AVANT. All three of these REM manufacturers utilize live speech mapping (LSM) as a tool for verifying a hearing aid fitting. LSM incorporates the belief that calibrated and recorded speech

is the most realistic stimuli to evaluate the advanced signal processing within current hearing aids (Fabry, 2003). Using this method, the patient's real-ear unaided response (REUR) is first measured. Next, the audiogram is entered into the processor the real-ear equipment. Although this paper has cautioned against using predicted measures, once the hearing threshold information is tabulated, the software can then predict the LDL (dB HL), and threshold and LDL in dB SPL. These predictions are made possible by adding the ANSI S3.6-1989 mean real-ear-to-dial-difference (REDD) to the entered thresholds and predicted LDL in dB HL to derive the predicted hearing thresholds and LDL in dB SPL. The REDD is simply the difference in dB, across frequencies, between the SPL measured in the real-ear and the audiometer dial setting that produced it. If the measured LDL is not entered, the REDD is added to Pascoe's (1988) mean LDL for Frye and MedRX AVANT equipment or at one standard deviation below Pascoe (1988) mean LDL for the AudioScan Verifit.

After the patient's DR is measured or predicted, the real ear aided responses (REAR) for input levels of 50, 65 and 80 dB SPL is plotted using either the NAL-NL1 or DSL [i/o] prescription. Three in situ SPL measurements (e.g. soft = 55 dB SPL, comfortable = 65 dB SPL, and loud = 80 dB SPL) are made with the hearing aid turned on while seated in the ear. Soft input should always be audible (i.e., above the patient's thresholds); comfortable input should fall near the level speech is assumed to be comfortable, and loud speech input should always be below the measured or predicted LDL. Thus, all incoming speech should lie within the patient's DR.

Even though utilizing the Pascoe (1988) average has become the accepted norm to predict the LDL in most manufacturer modules and REM equipment, problems using this method exist. Since Pascoe (1988) recorded the LDL at "too loud," this prediction is apt to lead

to the output of a hearing aid fitting at which high input sound levels, that would not be tolerable, even for a brief period of time, could potentially exceed the LDL of a patient.

Problems associated with using predicted LDL

Other problems associated with using a predicted LDL include the descriptive anchor and intersubject variability. Dillon and Storey (1998) reported that the maximum OSPL90 could be predicted by estimating LDL from hearing threshold. However, several studies disagree (Kamm et al, 1978; Dillon et al, 1984; Hawkins et al, 1987; Pascoe, 1988; and Sammeth et al, 1993; Mueller and Bentler, 1994; Valente and Vliet, 1997). Although many audiologists choose to use predicted threshold (dB SPL) and LDL (dB HL and dB SPL) measurements, past studies reveal that the individual LDL cannot be confidently predicted from threshold. In fact, research has reported that for hearing thresholds less than 40 to 60 dB, there is little or no relation between the hearing threshold and individual LDL. Although trend lines reveal that the LDL increases as hearing loss increases above 60 dB, there has been little correlation found due to large intersubject variability (Kamm et al, 1978; Pascoe, 1988, Bentler and Cooley, 2001).

First, the descriptive anchor is the loudness rating of the level at which the LDL is recorded. Historically, many descriptive anchors have been used to define the LDL. One of the earliest papers on the LDL (Silverman, 1947) reported the LDL as the “threshold of pain” or “threshold of feeling.” Silverman (1947) believed three loudness thresholds existed: “threshold of discomfort,” “threshold of tickle,” and “threshold of pain.” The “threshold of tickle” was the threshold at which the patient would denote a ‘tickling sensation’ in the ear. The “threshold of pain” was the threshold at which the patient would feel a physical, sharp pain in the ear. Silverman (1947) reported that these “pain” and “tickle” thresholds are present at approximately 140 dB SPL! Currently, clinicians do not attempt to reach such high levels. More recently,

terms such as uncomfortable loudness (UCL), uncomfortable loudness level (ULL), and threshold of discomfort (TD) have been used to determine the level at which input signals become uncomfortable.

The incongruity of descriptive anchors is obvious by reviewing Pascoe (1988) and the IHAFF protocol (Valente and Vliet, 1997). The IHAFF protocol defines the LDL as the level at which input signals are ‘loud, but ok.’ Recall that the OSPL90 is programmed based upon the LDL (ie., OSPL90 below LDL). If the IHAFF procedure is used to measure individual LDLs, the resulting target would – ignoring other variables involved in determining the OSPL90 – lead to a hearing aid fitting resulting in incoming signals that should never be judged to be uncomfortable. IHAFF bases this guideline upon the 1987 study by Hawkins and his colleagues.

Pascoe (1988), on the other hand, defined the LDL as the level at which a subject reported that the input signal became “too loud.” Using this descriptive anchor, the mean LDL data represents input sound levels that would not be tolerable, even for a brief period. Using Pascoe’s (1988) description would allow the output of the hearing aid to be adjusted to a level at which uncomfortable sounds could potentially reach the patient’s ear. As discussed earlier, most manufacturers use this data to determine the appropriate output adjustment for the hearing aid. Therefore, it is possible, that by using the data reported by Pascoe (1988), the resulting output could be judged uncomfortably loud by a patient as opposed to adjusting the output to a patient’s judgment of when incoming signals as “loud, but ok.” If there is a significant difference between Pascoe (1988) and Washington University School of Medicine (WUSM) LDL (ie., IHAFF procedure), the author hypothesizes that utilizing the Pascoe (1988) LDL data would result in a louder output, which, in turn, could lead to judgments that amplification is “too loud.”

A second problem with predicting the LDL is that individual preferences for LDLs could vary greatly (Kamm et al, 1978; Dillon, et al, 1984; Sammeth et al, 1993; Munro et al, 1996; Valente et al, 1997; Eberling 1999; Bentler and Cooley, 2001). As intersubject variability increases, the ability to accurately predict the individual LDL decreases and the individual LDL can vary approximately 30 dB (± 15 dB) or more at any frequency for the same hearing threshold (Kamm et al, 1978; Dillon et al, 1984; Munro et al, 1996). Using speech stimuli, Sammeth et al (1993) reported intersubject variability of 38 dB (mean of 87.8; SD of 19.2) for 50 hearing impaired subjects. Munro et al (1996), in a study of 21 hearing impaired adults, reported that 95% of subjects (ie., two SD) lie within a range of 40 dB. Valente et al (1997), who reported a range of intersubject variability of less than 30 (at 1000 Hz using a ER-3A) to 51 dB (at 3000 Hz using a TDH-50P) from their study of 31 hearing impaired ears.

In a comparison of five studies (N=433) Bentler and Cooley (2001) reported that intrasubject variability could be as great as 60 dB (re: 2cc coupler). These data contradict Eberling (1999), who conducted a retrospective study of four studies, and concluded that for 70% of the hearing impaired, the LDL could be predicted from hearing threshold within ± 5 dB. He defined the remaining 30% as outliers that consist of “sound sensitive,” (12%; lower LDL than predicted) and “sound addicts,” (17%; higher LDL than predicted).

Variables when Measuring the LDL

Until now, the author has focused on the potential dissatisfaction with amplification that could arise from predicting LDL from threshold measurements. The author has not only examined the large percent of audiologists that appear to prefer predicted measures, but also demonstrated, through a literature review, that measuring the individual LDL could potentially benefit the clinician and patient.

However, this paper would not be complete without including the many variables that can confound clinical findings when measuring the LDL. Even though the LDL is commonly used in audiology clinics, stimulus materials and test procedures are poorly standardized (Punch et al, 2004). Clinicians must take into consideration that it is conceivable that variables such as stimuli, instruction to the patient, psychophysical method, dB HL versus dB SPL, transformation function from dB HL to dB SPL, and the effect of training could all influence results significantly. Each of these factors is briefly discussed in the following pages.

Type of Stimuli

Three signal types are typically used to measure LDL: pure tones, narrow band noise (NBN), or speech. Several studies (Morgan et al, 1974; Hawkins, 1980; Bentler and Nelson, 2001) have reported no significant difference in mean LDL as stimuli was changed. For example, Hawkins (1980) studied the differences stimuli might play on the measured LDL in normal hearing subjects by comparing pure tones, noise (one-third octave band wide-band, and multi-talker babble), as well as spondaic words and sentences. Nineteen normal hearing subjects, evaluated from 250-4000 Hz, were found to have no statistical differences in mean LDL regardless of stimulus. Bentler and Nelson (2001) studied the effects that spectral shaping and content stimuli (multitone, continuous discourse and pure tones) play on the LDL on 20 normal hearing and 20 hearing impaired subjects. Again, no significant differences were found in mean LDLs for the hearing impaired subjects with the different stimuli used in the study. There have been, however, other studies to contradict these findings. Bentler and Pavlovic (1989) compared LDLs measured with multi-tone complexes (speech-like stimuli) to LDLs measured with pure tone signals. They reported the mean LDL for a multi-tone complex is lower than the mean LDL for pure tone stimuli.

Those that have performed research on speech LDLs have discovered that more frequency specific information is warranted (Hawkins et al, 1987). With that in mind, when choosing a stimulus to measure LDLs, approximately 70% of audiologists select speech because it most closely represents what the patient will be listening to when listening in their “real world” (Mueller, 2003). Since speech is a broadband signal, important frequency specific information, which can range greatly from one frequency to the next, is not obtained. Without this frequency-specific information, determining the Real Ear Saturation Response with a 90 dB Input (RESR90) and OSPL90 target becomes difficult (Mueller and Bentler, 2005). Also, since the LDL changes as a function of frequency, a speech signal may result in a comfortable loudness judgment in one testing, but exceed loudness comfort at specific frequencies in another testing (Hawkins, 1980; Hawkins, 1984).

Suppose an audiologist records a speech LDL. Since few manufacturers allow an audiologist to manually enter a speech LDL when creating a RESR90 target, the audiologist would still need to ‘predict’ the LDL for frequency-specific information. Thus, the important benefits of measuring the LDL are ineffective. Therefore, by choosing to measure the LDL using only speech, many audiologists are predicting the LDL without even realizing it. These reasons convince this author that an audiologist should measure both pure tone and speech LDL.

Impact of instruction on LDL

Several researchers have noted that instructions to the patient can strongly influence the reliability, accuracy, and threshold of responses (Beattie et al, 1980; Bornstein and Musiek, 1993; Munro et al, 1996). Instructions can influence a patient from the ‘initial point of discomfort’ to ‘extreme discomfort’ (Wallenfels, 1967). During “extreme discomfort,” the subject may experience pain or show physical signs of discomfort.

As instructions change, so can the LDL (Beattie et al, 1979). For this reason, WUSM only incorporates the instructions used by Cox (1995). These instructions, which are fully described in the methodology section, measure the patient's average response to loudness discomfort.

Psychophysical Method on the LDL

Examples of psychophysical methods include the ascending approach (i.e., up-down procedure), loudness scaling, and the method of adjustment, which was used with Bekesy tracking. The primary difference between the ascending approach and loudness scaling is the starting point. Whereas the ascending approach starts at the patient's most comfortable level (MCL), loudness scaling starts at a patient's threshold and increases in 2 or 5-dB steps until LDL is reached. The Cox Contour Test (Cox, 1995), which is used at WUSM to report patient LDLs, is a common form of ascending method and will be explained in the methodology section.

Beattie and Sheffler (1981) reported the effects of a psychophysical method (adjustment versus limits) on the speech LDL. They found that mean LDLs for the respective methods of adjustment and limits were 86.8 and 92.9 dB SPL, respectively. These differences were found to be statistically significant. The authors also reported an order effect. When the method of limits was used first, the patient LDL was an average of 16 dB higher than when the method of adjustment was presented first.

Jenstad et al (1997) determined that measurement procedure effects occur for loudness perception data. In this study, LDLs were obtained for 40 normal hearing adults on a 9-point categorical scale. The authors found that there were significant interactions between sequences (random versus sequential). Loudness function exponents were greater when the stimulus levels were presented in sequential order ($M= 1.5$, $SD = 0.2$) than when the stimulus levels were

presented randomly ($M=1.3$, $SD = 0.2$). Further examination revealed that an ascending approach leads to higher LDLs than an descending approach.

dB HL versus dB SPL Measurements

Another variable is whether the LDL is measured in dB HL or dB SPL. Ideally, all measurements should be consistent, whether in dB HL or dB SPL. Most manufacturers of real ear equipment transform both threshold and LDL data from dB HL to dB SPL by adding average REDD conversions to the measured data while the probe tube remains in the patient's ear. It is important to note that if direct measurements are not made, REDD conversions are made using averaged data (Fabry, 2003).

However, direct measurements, at the position of where the hearing aid will be placed, has resulted in SPL measures to become increasingly popular. As previously mentioned, this direct measurement accounts for specific resonances in the EAC, which, if performed accurately, decreases the chance for error. However, audiologists should be aware of the potential for miscalculation if the probe is placed greater than 4-6 mm from the tympanic membrane. As the distance from the tympanic membrane increases, the greater the potential for error in the higher frequencies becomes because of the effect of standing waves in the EAC (Dirks and Kincaid, 1987).

Transformation Function from dB HL to dB SPL

As reported above, the measured LDL can be impacted by many variables. If measurements are made in dB HL, how variable is the transformation to dB SPL? Clinically, the REDD is used primarily when the transducer is headphones. Research on the REDD has become necessary with the increasingly popular use of measuring the output during real-ear measures (ie., AudioScan Verifit; MedRx Live Speech Mapping; Frye SPL-Mode). Mean REDD values

have been reported for several prescriptive fitting formulae (e.g., DSL [i/o]; NAL-NL1). Both DSL [i/o] and NAL-NL1 allow the clinician to enter individual REDD measures. Mean REDD values presented in the ANSI S3.6 – 1989 are 12.0, 9.0, 15.0, 13.0 dB for 500, 1000, 2000, and 4000 Hz respectively.

This average conversion has been determined to be accurate within 2.3 dB of the “true” real ear SPL (Scollie et al., 1998). Scollie et al (1998) measured the REDD by sweeping pure tones at 70 dB across frequencies in 24 normal-hearing subjects. They reported that the REDD has a high test-retest reliability by testing subjects twice within the same test session (intra-test retest reliability). To ensure accuracy, the probe tube and earphone was completely removed and reinserted during the same visit. They reported mean test-retest differences that ranged from 0.9 to 2 dB across test frequencies of 250 to 4000 Hz. The mean or predicted REDD fell within 3 dB from the measured REDD with confidence intervals of +/- 2.4 dB for 1000 and 2000 Hz and approximately +/-5 dB for 250, 500, and 4000 Hz. Valente et al (1997) measured the LDL in dB HL and dB SPL and reported mean REDD values of 12.0, 8.8, 15.5 and 7.4 dB for 500-4000 Hz respectfully. These mean values closely correlate with the mean REDD used by NAL-NL1 and ANSI S3.6-1989 for 500 – 2000 Hz. However, at 4000 Hz the Valente et al (1997) data is approximately 6 dB lower.

Effect of Training on the LDL

Do LDLs change with patient practice? Studies suggest that a small “training effect” might occur when re-testing LDLs. For example, 95% of test-retest differences in the LDL, measured with speech, was within 4 to 8 dB from the individual LDL (Beattie et al, 1979; Sammeth et al, 1989). Beattie and Sheffler (1981) reported that 50% of their listeners had test-retest differences of +/-2 dB and nearly all subjects obtained LDLs within 8 dB.

Purpose

The major purpose of the present study is to determine if the measured WUSM LDLs (ie., the IHAF procedure) are significantly different than mean LDL data reported by Pascoe (1988). If there is a significant difference between the two methods, the author will hypothesize that utilizing the Pascoe (1988) LDL data will result in a louder output that could lead to increased probability of patients reporting the amplified sound is too loud, which has been reported to result in increased reports of amplification. The study will also address the following experimental questions:

1. How does the intersubject variability of the WUSM method to measure the LDL vary from previous studies (Kamm et al, 1978; Pascoe 1988; Valente et al, 1997; Eberling 1999; Bentler and Cooley, 2001)?
2. Can the LDL be accurately predicted from threshold measurements or averaged group data?
3. Is the mean LDL frequency dependent?
4. Does the mean LDL increase with hearing loss?
5. Can the REDD be accurately predicted from averaged group data?
6. How does the mean, frequency-specific WUSM REDD data compare to mean REDD data for NAL-NL1, DSL [i/o], ANSI S3.6-1989 and Valente et al (1997)?

METHODOLOGY

Measuring Loudness Discomfort levels (LDL)

Data collection

Data included in this study were acquired from two sources. First, a retrospective chart review was undertaken at three Washington University School of Medicine (WUSM) at St. Louis clinical sites: Center for Advanced Medicine (CAM), West County ENT Clinic (WC), and Central Institute for the Deaf (CID). All clinical sites utilize the same protocol for measuring the LDL. Ear specific LDL (500, 1000, 2000, 3000, and 4000 Hz) was measured in dB HL and dB SPL and recorded in the subject's chart at the time of the hearing aid evaluation (HAE). In order to be included for this study, chart information had to include threshold and LDL measures for at least one frequency. LDLs were measured on subjects who were going to be fit with hearing aids. A limited number of subjects with normal hearing were available. Therefore, additional subjects with normal to mild hearing losses were recruited for threshold and LDL measures. Attention was taken to follow the same protocol used at the three WUSM clinical sites.

Subjects

A total of 435 ears from 306 subjects, ranging in age from 17 to 97 years (Mean = 67.6 years, SD = 17 years), were retrieved from the retrospective chart review. The magnitude of hearing loss ranged from normal to profound. An additional 19 normal hearing subjects (N = 38 ears) were recruited by the investigator and LDLs were measured in the same manner as were measured at the three clinical sites. Combined, 51.9% were male and 48.1% were female.

Procedures

All thresholds and LDLs were measured utilizing calibrated (ANSI-1996) Grason-Stadler GSI-61 audiometers to generate the pure tone signals at 500, 1000, 2000, 3000 and 4000 Hz. Thresholds were measured during the subject's initial visit, and LDL was typically measured at the second visit during the hearing aid evaluation (HAE). Threshold and LDL measurements

were made on the first visit for the normal hearing subjects. All data was collected in a double-walled sound-treated audiometric booth at two (CAM and WC) of the three sites and a single-walled suite at the other site (CID).

Threshold and LDL measurements were obtained using either TDH-50P or ER-3A earphones, depending upon clinician preference. The ER-3A earphones were coupled to the ear canal using either a foam plug or a Grason Stadler immittance probe cuff on a plastic adapter (ER3-06) 96 (Valente et al, 1997).

To measure LDL in SPL dB near the eardrum, a probe tube was marked 30 mm from the tip and inserted into the ear canal so the mark was adjacent to the intertragal notch. On the average ear, this ensures placement of the probe tip ~4-6 mm from the tympanic membrane (Dirks and Kincaid, 1987; Valente et al, 1997). The probe tube was then taped in place to ensure it would not shift during measurements.

Using a calibrated Frye 6500, the reference microphone was “Disabled.” All LDL readings (in dB SPL) were obtained by reading the “Probe” output under the “Calibrate Probe” mode of the Frye 6500.

To measure the LDL, the patient was provided a laminated page of descriptive anchors ranging from “zero” to “seven.” The categorical loudness anchors and instructions are modified from the Cox Contour Test (Cox, 1995). The reliability of this test has been verified (Cox et al., 1997; Palmer and Lindely 1998; Ricketts and Bentler, 1996). This seven point scale includes:

- 7 “uncomfortably loud,”
- 6 “loud but okay,”
- 5 “comfortable, but slightly loud,”
- 4 “comfortable,”

- 3 “comfortable, but slightly soft,”
- 2 “soft,”
- 1 “very soft,” and
- 0 “cannot hear at all”

Instructions to the subjects closely followed those suggested by Cox (1995):

“The purpose of this test is to find your judgments of the loudness of different sounds. You will hear sounds that increase and decrease in volume. I want you to make a judgment about how loud the sounds are. Pretend you are listening to the radio at that volume. How loud would it be? After each sound, tell me which of these categories best describes the loudness. Keep in mind that an uncomfortably loud sound is louder than you would ever want on your radio no matter what mood you were in. When responding to each sound, it is OK to skip a category, or to repeat a category. Do you have any questions?”

After a practice run, the clinician initially starts at 20 dB SL re: the patient’s threshold at 1000 Hz. Utilizing an ascending procedure in 10 dB steps the clinician reaches the “uncomfortably loud” level. The clinician decreases 10 dB then repeats an ascending procedure, in 5 dB steps. The recorded LDL (in dB HL and dB SPL) is the level where “loud, but ok” is reached 50% of the time. Usually, this was accomplished in three runs. The same procedure was completed for 500, 2000, 3000, and 4000 Hz. The difference between LDL (dB HL) and LDL (dB SPL) is referred to as the real-ear-to-dial-difference (REDD) and this value was recorded for all LDL measures.

Recording the Data

Data was reorganized into a second Excel spreadsheet by hearing threshold, LDL dB HL and LDL dB SPL. Data for hearing thresholds from 0-115 dB was recorded. If a patient had three hearing thresholds that were 0 dB HL, he or she was recorded three times in the spreadsheet, once for each threshold. Next, data was sorted by frequency in ascending order for LDL dB HL and LDL dB SPL. This sorting allowed for the acquisition of mean, standard deviation, and the number of subjects for each frequency at each hearing threshold. All results were calculated from this spreadsheet.

Pascoe's (1988) charts and tables were recreated for easy comparison between the two studies. Since Pascoe (1988) only computed LDL measurements in dB HL, the mean ANSI3.9 – 1989 REDD data was added to mean LDL HL data to derive mean LDL SPL predictions.

RESULTS

The mean age for the 325 (306 hearing impaired; 19 normal) subjects was 67.6 years (SD = 16.7 years). Over 100 measurements were made for each threshold between 30-65 dB HL. Greater than 50 measurements were taken for each hearing threshold between 5-25 dB HL and 70 dB HL. However, for each threshold between 90-120 dB HL, at each frequency, less than 10 measurements were recorded. The least number of data points (316 in Table 1) was recorded for 4000 Hz. This is because the output of the audiometer was reached prior to a descriptive anchor of six (i.e. 'loud, but ok') could be obtained for some patients. The mean hearing threshold (in dB HL) for combined left and right ears were: 33.8 dB for 500 Hz, 39.0 dB for 1000 Hz, 45.4 dB for 2000 Hz, and 54.8 dB for 4000 Hz. This audiometric configuration represents the typical, sloping high frequency hearing loss observed during the chart review.

Mean Differences between WUSM and Pascoe (1988)

dB HL

Table 1 reports the mean, standard deviation (SD), and the number of data points for the LDL (dB HL) measured at WUSM. Also provided are the minimum, maximum, range at each frequency and grand mean, SD, and the grand number of data points at 500, 1000, 2000, and 4000 Hz. The mean LDL (dB HL) and SD (dB HL) values were as follows: 98.9 (9.7), 99.4 (9.3), 99.6 (10.5), 102.5 (10.7) at 500, 1000, 2000, and 4000 Hz, respectively. The grand mean LDL (dB HL) as a function of hearing threshold was derived by averaging individual LDL HL measurements across the four frequencies. These values ranged from 96.3 – 115.0 dB HL.

Table 2 reports the mean, standard deviation (SD), and the number of data points for the LDL (dB HL) measured by Pascoe (1988). Also provided are the minimum, maximum, range at each frequency and grand mean, SD, and the grand number of data points at 500, 1000, 2000, and 4000 Hz. Further calculations were completed to allow for easy comparison of the two groups. For instance, the mean LDL dB HL was computed by averaging mean LDL dB HL for each of the hearing thresholds; however, the mean standard deviations were taken from individual LDL measurements as recorded by Pascoe (1988). The mean LDL (dB HL) and SD (dB HL) values were as follows: 106.5 (7.9), 109.4 (7.4); 111.8 (7.3) and 114.9 (7.0) at 500, 1000, 2000, and 4000 Hz, respectively. The grand mean LDL as a function of hearing threshold was derived by averaging mean LDL (dB HL) across the four frequencies. These values ranged from 96.3 – 115.0 dB HL.

Table 1. Mean, standard deviation (SD), and number of data points (N), for LDL at four frequencies (in dB HL) measured at WUSM. Also provided are minimum, maximum, range at each frequency and grand mean, SD, and N.

<u>HL</u>	<u>500 Hz</u>			<u>1000 Hz</u>			<u>2000 Hz</u>			<u>4000 Hz</u>			<u>Grand Mean</u>	<u>Grand SD</u>	<u>Grand N</u>
	<u>Mean</u>	<u>SD</u>	<u>N</u>	<u>Mean</u>	<u>SD</u>	<u>N</u>	<u>Mean</u>	<u>SD</u>	<u>N</u>	<u>Mean</u>	<u>SD</u>	<u>N</u>			
0	100.7	9.3	7	104.0	4.2	5	97.7	8.6	13	96.3	5.7	12	98.6	7.6	37
5	95.9	10.5	20	99.6	8.8	27	96.7	6.7	15	93.8	8.9	13	97.0	9.0	75
10	100.4	8.6	26	100.6	8.3	12	100.0	11.1	11	92.5	6.0	8	99.2	9.0	57
15	98.2	9.3	35	99.9	7.9	33	94.6	2.9	8	90.0	7.1	6	97.9	8.5	82
20	99.1	10.2	44	97.9	9.7	25	96.9	2.2	7	100.0	--	1	98.5	9.5	77
25	98.9	9.8	44	94.7	9.2	33	99.1	5.6	8	100.0	0.0	2	97.4	9.3	87
30	97.2	8.1	40	98.8	7.3	46	94.3	10.6	23	97.1	7.0	7	97.3	8.3	116
35	95.9	10.8	41	98.1	10.8	40	94.5	9.9	39	99.1	9.8	7	96.3	10.5	127
40	98.2	9.7	37	96.9	9.2	44	97.6	10.0	47	101.9	11.3	8	97.8	9.7	136
45	98.7	9.9	28	95.9	8.4	43	96.4	12.0	44	98.0	10.2	25	97.0	10.2	140
50	98.3	10.2	26	101.0	10.2	40	98.3	10.4	54	97.6	10.0	17	99.0	10.2	137
55	102.0	8.3	22	98.9	10.0	21	99.5	10.9	53	101.9	11.4	38	100.5	10.5	134
60	99.4	9.8	19	99.5	8.9	35	103.3	8.6	49	101.1	11.3	44	101.2	9.8	147
65	100.9	9.0	18	105.4	6.9	23	106.8	9.9	22	105.0	9.0	43	104.8	8.9	106
70	105.1	8.5	7	107.2	8.7	12	107.4	6.7	21	107.1	8.7	36	107.0	8.0	76
75	106.6	6.1	8	108.7	7.0	6	106.4	6.6	9	108.5	8.4	19	107.7	7.3	42
80	117.3	4.6	3	107.1	5.5	9	117.3	3.2	4	113.2	5.5	11	112.2	6.3	27
85	103.3	12.6	4	109.0	3.6	3	112.8	6.7	6	112.3	6.5	9	110.4	8.0	22
90	105.0	--	1	110.0	--	1	104.5	0.7	2	113.2	7.7	5	110.0	6.8	9
95	--	--	--	115.0	--	1	116.2	4.1	5	117.5	3.5	2	116.4	3.5	8
100	--	--	--	105.0	7.1	2	--	--	--	112.0	--	1	107.3	6.4	3
105	--	--	--	--	--	--	--	--	--	110.0	--	1	110.0	--	1
110	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
115	--	--	--	--	--	--	--	--	--	115.0	--	1	115.0	--	1
120	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
GRAND	98.9	9.7	430	99.5	9.3	461	99.6	10.5	440	102.5	10.7	316			1647
Minimum	74			72			74			78					
Maximum	120			120			120			120					
Range	46			48			46			42					

Table 2. Mean, standard deviation (SD), and number of data points (N), for LDL at four frequencies (in dB HL) measured by Pascoe (1988). Also provided are minimum, maximum, range at each frequency and grand mean, SD, and N.

<u>HL</u>	<u>500 Hz</u>			<u>1000 Hz</u>			<u>2000 Hz</u>			<u>4000 Hz</u>			<u>Grand</u>	<u>Grand</u>	<u>Grand</u>
	<u>Mean</u>	<u>SD</u>	<u>N</u>	<u>Mean</u>	<u>SD</u>	<u>N</u>	<u>Mean</u>	<u>SD</u>	<u>N</u>	<u>Mean</u>	<u>SD</u>	<u>N</u>	<u>Mean</u>	<u>SD</u>	<u>N</u>
0	95.4	9.9	14	101.7	6.8	6	--	--	--	--	--	--	97.3	8.4	20
5	97.7	8.7	22	101.3	6.4	12	--	--	--	--	--	--	99.0	8.1	34
10	99.1	7.3	39	100.3	6.9	15	110.0	--	1	90.0	--	1	99.5	7.3	56
15	97.5	6.8	57	99.1	6.5	32	100.0	--	1	100.0	--	1	98.1	6.7	91
20	95.0	7.8	45	100.0	7.1	26	95.0	2.9	2	91.7	2.9	3	96.6	7.7	76
25	100.4	8.9	40	102.4	8.5	40	100.0	9.6	5	100.0	--	2	101.3	8.6	87
30	102.5	8.4	54	100.8	8.1	40	100.5	9.1	11	106.7	12.6	3	101.8	8.4	108
35	97.9	8.7	57	103.3	7.4	51	103.3	7.3	24	110.0	--	1	101.1	8.4	133
40	100.2	9.7	45	102.1	9.5	63	104.6	9.0	52	105.0	7.1	15	102.6	9.3	175
45	106.6	9.5	32	104.2	8.9	63	105.6	8.9	63	106.3	11.4	19	105.4	9.3	177
50	105.3	9.8	33	108.2	9.2	47	107.7	9.1	84	107.8	10.2	34	107.5	9.4	198
55	106.7	8.5	27	108.4	7.8	44	107.6	8.3	82	107.8	8.9	66	107.7	8.4	219
60	108.2	6.1	17	108.3	6.5	26	110.6	7.1	78	110.3	8.4	67	110.0	7.5	188
65	108.3	7.5	6	112.9	6.8	22	115.2	8.0	30	113.9	9.3	58	113.8	8.5	116
70	103.8	4.8	4	114.3	6.7	7	115.5	8.4	22	115.4	7.7	61	114.8	8.0	94
75	116.3	6.4	8	122.5	10.6	2	116.9	7.2	24	116.6	5.7	57	116.8	6.2	91
80	115.0	5.8	4	120.0	10.0	3	119.7	4.6	16	120.5	6.0	33	119.8	5.9	56
85	125.0	--	1	117.5	4.2	6	120.7	6.7	7	120.2	6.5	22	120.0	6.1	36
90	130.0	--	1	127.5	3.5	2	120.0	--	1	123.3	5.6	26	123.7	5.6	30
95	120.0	--	1	--	--	--	--	--	--	130.5	7.2	10	129.6	7.6	11
100	--	--	--	--	--	--	--	--	--	126.6	6.3	16	126.6	6.3	16
105	--	--	--	--	--	--	135.0	7.1	2	132.5	4.2	6	133.1	4.6	8
110	--	--	--	135.0	--	1	136.7	2.9	3	131.7	7.6	3	134.3	5.4	7
115	--	--	--	--	--	--	--	--	--	136.7	5.8	3	136.7	5.8	3
120	--	--	--	--	--	--	--	--	--	140.0	--	1	140.0	--	1
GRAND	106.5	7.9	507	109.5	7.4	507	111.8	7.3	503	114.9	7.0	479			2031
Minimum	95			99			95			90					
Maximum	130			135			137			140					
Range	35			36			42			50					

Figure 1 reports the mean WUSM LDL in dB HL for 500, 1000, 2000, 4000 Hz. Hearing threshold (0 to 120) is along the abscissa and mean LDL (HL) is along the ordinate. The solid line represents the hearing threshold level (in dB HL). The solid line (with the squares) illustrates the grand mean. Note how the frequencies do not significantly depart from this grand mean. This is in agreement with Pascoe's (1988) findings, which are replicated in **Figure 2**.

Figure 3 reveals scatterplot, Y intercept, regression coefficients, number of data points, and the line of best fit for 500, 1000, 2000, and 4000 Hz for dB HL. For each scatterplot, hearing threshold is along the abscissa and LDL (dB HL) is along the ordinate. The diamonds represent the individually measured LDL (dB HL) as a function of hearing threshold. Since subjects with the same hearing threshold and LDL are reported as one dot, the number of dots does not accurately represent the recorded data. The LDL in dB HL can be predicted from hearing threshold by using the formula:

$$Y = b(X) + a$$

Where Y is the predicted LDL in dB HL, b is the regression coefficient, and X is the measured hearing threshold (in dB HL) and a is the Y intercept. For example, using the data in **Figure 3** at 1000 Hz, the predicted LDL (dB HL) for a hearing loss of 40 dB would be $0.0891(40) + 95.991$ or 99.6 dB HL (Valente et al, 1997). The Pearson product-moment correlations (r), which were used to derive the R^2 values, ranged from .130 (at 500 Hz) to .447 (at 4000 Hz). This correlation reports how much individual LDL measurements vary around the line of best fit at each frequency. For example, the poorest correlation ($r = .130$ at 500 Hz), signifies that only 1.7% ($R^2 = .0167$) of data was explained by the line of best fit in 500 Hz. This indicates that 98.3% of the

data was unaccounted for. However, for 2000 and 4000 Hz, correlations improve, and fall into the moderately correlated.

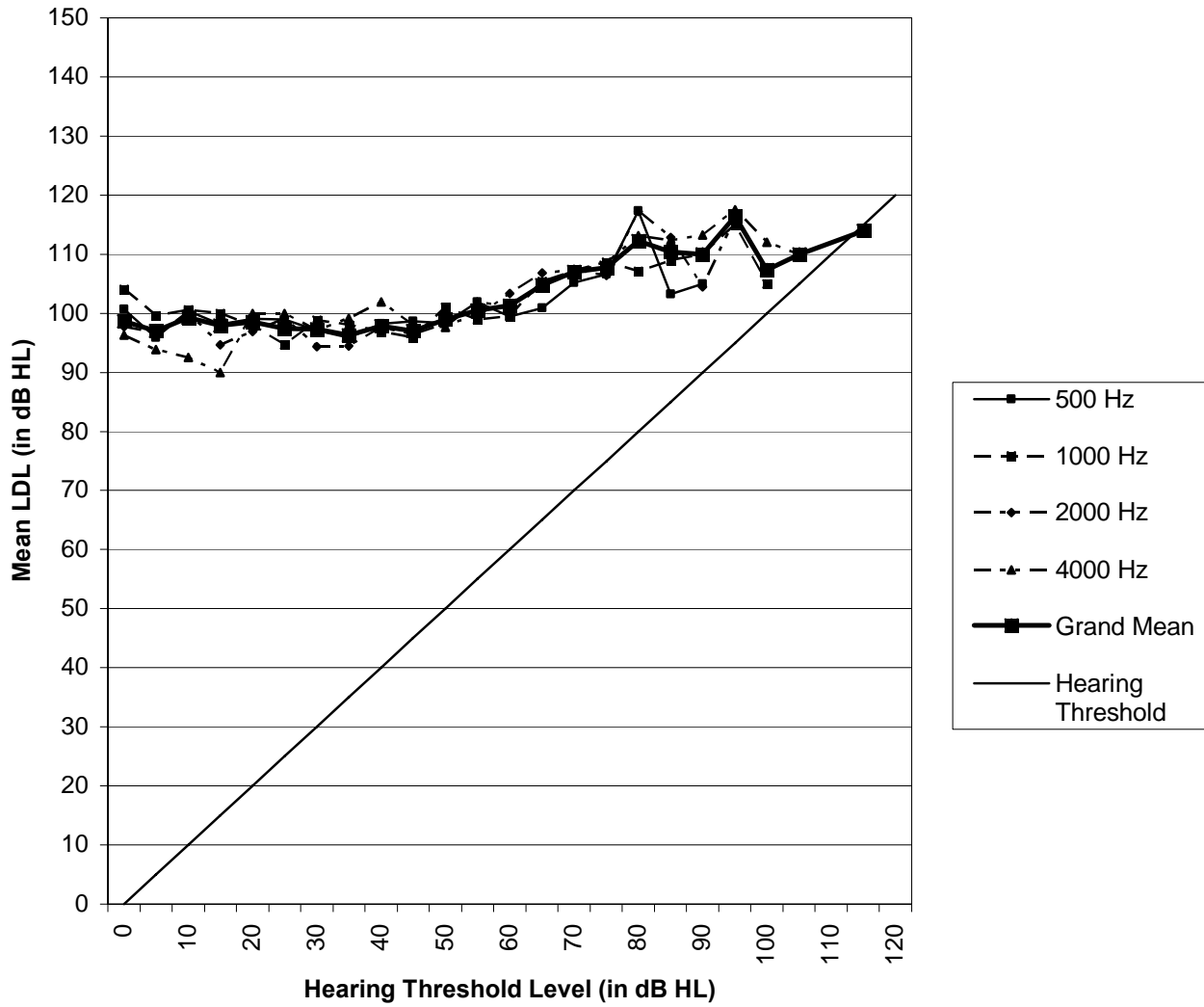


Figure 1. WUSM mean LDL (dB HL) as a function of frequency. Also provided is the grand mean and hearing threshold.

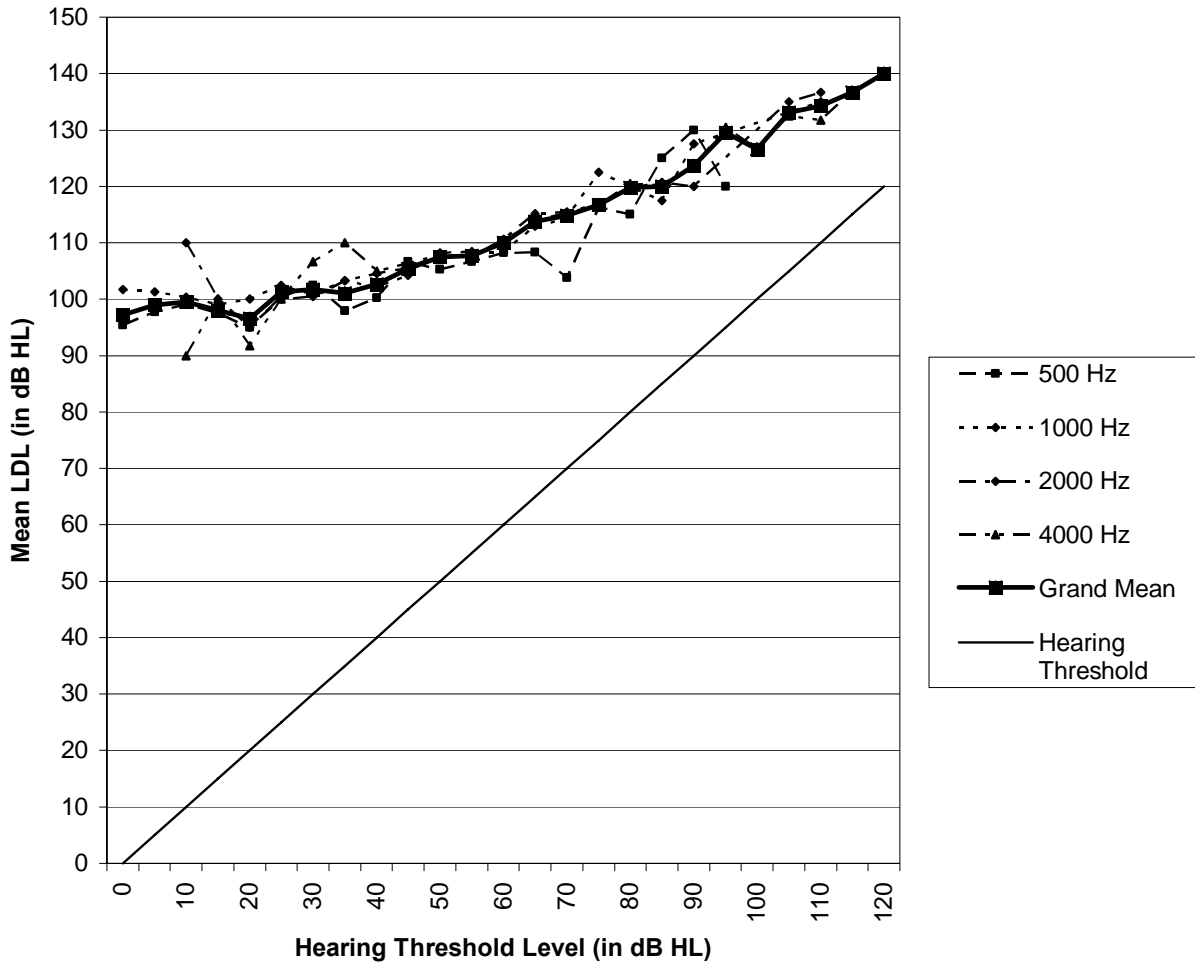


Figure 2. Pascoe (1988) mean LDL (dB HL) as a function of frequency. Also provided is the grand mean and hearing threshold.

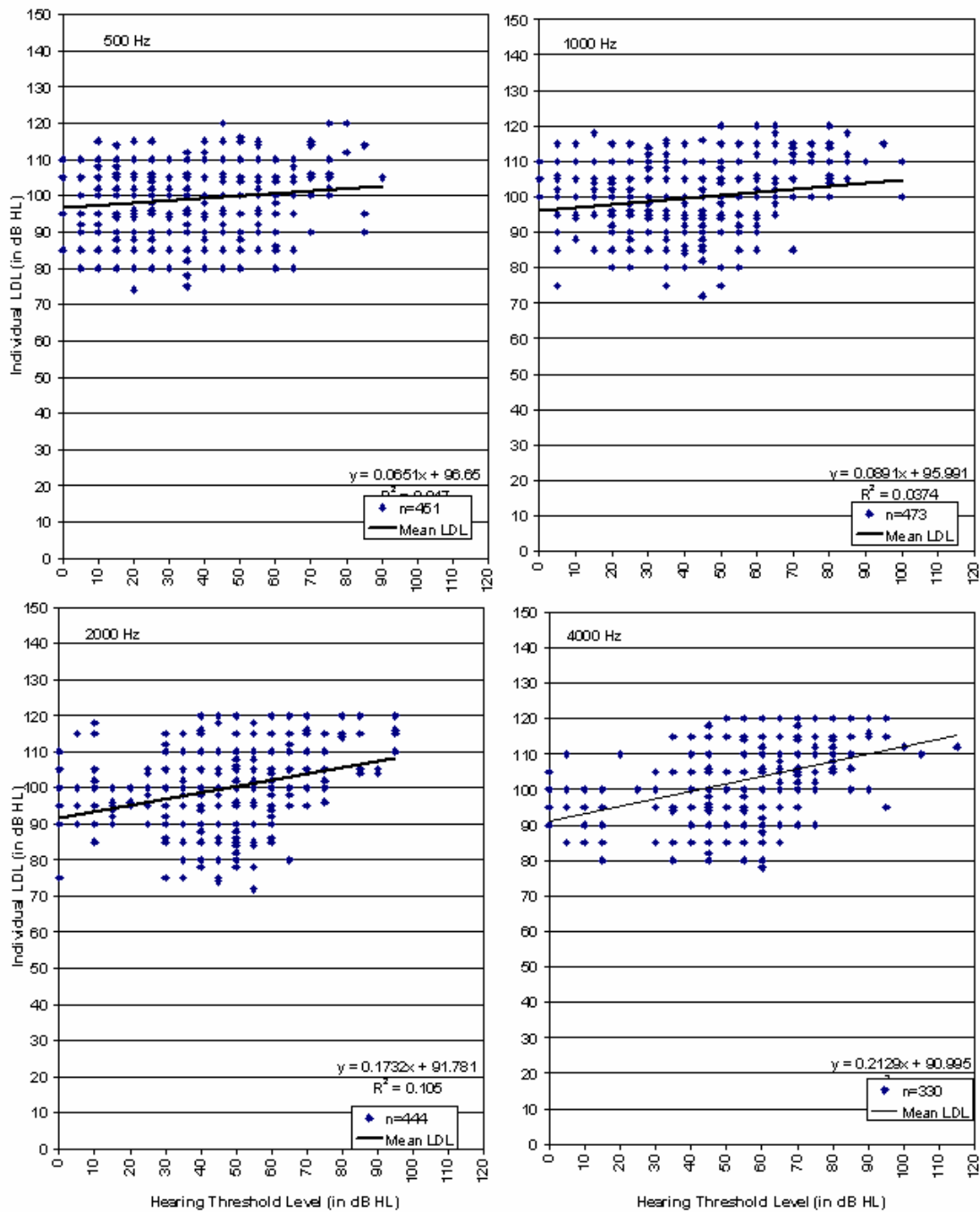


Figure 3. Scatterplots at 500 - 4000 Hz for WUSM LDLs (in dB HL) as a function of hearing level. Also provided is the line of best fit (mean LDL), number of data points (N), R^2 values and the equation for the line of best fit

range ($r=0.324$ and 0.447 respectively). These weak correlations signify that hearing threshold measurements are better predictors of LDL at higher frequencies. This is in agreement with previous studies (Kamm et al, 1978; Dillon et al, 1984; Hawkins et al, 1987; Pascoe, 1988).

A two-factor repeated measures ANOVA (LDL (dB HL) \times frequency) revealed a significant two-factor interaction ($F = 9.573$; $df = 3/1693$; $p < .001$), indicating that significant differences were present in the mean LDLs for at least one of the test frequencies. Post-hoc analysis (Bonferroni/Dunn test of multiple pair-wise comparisons) indicated that the main effect of frequency was driven by significant differences between 4000 Hz to the three other test frequencies ($p < .001$). Results (the mean difference and critical difference in parenthesis) are: -3.6 (1.9), -3.042 (1.8), -2.9 (1.9) for 500 to 4000, 1000 to 4000, and 2000 to 4000, respectively. Therefore, results reveal that there is no statistically significant difference between mean LDLs measured at 500, 1000 and 2000 Hz, but the mean LDL at 4000 Hz was statistically significantly different from the mean LDLs measured at 500, 1000 and 2000 Hz.

To illustrate the findings from **Table 2**, **Figure 4** superimposes Pascoe's (1988) data with WUSM mean LDL (in dB HL) as a function of hearing threshold for 500, 1000, 2000, and 4000. **Figure 5** superimposes grand mean LDL HL and standard deviation bars from Pascoe (1988) and WUSM. The squares represent the mean WUSM LDL HL and the triangles represent the mean Pascoe (1988) LDL HL. Standard deviation bars report ± 2 standard deviations (SD), or the point for which 95% of all data can be accounted. For hearing thresholds where less than two data points was reported, standard deviation bars are not reported.

The reader may notice the increased difference between WUSM and Pascoe mean LDL as hearing threshold becomes poorer. Generally, the Pascoe (1988) mean LDL departs from WUSM

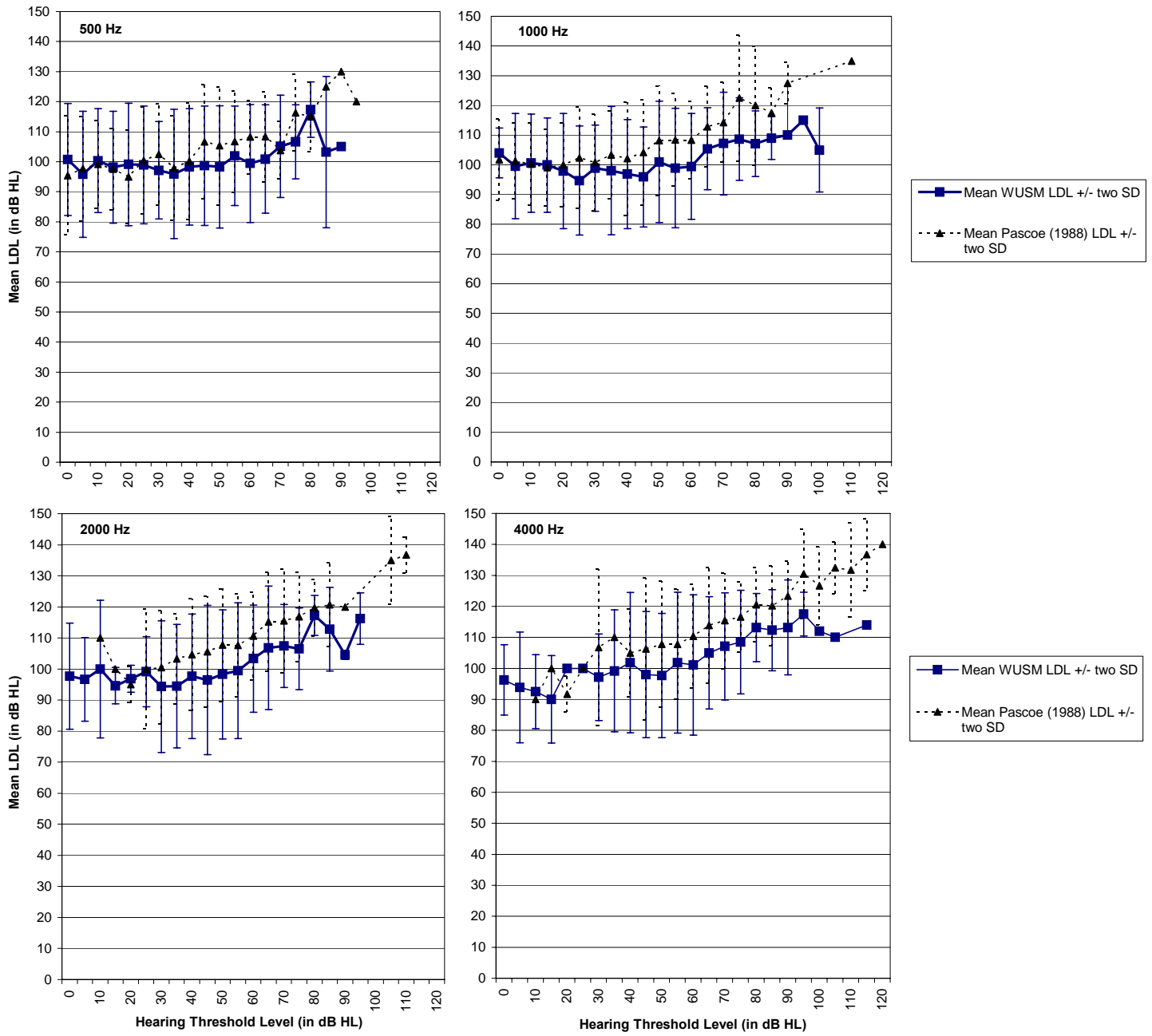


Figure 4. Comparison of mean LDL (in dB HL) as a function of frequency for Pascoe (1988) and WUSM data. Error bars represent +/- 2 SD.

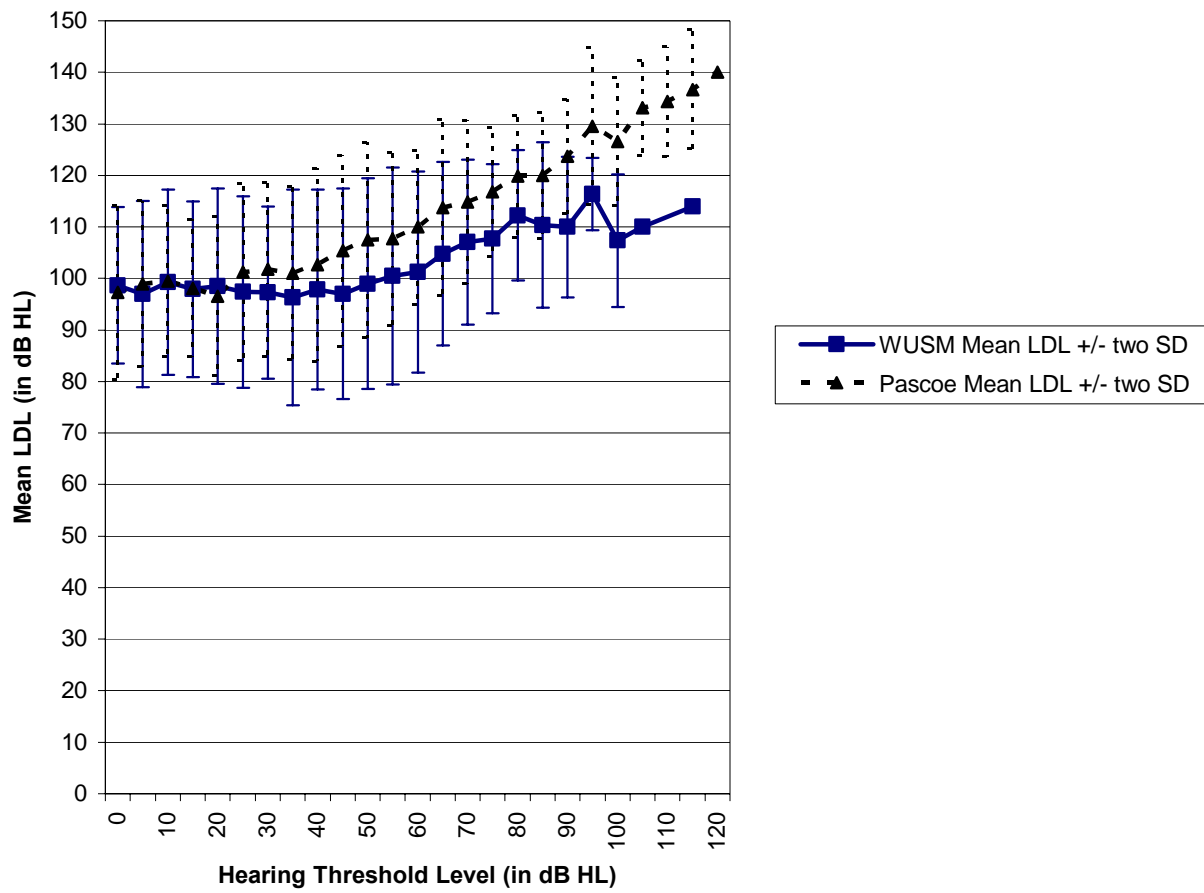


Figure 5. Comparison of Grand Mean WUSM LDL (dB HL): re Pascoe's Data on Hearing Threshold Levels. Error bars represent +/- 2 SD.

mean LDL data at approximately 30 dB for each of the tested frequencies (500, 1000, 2000, and 4000 Hz) (**Figure 4**) and the grand mean (**Figure 5**). This difference widens as hearing loss decreases with increasing hearing threshold, with the Pascoe (1988) mean LDLs greater than the mean WUSM LDLs.

Comparisons between Pascoe (1988) and WUSM were analyzed via repeated measures ANOVA (Pascoe (1988) versus WUSM \times frequency \times Pascoe (1988) versus WUSM \times frequency). In general, the main effect for Pascoe (1988) versus WUSM indicated a significant difference between the two sets of data ($F = 33.022$; $df = 1/157$; $p < .01$). There was not a main effect for frequency ($F = 2.446$; $df = 3/157$; $p = .06$) or an interaction between frequency and test site ($F = .862$; $df = 3/157$; $p = .46$).

dB SPL

Table 3 reports the mean, standard deviation (SD), and the number of data points for the LDL (dB SPL) measured at WUSM. Also provided are the minimum, maximum, range at each frequency and grand mean, SD, and the grand number of data points at 500, 1000, 2000, and 4000 Hz. The mean LDL (dB SPL) and SD (dB SPL) values were as follows: 110.9 (9.9), 109.2 (9.9), 114.8 (11.1), 113.9 (11.0) at 500, 1000, 2000, and 4000 Hz, respectively. The grand mean LDL (dB SPL) as a function of hearing threshold was derived by averaging individual LDL SPL measurements across the four frequencies. These values ranged from 108.41 – 128.50 dB SPL

Table 4 reports the mean, standard deviation (SD), and the number of data points for the LDL (dB SPL) measured by Pascoe (1988). Also provided are the minimum, maximum, range at each frequency and grand mean, SD, and the grand number of data points at 500, 1000, 2000, and 4000 Hz. As previously mentioned, the Pascoe (1988) dB SPL values were computed by

Table 3. Mean, standard deviation (SD), and number of data points (N), for LDL at four frequencies (in dB SPL) measured at WUSM. Also provided are minimum, maximum, range at each frequency and grand mean, SD, and N.

<u>HL</u>	<u>500 Hz</u>			<u>1000 Hz</u>			<u>2000 Hz</u>			<u>4000 Hz</u>			<u>Grand</u>	<u>Grand</u>	<u>Grand</u>
	<u>Mean</u>	<u>SD</u>	<u>N</u>	<u>Mean</u>	<u>SD</u>	<u>N</u>	<u>Mean</u>	<u>SD</u>	<u>N</u>	<u>Mean</u>	<u>SD</u>	<u>N</u>	<u>Mean</u>	<u>SD</u>	<u>N</u>
0	111.8	8.3	7	112.1	2.7	5	112.4	9.3	13	108.6	4.5	12	111.0	7.1	37
5	108.8	10.0	20	106.9	8.7	27	106.9	8.7	15	106.5	8.3	13	108.4	8.7	75
10	112.1	8.2	26	108.6	8.0	12	116.3	10.3	11	108.3	10.1	8	111.6	9.1	57
15	110.1	9.4	35	109.6	7.8	33	112.7	8.7	8	102.6	5.8	6	109.6	8.6	82
20	110.5	11.9	44	109.4	10.0	25	115.1	4.0	7	124.2	--	1	110.8	10.8	77
25	111.0	10.6	44	104.2	9.4	33	115.6	6.1	8	116.4	1.3	2	109.0	10.4	87
30	110.1	8.7	40	108.0	7.4	46	107.5	11.8	23	106.8	9.1	7	108.5	8.9	116
35	107.0	11.1	41	108.8	12.0	40	109.9	9.6	39	108.5	8.3	7	108.5	10.8	127
40	109.2	9.2	37	106.2	9.6	44	112.1	10.2	47	109.3	10.2	8	109.2	9.9	136
45	109.2	8.4	28	103.6	9.3	43	111.4	12.2	44	113.0	11.8	25	108.9	11.1	140
50	110.4	8.5	26	110.9	10.1	40	113.4	10.3	54	106.9	10.2	17	111.3	10.0	137
55	115.2	10.2	22	109.0	8.9	21	114.7	12.1	53	112.5	10.7	38	113.3	11.0	134
60	113.5	9.9	19	109.8	9.6	35	118.5	10.8	49	112.1	11.8	44	113.9	11.2	147
65	114.3	9.1	18	116.1	7.4	23	121.7	10.1	22	115.6	11.2	43	116.7	10.1	106
70	116.7	9.3	7	120.0	8.5	12	121.0	8.8	21	119.1	7.9	36	119.5	8.3	76
75	118.8	5.1	8	116.4	5.0	6	121.8	6.0	9	119.0	8.1	19	119.2	6.8	42
80	127.3	1.2	3	118.9	7.4	9	134.7	6.0	4	126.2	6.5	11	125.1	8.1	27
85	119.7	10.1	4	117.9	10.1	3	126.2	5.3	6	121.1	7.9	9	121.8	8.0	22
90	120.9	--	1	124.0	--	1	127.0	8.5	2	124.7	9.0	5	124.7	7.3	9
95	--	--	--	119.0	--	1	131.4	11.2	5	126.0	5.7	2	128.5	9.8	8
100	--	--	--	124.9	4.9	2	--	--	--	123.0	--	1	124.2	3.6	3
105	--	--	--	--	--	--	--	--	--	122.7	--	1	122.7	--	1
110	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
115	--	--	--	--	--	--	--	--	--	115.0	--	1	115.0	--	1
120	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
GRAND	110.9	9.9	430	109.2	9.9	461	114.8	11.1	440	114.0	11.0	316			1647
Minimum	85			76			85			90					
Maximum	135			134			145			143					
Range	50			58			60			53					

Table 4. Mean, standard deviation (SD), and number of data points (N), for LDL at four frequencies (in dB SPL) measured by Pascoe (1988). Also provided are minimum, maximum, range at each frequency and grand mean, SD, and N.

HL	<u>500 Hz</u>			<u>1000 Hz</u>			<u>2000 Hz</u>			<u>4000 Hz</u>			<u>Grand</u>	<u>Grand</u>	<u>Grand</u>
	<u>Mean</u>	<u>SD</u>	<u>N</u>	<u>Mean</u>	<u>SD</u>	<u>N</u>	<u>Mean</u>	<u>SD</u>	<u>N</u>	<u>Mean</u>	<u>SD</u>	<u>N</u>	<u>Mean</u>	<u>SD</u>	<u>N</u>
0	107.4	9.9	14	110.7	6.8	6	--	--	--	--	--	--	109.1	8.4	20
5	109.7	8.7	22	110.3	6.4	12	--	--	--	--	--	--	110.0	8.1	34
10	111.1	7.3	39	109.3	6.9	15	125.0	--	1	103.0	--	1	112.1	7.3	56
15	109.5	6.8	57	108.1	6.5	32	115.0	--	1	113.0	--	1	111.4	6.7	91
20	107.0	7.8	45	109.0	7.1	26	110.0	2.9	2	104.7	2.9	3	107.7	7.7	76
25	112.4	8.9	40	111.4	8.5	40	115.0	9.6	5	113.0	0.0	2	113.0	8.6	87
30	114.5	8.4	54	109.8	8.1	40	115.5	9.1	11	119.7	12.6	3	114.9	8.4	108
35	109.9	8.7	57	112.3	7.4	51	118.3	7.3	24	123.0	--	1	115.9	8.4	133
40	112.2	9.7	45	111.1	9.5	63	119.6	9.0	52	118.0	7.1	15	115.2	9.3	175
45	118.6	9.5	32	113.2	8.9	63	120.6	8.9	63	119.3	11.4	19	117.9	9.3	177
50	117.3	9.8	33	117.2	9.2	47	122.7	9.1	84	120.8	10.2	34	119.5	9.4	198
55	118.7	8.5	27	117.4	7.8	44	122.6	8.3	82	120.8	8.9	66	119.9	8.4	219
60	120.2	6.1	17	117.3	6.5	26	125.6	7.1	78	123.3	8.4	67	121.6	7.5	188
65	120.3	7.5	6	121.9	6.8	22	130.2	8.0	30	126.9	9.3	58	124.8	8.5	116
70	115.8	4.8	4	123.3	6.7	7	130.5	8.4	22	128.4	7.7	61	124.5	8.0	94
75	128.3	6.4	8	131.5	10.6	2	131.9	7.2	24	129.6	5.7	57	130.3	6.2	91
80	127.0	5.8	4	129.0	10.0	3	134.7	4.6	16	133.5	6.0	33	131.1	5.9	56
85	137.0	--	1	126.5	4.2	6	135.7	6.7	7	133.2	6.5	22	133.1	6.1	36
90	142.0	--	1	136.5	3.5	2	135.0	--	1	136.3	5.6	26	137.5	5.6	30
95	132.0	--	1	--	--	--	--	--	--	143.5	7.2	10	137.8	7.6	11
100	--	--	--	--	--	--	--	--	--	139.6	6.3	16	139.6	6.3	16
105	--	--	--	--	--	--	150.0	7.1	2	145.5	4.2	6	147.8	4.6	8
110	--	--	--	144.0	--	1	151.7	2.9	3	144.7	7.6	3	146.8	5.4	7
115	--	--	--	--	--	--	--	--	--	149.7	5.8	3	149.7	5.8	3
120	--	--	--	--	--	--	--	--	--	153.0	--	1	153.0	--	1
GRAND	118.5	1.5	507	118.5	1.8	508	126.8	2.1	508	127.9	2.9	508			2031
Minimum	107			108			110			103					
Maximum	142			144			152			153					
Range	35			36			42			50					

adding the prescribed ANSI3.9 – 1989 REDD to mean LDL HL data. The mean LDL (dB SPL) and SD (dB SPL) values were as follows: 118.5 (7.9), 118.5 (7.4), 126.8 (7.3) and 127.9 (7.0) at 500, 1000, 2000, and 4000 Hz, respectively. The grand mean LDL as a function of hearing threshold was derived by averaging mean LDL (dB SPL) across the four frequencies. These values ranged from 107.68 (at 20 dB) – 153.00 dB SPL (at 120 dB).

Figure 6 reports the mean WUSM LDL in dB SPL for frequency 500, 1000, 2000, 4000 Hz. Hearing threshold (0 to 120) is along the abscissa and mean LDL SPL is along the ordinate. The solid line represents the hearing threshold level (in dB HL). The solid line with the squares illustrates the grand mean. Note how the frequencies do not significantly depart from this grand mean, except in cases where a small number of measurements are recorded. This is in agreement with Pascoe's (1988) findings, which are replicated in **Figure 7**.

Figure 8 reveals the scatterplot, Y intercept, regression coefficients, number of data points, and the line of best fit for 500, 1000, 2000, and 4000 Hz for dB SPL. For each scatterplot, hearing threshold is along the abscissa and LDL dB SPL is along the ordinate. The diamonds represent the individually measured LDL dB SPL as a function of hearing threshold. Since subjects with the same hearing threshold and LDL are reported as one dot, the number of dots does not accurately represent the recorded data. The LDL in dB SPL can also be predicted from hearing threshold by using the formula:

$$Y = b(X) + a$$

Where Y is the predicted LDL in dB SPL, b is the regression coefficient, and X is the measured hearing threshold (in dB HL) and a is the Y intercept. For example, using the data in Figure 8 at 1000 Hz, the predicted LDL (dB SPL) would be $0.1187(40) + 104.57$ or 109.32 dB SPL (Valente et al, 1997). The Pearson product-moment correlations (r), which were used to derive the R²

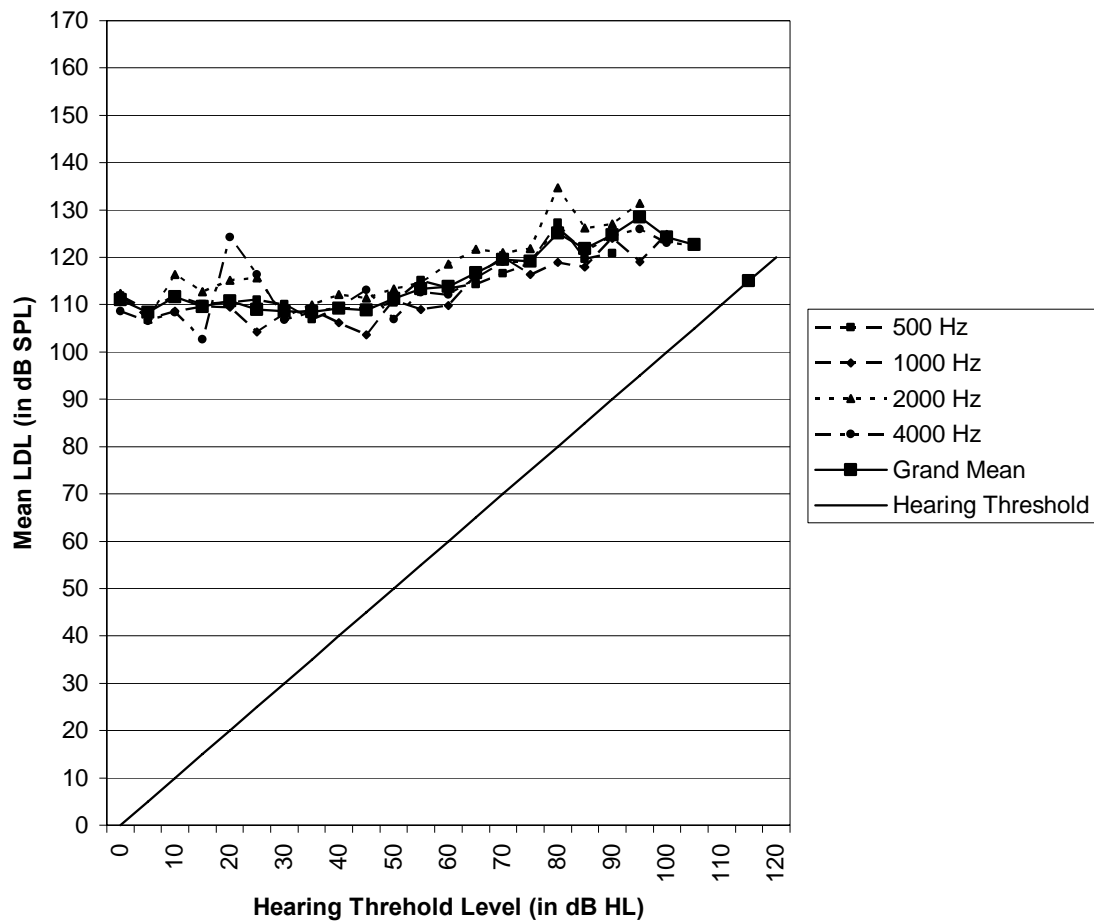


Figure 6. WUSM mean LDL (dB SPL) as a function of frequency. Also provided is the grand mean and hearing threshold.

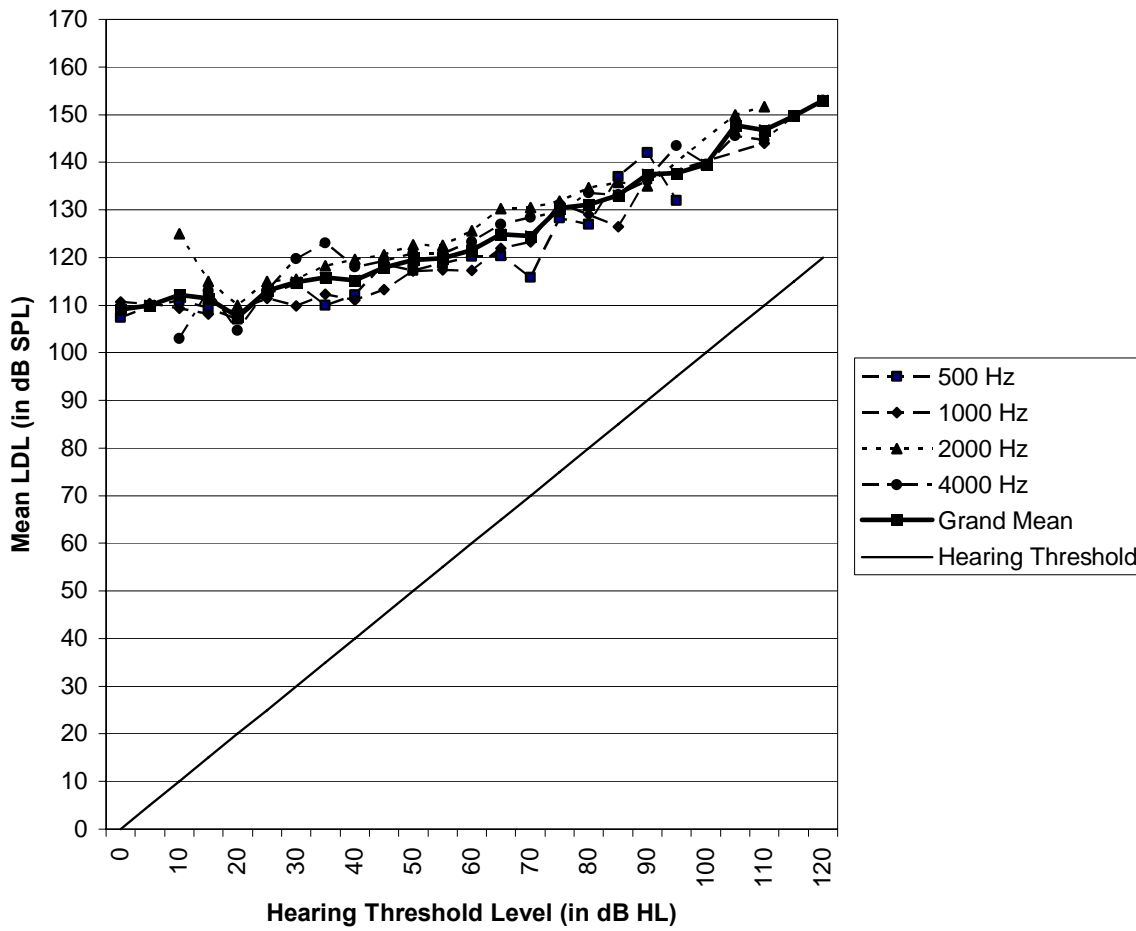


Figure 7. Pascoe (1988) mean LDL (dB SPL) as a function of frequency. Also provided is the grand mean and hearing threshold.

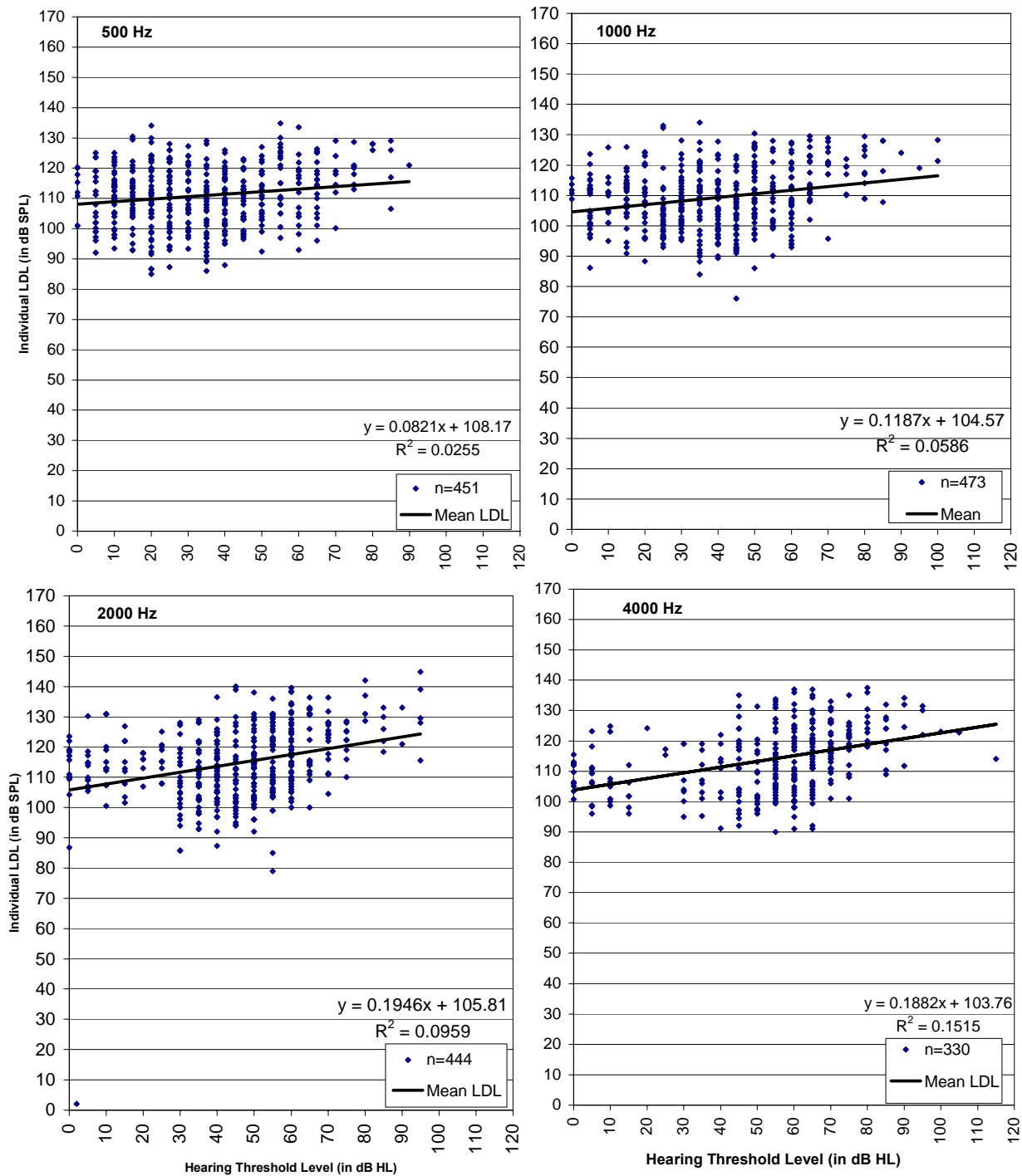


Figure 8. Scatterplots at 500 - 4000 Hz for WUSM LDLs (in dB SPL) as a function of hearing level. Also provided is the line of best fit (mean LDL), number of data points (N), R^2 values and the equation for the line of best fit

values, ranged from .161 at 500 Hz to .389 at 4000 Hz. However, for 2000 and 4000 Hz, correlations increase, and fall into the mildly correlated range ($r=.295$ and $.389$ respectively). These correlations, as also reported with LDL HL measurements, signify that threshold measurements are weak but better predictors of LDL SPL at higher frequencies.

A two-factor repeated measures ANOVA (LDL dB SPL \times frequency) revealed a significant two-factor interaction ($F = 27.479$; $df = 3/1693$; $p < .001$), indicating that significant differences were present in the mean LDLs for at least one of the test frequencies. Post-hoc analysis (Bonferroni/Dunn test of multiple pair-wise comparisons) indicated that the main effect of frequency was driven by significant differences between 500 and 2000 Hz ($p < .001$), 500 and 4000 Hz ($p < .001$), 1000 and 2000 Hz ($p < .001$), and 1000 and 4000 Hz ($p < .001$). Results (the mean difference and critical difference in parenthesis) are as follows: -3.9 (1.8), -2.9 (2.0), -5.6 (1.8) and -4.7 (2.0) for 500 and 2000 Hz, 500 and 4000 Hz, 1000 and 2000 Hz, and 1000 and 4000 Hz, respectively. Therefore, there is no significant difference between the lower frequencies (500 and 1000 Hz) or the higher frequencies (2000 and 4000 Hz).

To illustrate the findings from **Table 4, Figure 9** superimposes Pascoe's (1988) data with WUSM mean LDL (in dB SPL) as a function of hearing threshold for 500, 1000, 2000, and 4000. **Figure 10** superimposes grand mean LDL SPL and standard deviation bars from Pascoe (1988) and WUSM. The squares represent the mean WUSM LDL SPL, and the triangles represent the mean Pascoe (1988) LDL SPL. Standard deviation bars report ± 2 standard deviations (SD), or the point for which 95% of all data can be accounted. For hearing thresholds where less than two data points was reported, standard deviation bars are not reported.

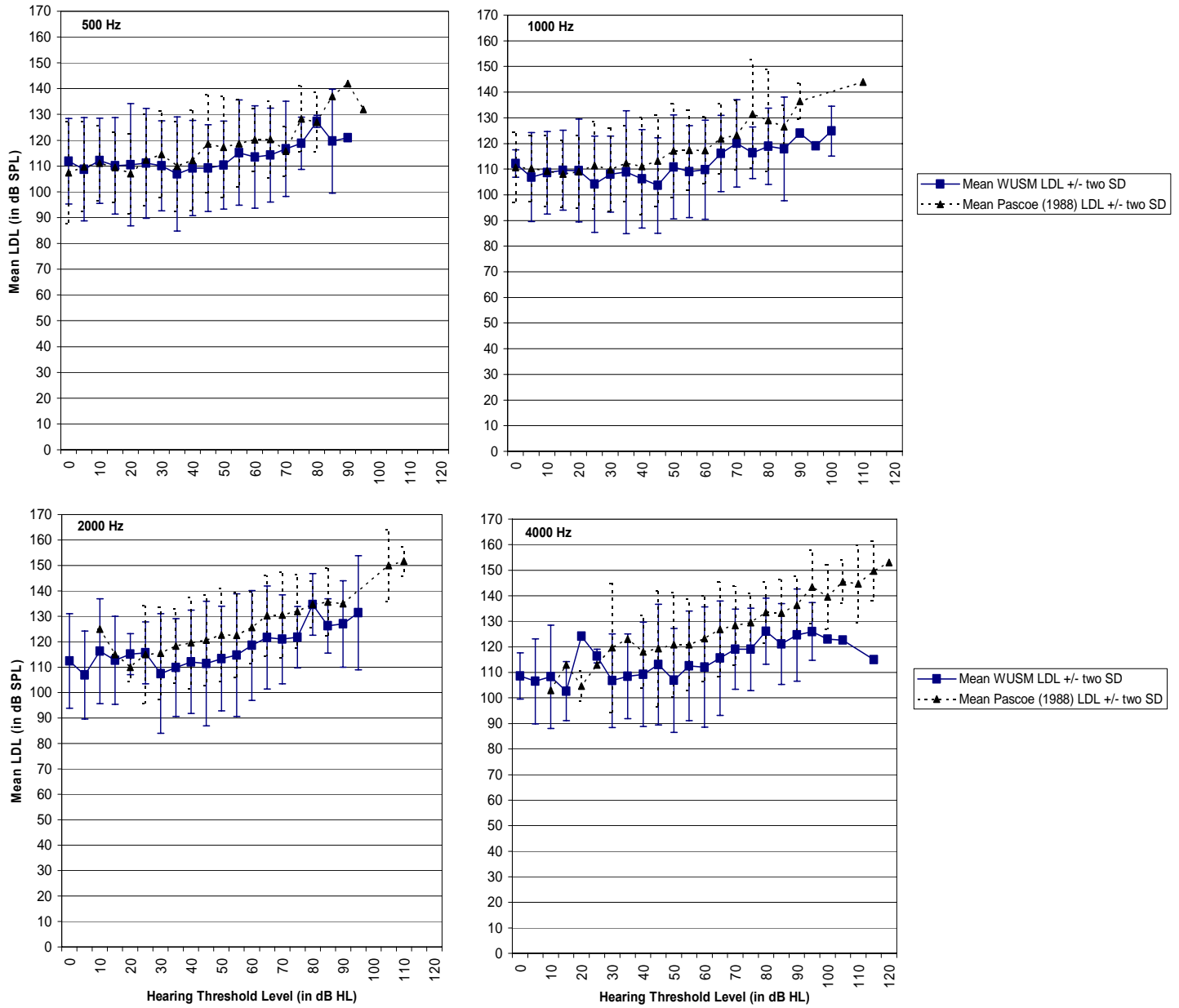


Figure 9. Comparison of mean LDL (in dB SPL) as a function of frequency for Pascoe (1988) and WUSM data. Error bars represent +/- 2 SD.

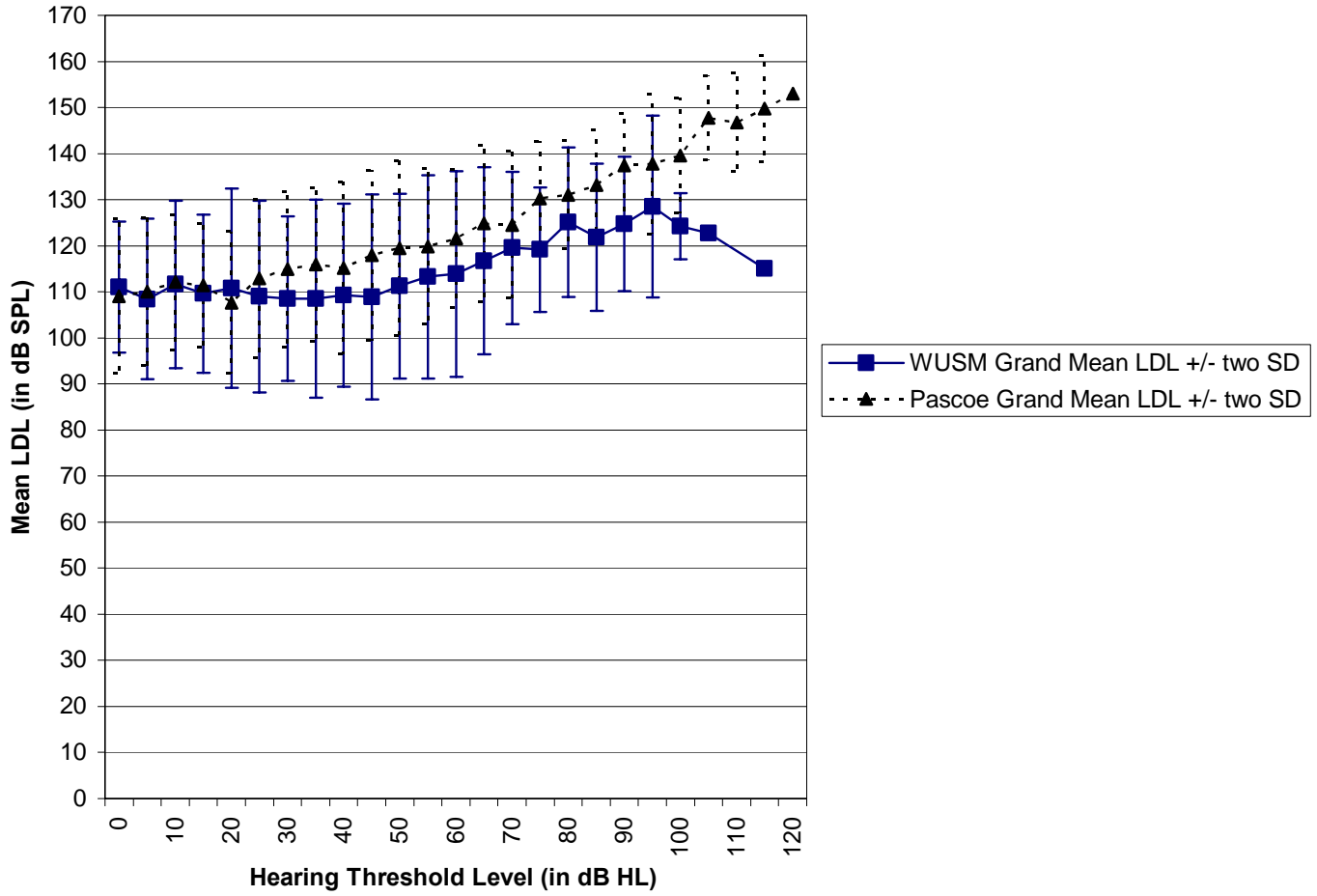


Figure 10. Comparison of Grand Mean WUSM LDL (dB SPL): re Pascoe's Data on Hearing Threshold Levels. Error bars represent +/- 2 SD.

The reader may notice the increased difference between WUSM and Pascoe mean LDL as hearing threshold becomes poorer. Generally, the Pascoe (1988) mean LDL departs from WUSM mean LDL data at approximately 30 dB (HL) for each frequency (500, 1000, 2000, and 4000 Hz) (**Figure 9**) and the grand mean (**Figure 10**). This difference increases as hearing loss decreases with increasing hearing threshold, with the Pascoe (1988) mean LDLs greater than the mean WUSM LDLs.

Comparisons between Pascoe (1988) and WUSM were analyzed via repeated measures ANOVA (site \times frequency). In general, the main effect for site indicated a significant difference between the two sets of data ($F = 31.601$; $df = 1/157$; $p < .01$). There was not a main effect for frequency ($F = 5.680$; $df = 3/157$; $p = .001$) or an interaction between frequency and test site ($F = 1/530$; $df = 3/157$; $p = .209$).

Intersubject variability

The first experimental question addressed earlier in this paper was whether or not LDL can be accurately predicted from threshold. The intersubject variability is the range of threshold that fall above or below the line of best fit. A small degree of variability would indicate that LDL could be accurately predicted from an individual's hearing threshold. As the intersubject variability range increases, the ability to predict the LDL from the hearing threshold decreases.

Figure 3 and Figure 8 illustrates the large intersubject variability in WUSM data. This range could be as small as small as 10 dB or as large as to 59 dB depending upon frequency (500, 1000, 2000 or 4000 Hz) and measurement method (dB HL or dB SPL). This intersubject variability was further illustrated by **Figures 3 and 8**, which revealed scatterplots, Y intercept (a), regression coefficient (b), and the line of best fit for 500 to 4000 Hz for WUSM data in dB

HL and dB SPL, respectively. The line of best fit represents the “predicted” LDL in dB (HL) if only hearing threshold was known. However, this line only accounts for 8.6% of the data in dB HL and 8.9% of the data in dB SPL, due to the large intersubject variability. For example, in **Figure 8** (at 1000 Hz), the LDL for a threshold of 45 dB HL could range from 76 – 123 dB SPL. This variability exists for other hearing levels and frequencies.

Simple linear regression computed the line of best fit for the prediction of LDL in dB (HL) from hearing threshold. This equation revealed that the correlation between hearing threshold and LDL in dB (HL) is weak, but significant ($r = .293$; $p < .001$; Fisher’s r to z). As noted in **Figures 3 and 8**, the mean line of best fit increases with frequency, indicating that mean LDL is greater as frequency and hearing level increases. This finding is in agreement with Kamm et al (1978), Dillon et al (1984), Hawkins et al (1987), and Pascoe (1988). Each of these studies reported the dangers of predicting the LDL (in dB HL) from threshold measurements due to the presence of large intersubject variability. WUSM includes that predicting the LDL SPL could also be detrimental.

REDD

WUSM

The individually measured REDD at 500, 1000, 2000 and 4000 Hz was computed by the following formula: LDL SPL – from the LDL measured in dB HL. **Figure 11** reports the mean REDD as a function of frequency (500, 1000, 2000, and 4000 Hz). For each scatterplot the four test frequencies are along the abscissa and REDD is along the ordinate. The dots represent measured REDD as a function of frequency. The mean LDL (dB SPL) and SD (dB SPL) values were as follows: 12.2 (4.8), 9.7 (4.6), 15.1 (4.9), and 11.4 (6.1) for 500, 1000, 2000 and 4000 Hz respectively. Overall, the intersubject variability

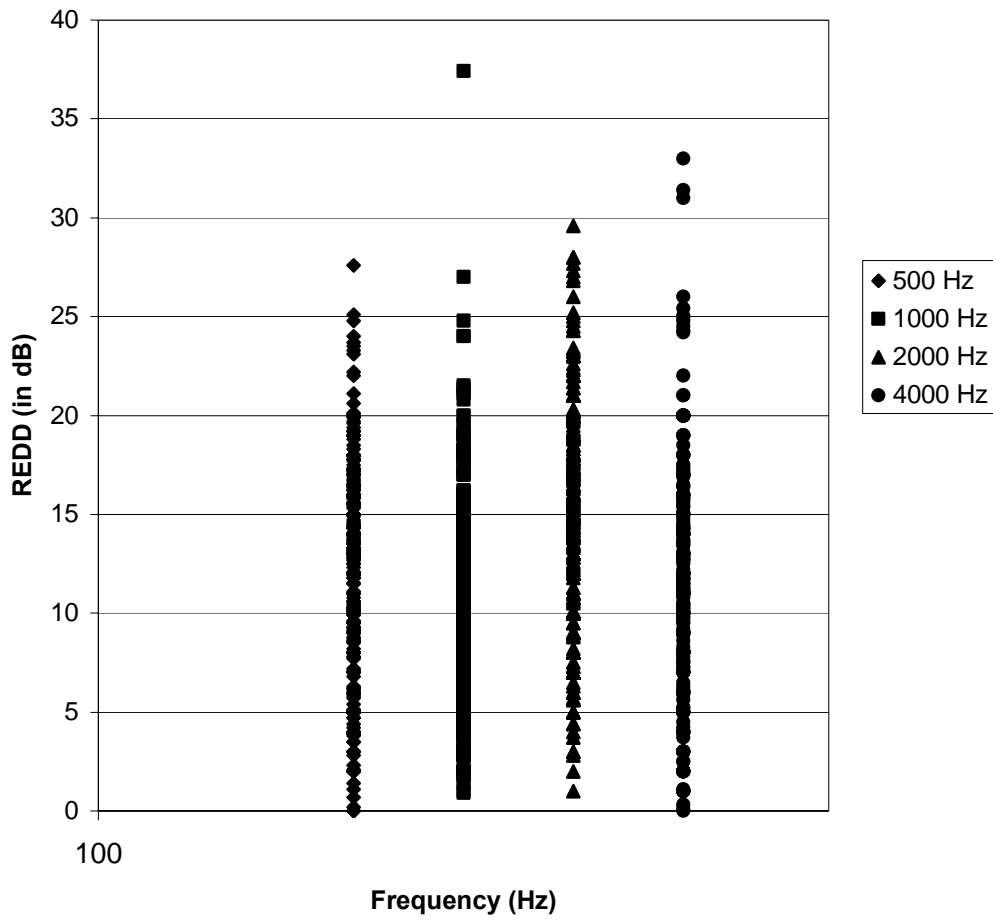


Figure 11. REDD as a function of frequency (WUSM)

ranges from 27.6 (at 500 Hz) to 37.4 (at 1000 Hz). However, the greatest intersubject variability (reported at 1000 Hz) is due to an outlier. If this outlier is excluded from data, the intersubject variability would be 27.0 dB at 1000 Hz.

DISCUSSION

Mean Differences between Pascoe (1988) and WUSM data

This study determined that the mean LDL measured at WUSM (ie., the IHAF procedure) is significantly different than the mean LDL reported by Pascoe (1988) for hearing losses greater than 30 dB (HL) and therefore suggests that utilizing the mean Pascoe (1988) LDL could result in a greater output. This could lead to overamplification and to possible hearing aid rejection, regardless of whether a prescriptive fitting method or real ear measurement (REM) system is used to verify the hearing aid fit. These results should act to caution audiologists against using the predicted LDL derived from Pascoe (1988) to predict individual LDL in dB HL or dB SPL. Overall, these findings support results from the Kochkin (2000) survey, which found that the overamplification of environmental sounds is the third most common reason for hearing aid dissatisfaction and Kochkin (2002), which reported that only 59% of those tested were satisfied with their hearing aids (re: loudness).

The author feels that the difference between Pascoe (1988) and WUSM that were reported (**Figures 4, 5, 9 and 10**) are most likely due to differences in the descriptive anchors used. Recall that the incongruity of descriptive anchors is obvious by reviewing Pascoe (1988) and the IHAF protocol (Valente and Vliet, 1997). The IHAF protocol defines the LDL as the level at which input signals are ‘loud, but ok.’ Pascoe (1988), on the other hand, defined the LDL as the level at which a subject reported that the input signal became “too loud.” Therefore, it is possible, that by using the data reported by Pascoe (1988), the resulting output could be

judged uncomfortably loud by a patient as opposed to adjusting the output to a patient's judgment of when incoming signals as "loud, but ok." The reported difference between Pascoe (1988) and WUSM LDL (ie., IHAF procedure), signifies that utilizing the Pascoe (1988) LDL data could result in a louder output, which, in turn, could lead to judgments that amplification is "too loud."

Second, variability could exist due to the extrapolation of data. For data greater than 120 dB HL, Pascoe (1988) assumed a linear increase of 5 dB for each remaining loudness category above 120 dB HL. For example, if a patient rated 120 dB HL as "OK" (number 5), 5 dB HL would be added for each remaining category, which would result in an LDL of 140 dB HL. WUSM does not extrapolate data. While this insured accuracy of data, one disadvantage was that data was rejected from if the descriptive anchor of 'loud, but ok,' could not be reported before the limits of the audiometer was reached. This created a greater amount of measures being excluded as hearing threshold increased. Overall, this may have led to the measured LDLs, for elevated thresholds, that were artificially low.

Even though reasons exist for the differences between the Pascoe (1988) and WUSM, Dillon and Storey (1998) also report that Pascoe's (1988) data leads to greater measures of LDL. Overall, by comparing their data to five other studies (for 500, 1000 and 2000 Hz) they concluded that utilizing Pascoe (1988) data led to mean LDLs 4 to 8 dB greater than the mean LDL reported by Pascoe (1988).

Intersubject Variability

Finally, WUSM reports a large intersubject variability for the measured LDL and REDD data. This agrees with findings from previous studies, which also reported that LDL cannot be accurately predicted from threshold (Kamm et al, 1978; Dillon et al, 1984; Hawkins et al, 1987;

Pascoe, 1988; Sammeth et al, 1993; Valente et al, 1997). This intersubject variability should act to caution clinicians of the possible error that could be introduced by using mean, normalized data to derive threshold and loudness discomfort measures for an individual. To address how LDL intersubject variability reported by WUSM LDL differs from previous studies, the author compared the WUSM data with Eberling's (1999) study, against two major studies: Bentler and Cooley (2001) and WUSM data. Recall that Eberling (1999) concluded that for 70% of the hearing impaired, the LDL could be predicted from hearing threshold within +/- 5 dB. He defined the remaining 30% as outliers that consist of "sound sensitive," (12%; lower LDL than predicted) and "sound addicts," (17%; higher LDL than predicted). A study by Mueller and Bentler (2005) reported that when using the +/- 5 dB criterion, less than 20% of Bentler and Cooley's (2001) data fell into this range. This finding would leave 80% of hearing aid users with a less than satisfactory output.

Table 5 reports the percentage of data that falls below -5dB, within +/-5 dB, and above +5 dB of the mean LDL for each of the four frequencies (500, 1000, 2000, and 4000 Hz). The percentage of data that fell below -5dB of the mean were: 500 Hz = 26, 25; 1000 Hz = 20, 23; 2000 Hz = 22, 27; 4000 Hz = 18, 27 for dB (HL) and dB (SPL), respectively. The percentage of data that fell within +/-5 dB of the mean was: 500 Hz = 45, 45; 1000 Hz = 60, 50; 2000 Hz = 56, 44; 4000 Hz = 62, 50 for dB (HL) and dB (SPL), respectively. The percentage of data that fell within above +5 dB of the mean were: 500 Hz = 45, 45; 1000 Hz = 60, 50; 2000 Hz = 56, 44; 4000 Hz = 62, 50 for dB (HL) and dB (SPL), respectively. **Figures 12** and **13**, which graphically reports the data from **Table 5**, reveals scatterplot, Y intercept, regression coefficients, number of data points for WUSM, and the line of best fit for 500, 10000, 2000, and 4000 Hz for dB HL (**Figure 13**) and dB SPL (**Figure 14**), respectively. For each scatterplot,

hearing threshold is along the abscissa and LDL is along the ordinate. The diamonds represent the individually measured LDL as a function of hearing threshold. The diamonds above the top, thin line represent data that fell above +5 dB from the mean. The diamonds between the two thin lines represent data that fell ± 5 dB. Lastly, the diamonds below the bottom, thin line is the data that fell -5 dB from the mean.

The overall 'sound sensitive' group (i.e., data points below the -5 dB line) increased to 22% (dB HL) and 26% (dB SPL) and the 'sound addict' group (i.e., data point below the +5 dB line) increased to 23% (dB HL) and 26% (dB SPL). Therefore, the data from WUSM is in opposition to that reported by Eberling (1999). That is approximately 50% of subjects fall *outside* of ± 5 dB. The author believes this, in agreement with Bentler and Cooley's (2001) finding of 80% ± 5 dB of the mean, reiterates the importance of measuring individual LDL. This finding is interesting, considering that Kochkin (2002) found that patient satisfaction was only 42% for 'comfort with loud sounds.' Could the remaining 52% of the respondents in Kochkin (2002) fall outside of ± 5 dB from the mean LDL?

Table 5. Percentage of data points that were above +5 dB of the mean LDL, below -5 dB and within +/-5 dB for dB (HL) and dB (SPL) for each frequency. Also included is the grand percentages.

	dB (HL)			dB (SPL)		
	% below 5 dB	% within +/-5 dB	% above +5 dB	% below 5 dB	% within +/-5 dB	% above +5 dB
500	0.26	0.45	0.29	0.25	0.45	0.30
1000	0.20	0.60	0.20	0.23	0.53	0.24
2000	0.22	0.56	0.22	0.27	0.44	0.29
4000	0.18	0.62	0.20	0.27	0.50	0.23
Grand	0.22	0.56	0.23	0.26	0.48	0.26

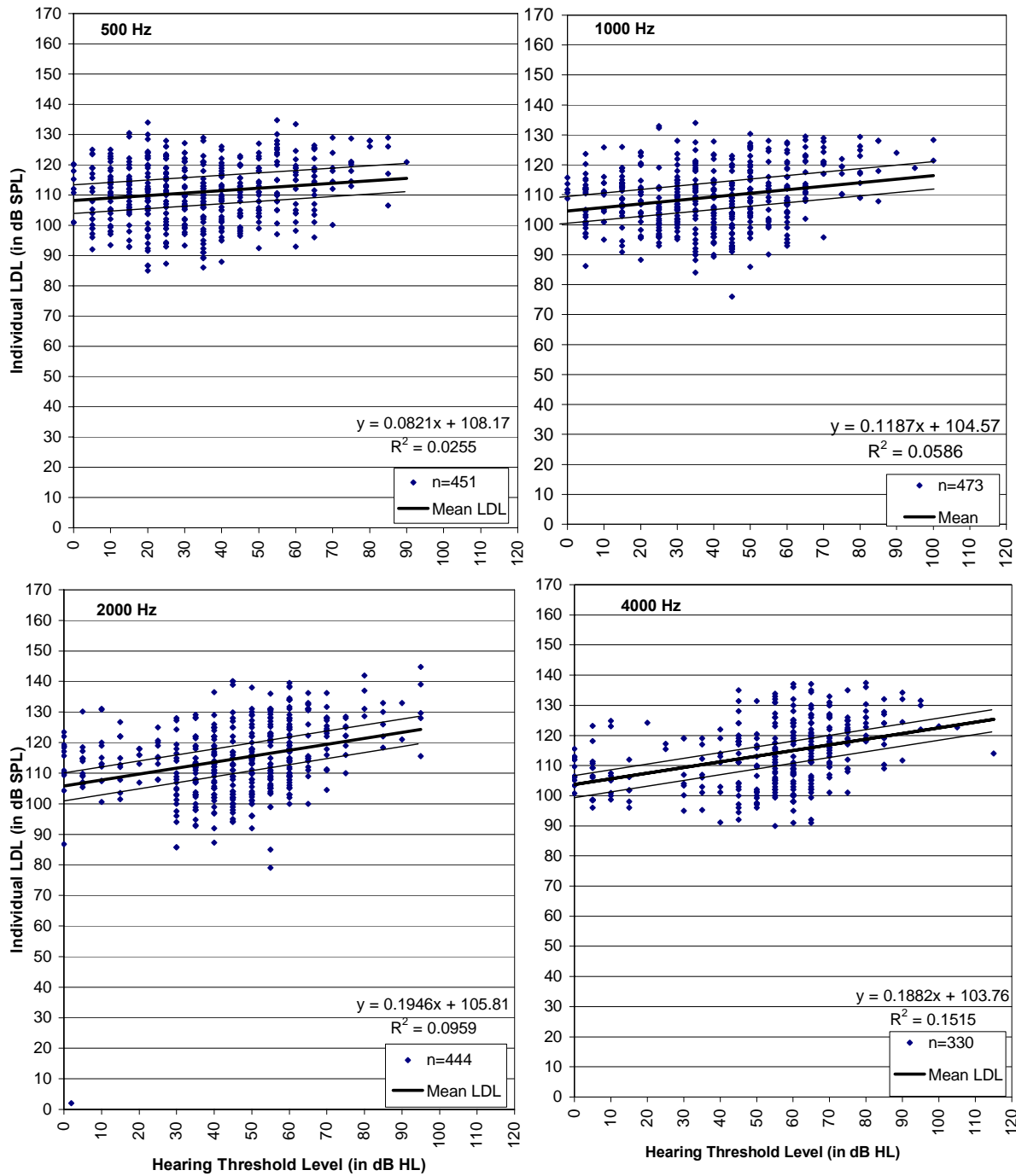


Figure 12. Scatterplots at 500 - 4000 Hz for WUSM LDLs (in dB SPL) as a function of hearing level. Also provided is the line of best fit (mean LDL), number of data points (N), R2 value and the equation for the line of best fit. The two thinner, solid lines represent +/- 5 dB above and below the mean.

Comparison of Studies

Figure 14 reports the mean REDD for 500 to 4000 Hz from four sources: WUSM; ANSI S3.6 – 1989; Valente et al (1997), and NAL-NL1. The mean REDD for the ANSI S3.6 – 1989 standards were as follows: 500 Hz = 12.0; 1000 Hz = 9.0; 2000 Hz = 15.0; 4000 Hz = 13.0 dB. The mean REDD for NAL-NL1 were as follows: 500 Hz = 11.7; 1000 Hz = 9.5; 2000 Hz = 15.9; 4000 Hz = 13.4 dB. Finally, the mean REDD from Valente et al (1997) were: 500 Hz = 11.8; 1000 Hz = 10.1; 2000 Hz = 16.8; 4000 Hz = 10.7 dB. Since individual REDD was not included in Valente et al (1997), the mean REDD was computed from the mean LDL dB (SPL) and (HL) using the ER-3A earphone (N=31).

How accurate are the REDD predictions from other sources? Differences between the WUSM and the ANSI S3.6 – 1989 REDD data ranged from 0.38 dB at 500 Hz to -1.67 dB at 4000 Hz. This data was then compared to the NAL-NL1 mean REDD transformation. Although WUSM and NAL-NL1 were similar at 500 – 2000 Hz, the mean WUSM REDD was lower by 2.7 dB at 4000 Hz. The smaller number of subjects for the 4000 Hz measurements could account for this decrease. Overall, the greatest variability between the four studies existed at 4000 Hz between Valente et al (1997) (10.7 dB) and NAL-NL1 (13.4 dB). This author believes that a mean difference of 2.7 is not clinically significant. Previously, the author reported that predicting the LDL from hearing threshold can lead to a large number of patients being inaccurately fit. Even though NAL-NL1, ANSI S3.6 - 1989, Valente et al (1997) and WUSM mean data vary a maximum of 2.7 dB, it cannot be assumed that predicted REDD will closely match the measured REDD (i.e., **Figure 14**).

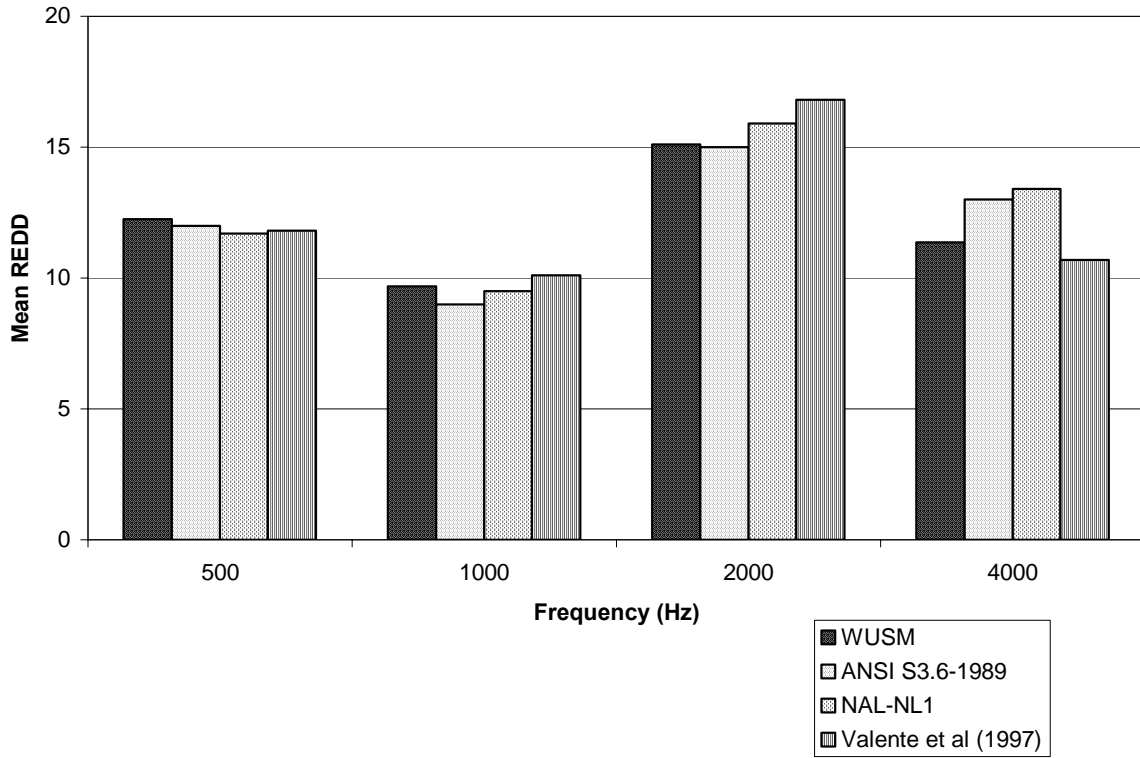


Figure 14: Comparison of mean REDD as a function of frequency (500 - 4000 Hz) as predicted by: WUSM, ANSI S3.6-1989, NAL-NL1, and Valente et al (1997).

On the contrary, although the intersubject variability for WUSM REDD is less than the range for LDL measurements, it is still significant. Although the average REDD relates closely with REDD findings from other reports, it must be assumed (since the actual data for the sources was not available), that individual LDL measurements varied as greatly as WUSM REDD data.

As mentioned previously, the Audioscan Verifit, Speechmap™ program utilizes live speech mapping (LSM) to verify hearing aid fittings. In this system, once threshold (dB HL) is entered, the software calculates the predicted LDL (dB HL) as well as predicted threshold and LDL (in dB HL) by applying the ANSI S3.6 – 1989 average REDD conversions. The LDL dB (SPL) can be measured two ways. One method is adding ANSI S3.6 – 1989 REDD conversions to the measured LDL (in dB HL). If the individual LDL dB HL is not measured, Verifit adds the ANSI S3.6 – 1989 REDD conversions to one standard deviation below Pascoe (1988) mean LDL HL. Finally, to record the predicted hearing threshold levels (in dB SPL), the software adds the ANSI S3.6 – 1989 REDD values (500 Hz = 13 dB, 1000 Hz = 10 dB, 2000 Hz = 15 dB, 4000 Hz = 15 dB) to measured hearing threshold (in dB HL).

The solid upper line (◆—◆) of **Figure 15A** represents the measured hearing threshold in dB HL for Subject 1. The solid lower line (■—■) represents the measured LDL in dB HL. The lower dashed line (▲--▲) is the predicted LDL (one SD below the mean LDL Pascoe (1988)) (in dB HL). Notice the close agreement between the predicted and measured LDL. In **Figure 15B**, the predicted threshold (dB SPL) is the dashed lower line (◆--◆). The solid line (■—■) is the measured LDL in dB SPL. The dashed upper line (▲--▲) is the predicted LDL in dB SPL. As can be seen, the agreement between the measured and predicted LDLs (in dB SPL) is quite close, although not as close as the predicted and measured values in dB HL.

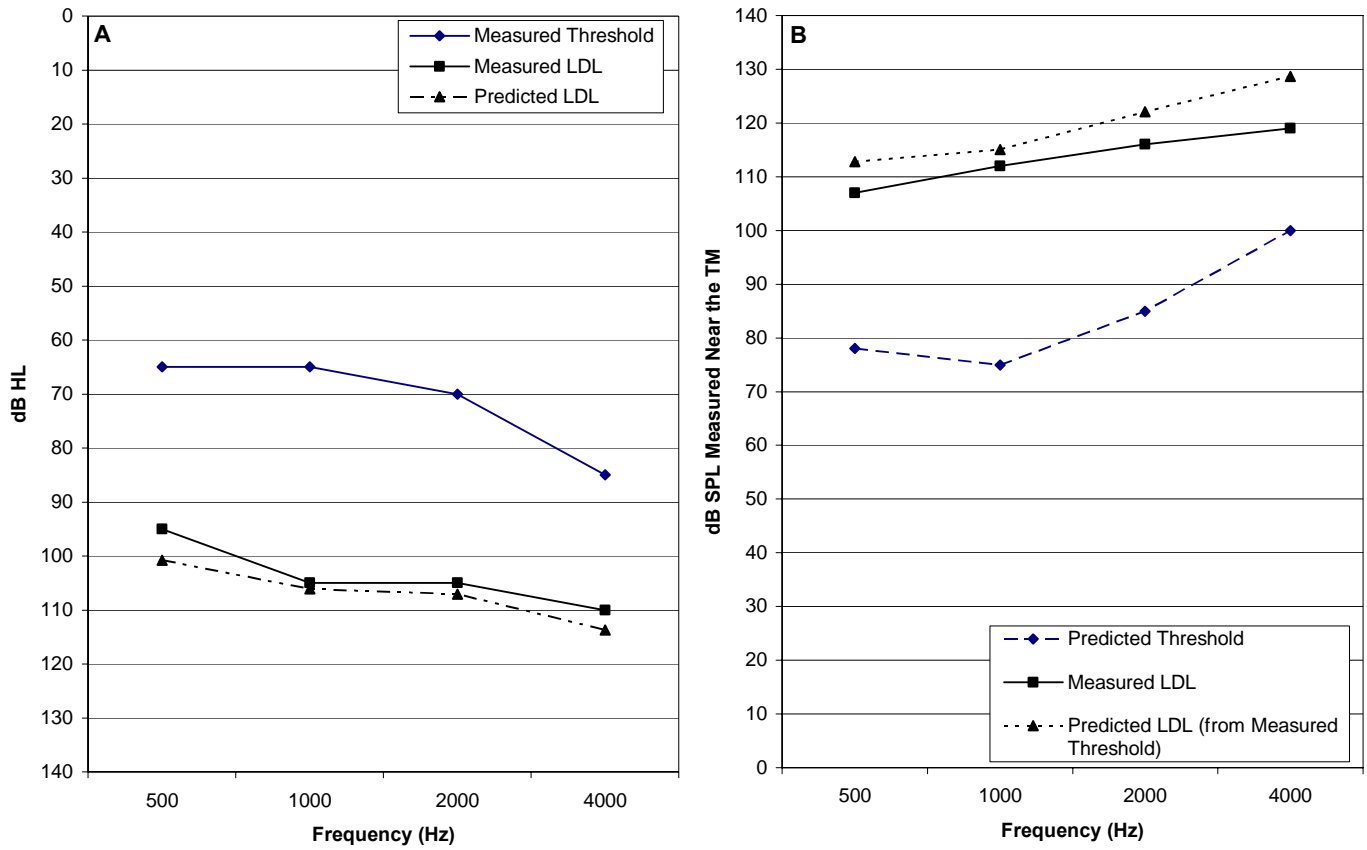


Figure 15. A) Measured (■-■) and predicted LDL (▲-▲) predicted from threshold (◆-◆) using AudioScan Speechmap™ software for Subject 1 (in dB HL). A) Measured (■-■) and predicted LDL (▲-▲) predicted from threshold (◆-◆) using AudioScan Speechmap™ software for Subject 1 (in dB SPL).

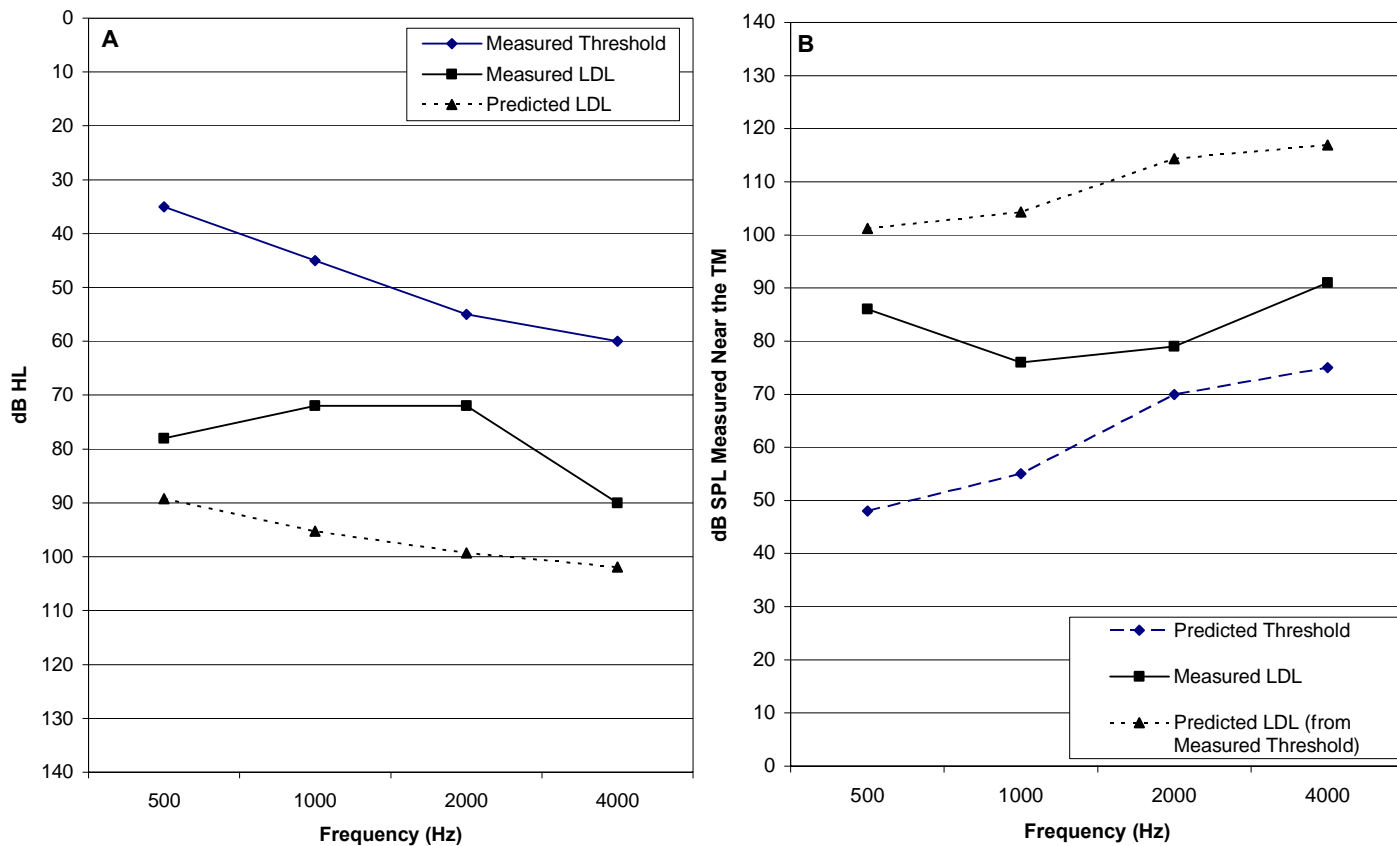


Figure 16. A) Measured (■-■) and predicted LDL (▲-▲) predicted from threshold (◆-◆) using AudioScan Speechmap™ software for Subject 2 (in dB HL). A) Measured (■-■) and predicted LDL (▲-▲) predicted from threshold (◆-◆) using AudioScan Speechmap™ software for Subject 2 (in dB SPL).

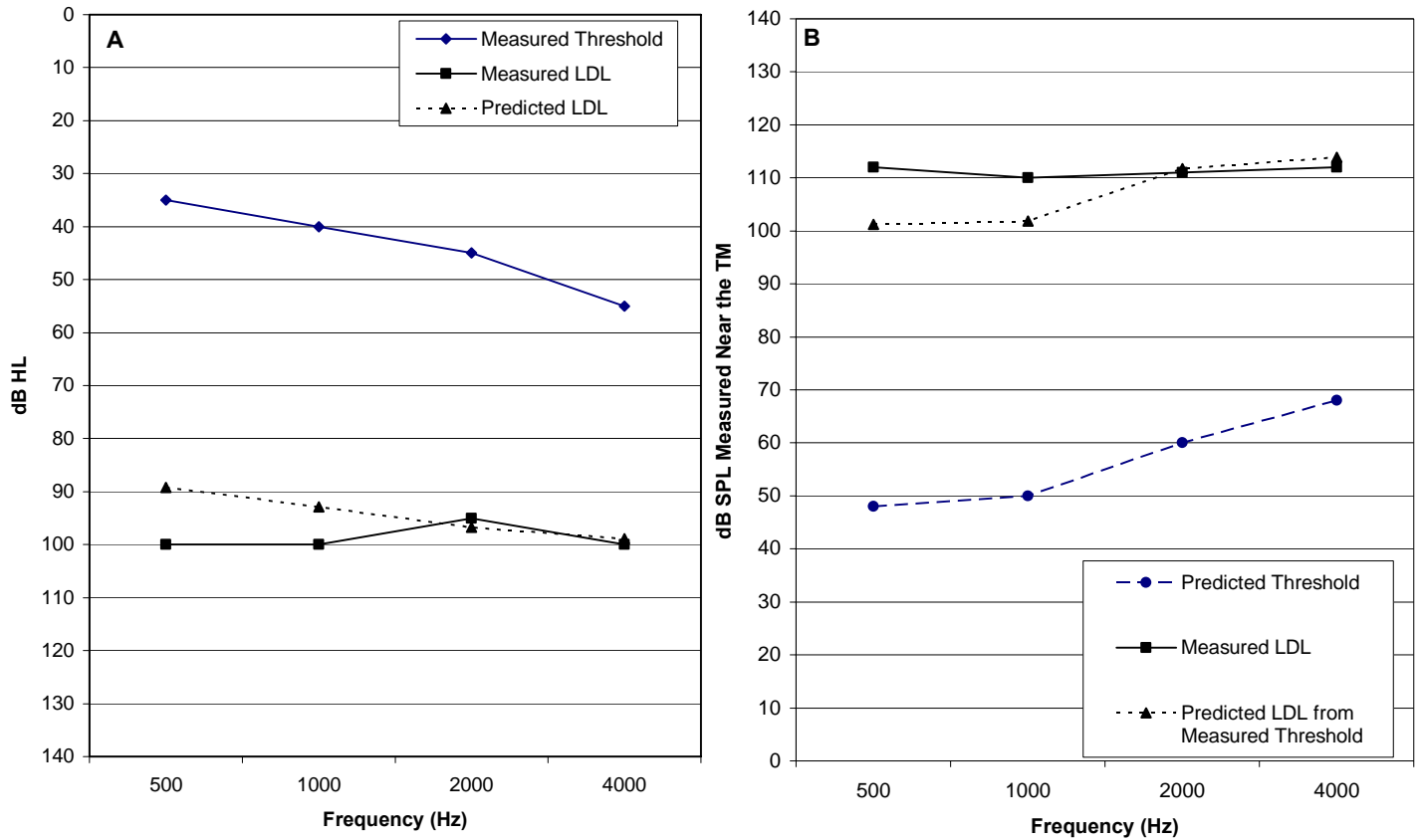


Figure 17. A) Measured (■-■) and predicted LDL (▲-▲) predicted from threshold (◆-◆) using AudioScan Speechmap™ software for WUSM mean LDL (in dB HL). A) Measured (■-■) and predicted LDL (▲-▲) predicted from threshold (◆-◆) using AudioScan Speechmap™ software for WUSM mean LDL (in dB SPL).

Using another subject, the solid upper line (◆—◆) in **Figure 16A** represents the measured hearing threshold (in dB HL) for Subject 2. The solid lower line (■—■) represents the measured LDL in dB HL. The lower dashed line (▲--▲) is the predicted LDL (dB HL). In this case, the agreement between the predicted and measured LDL values is quite poor. This difference, between the predicted LDL (HL) and the measured LDL HL, increases from 11.2 (at 500 Hz) to 27.3 dB HL (at 2000 Hz). This is due to the subject having a significantly reduced DR. Using the predicted LDL could result in a hearing aid fitting that would allow input levels to exceed the measured LDL by as much as 27 dB at some frequencies. This could lead to an intolerably loud fit and lead to the rejection of amplification as well as cause further hearing loss.

In **Figure 16B**, the predicted threshold (dB SPL) is the dashed lower line (◆--◆). The solid line (■—■) is the measured LDL in dB SPL. The dashed upper line (▲--▲) is the predicted LDL in dB (SPL). As can be seen, the agreement between the measured and predicted LDLs is poorer than that recorded for dB (HL). This difference between the predicted LDL (SPL) and the measured LDL SPL range increased to 15.2 (at 500 Hz) to 35.3 dB SPL (at 2000 Hz). This is because not only did the subject have a lower than predicted LDL (HL), but also had lower than predicted REDDs. Suppose the same patient was fit with a hearing aid using LDL predictions in dB (SPL). Similar to **Figure 16A**, this hearing aid fitting would allow input levels that could exceed the measured LDL to reach the patient's ear. However, using this prediction would allow input levels that are 35 dB above the measured LDL to reach the patient's ear at some frequencies! Overall, this represents one of many subjects that could experience discomfort with loud stimuli due to an OSPL90 that is set higher than LDL in dB (HL) or dB (SPL).

Finally, the solid upper line (◆-◆) on the left side of **Figure 17** represents the mean measured threshold in dB (HL) for WUSM mean data. The solid lower line (■—■) represents the mean measured LDL (HL). The lower dashed line (▲--▲) is the mean predicted LDL (HL) based on the average hearing threshold (dB HL). The agreement between the mean measured and predicted LDL is good for the high frequencies (2000 and 4000 Hz). On the right side of **Figure 17**, the predicted threshold is the dashed lower line (◆-◆). The solid line (■—■) is the measured LDL (SPL). The dashed upper line (▲--▲) is the predicted LDL in dB (SPL). Again, the agreement between the measured and predicted LDLs is good for the high frequencies (2000 and 4000 Hz); and, as seen in previous examples, the largest difference is between the predicted LDL (SPL) from hearing threshold and the measured LDL (SPL).

The three examples cited above reiterate the variability and significant errors that can arise from utilizing predicted measurements. WUSM mean data was graphed (**Fig. 17**) to visually report differences between Pascoe (1988) mean data and this study. However, the Audioscan Verifit, Speechmap™ program predicts LDLs at one standard deviation below Pascoe (1988). Therefore, it can be assumed that the differences between WUSM mean LDL and predicted values would be greater on real-ear equipment utilizing mean Pascoe (1988) data (Frye and MedRx). Since this study has reported that utilizing Pascoe (1988) data could result in a louder output, and thus lead to the rejection of amplification, some audiologists may be convinced that only the Audioscan Verifit, Speechmap™ program should be used to verify a hearing aid fitting if predicted values are to be used. However, there are benefits to using LSM from other real-ear manufacturers. For instance, the Frye 7000 allows the user to manually enter the individual REDD and LDL.

Since the large intersubject variability found in this study, and more specifically, these cases, reveals that, at any frequency, incorporating mean REDD to dB HL measurements to predict hearing threshold and LDL in dB SPL could lead to either over or underestimation, recording threshold levels and LDL in dB SPL should be considered the best case scenario. Since no transformations must be applied, these measurements are the most accurate. If these measurements can not be obtained, the second best option would be to measure individual LDL HL to reduce variability. Even though REDD transformations would be applied, the predicted LDL (in dB SPL) would be more precise, since Pascoe (1988) mean data would not be used. This study can not support predicting LDL from hearing threshold.

Conclusion

In conclusion, this study answered the remaining experimental questions:

1. The large intersubject variability reported by WUSM relates closely with findings reported by other studies (Kamm et al, 1978; Pascoe 1988; Eberling 1999; Bentler and Cooley, 2001). This large intersubject variability should act to caution clinicians of the possible error that could be introduced by using mean, normalized data to derive individual threshold and loudness discomfort measures.
2. The large intersubject variability and the weak correlation between hearing threshold and LDL (in dB HL and dB SPL) questions the validity of accurately using predicted LDL values. Overall, this study supports findings from Valente et al (1994) which states that the only situation when predicting individual performance from average group data is appropriate is in the case of evaluating children or the difficult-to-test population, where suprathreshold measurements are not always possible.

3. Although frequencies do not significantly depart from this grand mean, statistical analysis revealed that the mean LDL is frequency dependent. There is no statistically significant difference between mean LDL dB HL measured at 500, 1000 and 2000 Hz, but the mean LDL dB HL at 4000 Hz was statistically significantly different from the mean LDL dB HL measured at 500, 1000 and 2000 Hz. There is no significant difference between the mean LDL dB SPL measured at lower frequencies (500 and 1000 Hz) or the higher frequencies (2000 and 4000 Hz).
4. The mean LDL remains stable as hearing loss increases for lower frequencies (500 and 1000 Hz). However, for higher frequencies (3000 and 4000 Hz) the mean LDL gradually increases with hearing loss. This is in agreement with findings reported by Kamm et al (1978), Hawkins et al (1987), Pascoe (1988), and Valente et al (1997).
5. The large intersubject variability reported for the REDD questions the validity of accurately predicting individual REDD from averaged group data.
6. Comparisons between WUSM and mean REDD data (NAL-NL1, DSL[i/o], ANSI S3.6-1989 and Valente et al (1997) report that the mean REDD reported in this study relates closely with mean REDD from other studies.

References

- American National Standards Institute (1996). *American National Standards Institute Specification for Audiometers (ANSI S3.6-1996)* New York: Acoustical Society of America.
- American National Standards Institute (1989). *American National Standards Institute Specification for Audiometers (ANSI S3.6-1989)* New York: Acoustical Society of America.
- ASHA Ad Hoc Committee on Hearing Aid Selection and Fitting (1988) Guidelines for hearing aid fitting for adults. *Am J Audiol* 7: 5-13
- Beattie RC, Sheffler MV. (1981) Test-retest stability and effects of psychophysical methods on the speech loudness discomfort level. *Audiology* 20:143-56.
- Beattie RC, Dion SA, Carmen RE, Kunkel HA. (1980) Loudness discomfort level for speech: comparison of two instructional sets for saturation sound pressure level selection. *Ear Hear* 1:197-205.
- Beattie RC, Edgerton BJ, Gager DW. (1979) Effects of speech materials on the loudness discomfort level. *J Speech Hear Dis* 44:435-458.
- Bentler RA, Cooley LJ. (2001) An examination of several characteristics that affect the prediction of OSPL90 in hearing aids. *Ear Hear* 22:58-64.
- Bentler RA, Nelson JA. (2001) Effect of spectral shaping and content on loudness discomfort. *J Am Acad Audiol* 12:462-70.
- Bentler RA, Pavlovic CV. (1989) Comparison of discomfort levels obtained with pure tones and multitone complexes. *J Acoust Soc Am* 86:126-132.
- Bornstein SP, Musiek FE. (1993) Loudness discomfort level and reliability as a function of instructional set. *Scand Audiol* 22:125-31.
- Byrne D, Dillon H. (1986) The National Acoustics Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid. *Ear Hear* 7:257-265.
- Byrne D, Dillon H, Ching T, Katsch R, Keidser G. (2001) The NAL-NL1 procedure for fitting non-linear hearing aids: Characteristics and comparisons with other procedures. *J Am Acad Audiol* 12:37-51.
- Ching T, Dillon H, Katsch R, Byrne D. (2001) Maximizing effective audibility in hearing aid fitting. *Ear Hear* 22:212-224.
- Cornelisse LE, Seewald RC, Jamieson DG. (1995) The input/output formula: A theoretical approach to the fitting of personal amplification devices. *J Acoust Soc Am* 97:1854-1864.

Cox RM. (1995). Using loudness data for hearing aid selection: the IHAFf approach. *Hear J* 48:10,39-44.

Cox R, Alexander G, Taylor I, Gray G. (1997) The contour test of loudness perception. *Ear Hear* 18:388-400.

Dirks DD, Kincaid GE. (1987) Basic acoustic considerations of ear canal probe measurements. *Ear Hear* 8:60S-67S.

Dillon H, Chew R, Deans M. (1984) Loudness discomfort level measurements and their implications for the design and fitting of hearing aids. *Austral J Audiol* 6:73-79.

Dillon H, Storey L. (1998) The National Acoustic Laboratories' procedure for selecting the saturation sound pressure level of hearing aids: theoretical derivation. *Ear Hear* 19:255-266.

Elberling C. (1999) Loudness scaling revisited. *J Am Acad Audiol* 10:248-260.

Fabry D. (2003) Real ear measurements and digital hearing aids: realities, myths and measurement techniques. *Phonak Focus* 32:1-15.

Filion PR, Margolis RH. (1992) Comparison of clinical and real-life judgments of loudness discomfort. *J Am Acad Audiol* 3:193-199.

Franks JR, Beckmann NJ. (1985) Rejection of hearing aids: attitudes of a geriatric sample. *Ear Hear* 6:161-166.

Gitles T, Niquette P. (1995) FIG6 in ten. *Hear Rev* 5:28-30.

Haggard M, Hall J. (1982) Forms of binaural summation and the implications of individual variability for binaural hearing aids. *Scand Audiol* 15s:47-63.

Hawkins DB. (1980) Loudness discomfort levels: a clinical procedure of hearing aid evaluations. *J Speech Hear Dis* XLV:3-15.

Hawkins DB. (1984) Selection of a critical electroacoustic characteristic: SSPL90. *Hear Instr* 35:28-32.

Hawkins DB, Walden BE, Montgomery A, Prosek RA. (1987) Description and validation of an LDL procedure designed to select SSPL90. *Ear Hear* 8:162-169.

Heinz MG, Issa JB, Young ED. (2005) Auditory-nerve rate responses are inconsistent with common hypotheses for the neural correlates of loudness recruitment. *J Assoc Res Otol* 6:91-105.

Jenstad LM, Connelisse LE, Seewald RC. (1997) Effects of test procedure on individual loudness functions. *Ear Hear* 18:401-408.

- Kamm C, Dirks DD, Mickey MR. (1978) Effect of sensorineural hearing loss on loudness discomfort level and most comfortable loudness judgments. *J Speech Hear Res* 21:668-81.
- Kirkwood DH. (2003) Survey of dispensers finds little consensus on what is ethical practice. *Hear J* 56:19-26.
- Kochkin S. (2000) Customer satisfaction with single and multiple microphone digital hearing aids. *Hear Rev* 7:24-29.
- Kochkin S. (2002) Market Trak VI: on the issue of value: hearing aid benefit, price, satisfaction, and repurchase rates. *Hear Rev* 9:20-41
- Martin FN, Champlin CA, Chambers JA. (1998) Seventh survey of audiometric practices in the United States. *J Am Acad Audiol* 9:95-104.
- Morgan DE, Wilson RH, Dirks DD. (1974) Loudness discomfort level: selected methods and stimuli. *J Acoust Soc Am* 56:577-581.
- Mueller HG. (2003) Fitting test protocols are “more honored in the breach than the observance.” *Hear J* 50:19-24.
- Mueller HG, Bentler RA. (2005) Fitting hearing aids using clinical measures of loudness discomfort levels: an evidence-based review of effectiveness. *J Am Acad Audiol* 16:461-72.
- Mueller HG, Bentler RA. (1994) Measurements of TD: how loud is allowed? *Hear J* 47:42-46.
- Mueller HG, Hornsby B. (2002) Selection, verification, and validation of maximum output. Valente M (ed.) in *Hearing Aids: Standards, Options, and Limitations*. New York: Thieme Medical Publishers, 23-65.
- Munro KJ, Nind L, Cleaver VC. (1996) The relationship between uncomfortable loudness level and maximum power output in subjects recently fitted with NHS hearing aids. *J Audiol* 30:275-286.
- Munro KJ, Patel RK. (1998) Are clinical measurements of uncomfortable loudness levels a valid indicator of real-world auditory discomfort? *Scand Audiol* 32:287-293.
- National Institute on Deafness and Other Communication Disorders (NIDCD) (2001) Hearing Aids. NIH Pub. No. 99-4340. Retrieved January 2004 from <http://www.nidcd.nih.gov/health/hearing/hearngaid.asp/>
- Palmer CV, Lindley GA. (2002) Overview and rationale for prescriptive formulas for linear and nonlinear hearing aids. Valente M (ed.) in *Hearing Aids: Standards, Options, and Limitations*. New York: Thieme Medical Publishers, 1-22.

- Palmer CV, Lindley GA. (1998) Reliability of the contour test in a population of adults with hearing loss. *J Am Acad Audiol* 9:209-215.
- Palmer CV, Lindley GA. (1997) Fitting wide dynamic range compression hearing aids: DSL [i/o], the IHAFF protocol, and FIG6. *Am J Audiol* 6:19-28.
- Pascoe DP. (1988) Clinical measurement of the auditory dynamic range and their relation to formulas for hearing aid gain. In Jensen JH (ed). *Hearing Aid Fitting*. Copenhagen: Storgaard Jensen, 129-154.
- Preminger JE, Neuman AC, Cunningham DR. (2001) The selection and validation of output sound pressure level in multichannel hearing aids. *Ear Hear* 22:487-500.
- Punch JL, Joseph A, Rakerd B. (2004). Most comfortable and uncomfortable loudness levels: six decades of research. *Am J Audiol* 13:144-157.
- Ricketts TA, Bentler RA. (1996) The effect of test signal type and bandwidth on the categorical scaling of loudness. *J Acoust Soc Am* 99:2281-2287.
- Sammeth CA, Birman M, Hecox KE. (1989) Variability of most comfortable and uncomfortable loudness levels to speech stimuli in the hearing impaired. *Ear Hear* 10:94-100.
- Sammeth CA, Peek BF, Bratt GW, Bess FH, Amberg SM. (1993) Ability to achieve gain/frequency response and SSPL-90 under three prescription formulas with in-the-ear hearing aids. *J Am Acad Audiol* 4:33-41.
- Scollie SD, Seewald RC, Corneliss LE, Jenstad LM. (1998) Validity and repeatability of level-independent HL to SPL transforms. *Ear Hear* 19:407-413.
- Seewald R, Cornelisse L, Ramji K, Sinclair S, Mookie K, Jamieson D. (1997) *A Software Implementation of the Desired Sensation Level (DSL[i/o]) Method for Fitting Linear Gain and Wide Dynamic Range Compression Hearing Instruments: Version 4.1a. User's Manual*. London: University of Western Ontario.
- Silverman SR. (1947) Tolerance for pure tones and speech in normal and defective hearing. *Ann Otol Rhin Laryngol* 56:658-677.
- Stach B. (2003) *Comprehensive Dictionary of Audiology*: Illustrated. Clifton Park, NY : Thomson\Delmar Learning.
- Valente M, Potts LG, Valente M. (1997) Differences and intersubject variability of loudness discomfort levels measured in sound pressure level and hearing level for TDH-50P and ER-3A earphones. *J Am Acad Audiol* 8:59-67.

Valente M, Potts LG, Valente M, Vass W, Goebel J. (1994) Intersubject variability of real-ear sound pressure level: conventional and insert earphones. *J Am Acad Audiol* 5:390-398.

Valente M, Van Vliet D. (1997) The independent hearing aid fitting forum (IHAFF) protocol. *Trend Amplif* 2:6-35.

Valente M, Abrams H, Sweetow R, Benson D, Chilson T, Solodar H, Citron D. (2006) Guidelines for the Audiological Management of Adult Hearing Impairment. In review for publication. Retrieved March 2006 from <http://www.audiology.org/professional/positions/haguidelines.pdf>

Wallenfels H. (1967) *Hearing Aids on Prescription*. Springfield, IL: Charles Thomas.

Watson LA. (1944) Certain fundamental principles in prescribing and fitting hearing aids. *Laryngoscope* 56:531-558.