ANALYSIS OF CHINCHILLA TEMPORARY AND PERMANENT THRESHOLD SHIFTS FOLLOWING IMPULSIVE NOISE EXPOSURE

REPORT WRITTEN BY:

William J. Murphy, Ph.D.

Amir Khan

Peter B. Shaw, Ph.D.

REPORT DATE:

March 2011

REPORT NUMBER:

EPHB 338-05c

NIOSH Interagency Agreement/Army MIPR

08-19-09M1 / MIPR9J07586218

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

Centers for Disease Control and Prevention

National Institute for Occupational Safety and Health

Division of Applied Research and Technology

Engineering and Physical Hazards Branch

Hearing Loss Prevention Team

4676 Columbia Parkway, Mail Stop C-27

Cincinnati, Ohio 45226-1998

DISCLAIMER

Mention of company names or products does not constitute endorsement by the Centers for Disease Control and Prevention.

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the National Institute for Occupational Safety and Health.

ANALYSIS OF CHINCHILLA TEMPORARY AND PERMANENT THRESHOLD SHIFTS FOLLOWING IMPULSIVE NOISE EXPOSURE

Table of Contents

Executive Summary

The analysis of the chinchilla impulsive noise exposures evaluated six potential noise exposure hazard indices (HIs) for goodness-of-fit and discrimination. The candidate HIs were the MIL-STD 1474D, A-weighted equivalent 8-hour level (LAeq8hr), Auditory Hazard Assessment Algorithm for Human (AHAAH) in the Unwarned and Warned condition, Pfander C-duration, and Smoorenburg D-duration. The Auditory Research Laboratory at State University of New York at Plattsburgh and the US Army Aeromedical Research Laboratory (Fort Rucker) collected auditory evoked potentials (AEP) from more than 900 chinchilla following exposure to impulsive noise exposures. For each exposure condition, a representative waveform was digitally recorded and archived along with the baseline AEP threshold, temporary threshold shift, permanent threshold shift and histological data from each animal. The exposures investigated the effects of peak level, number of impulses (1, 10, or 100) and temporal spacing of impulses (6, 60 or 600 seconds). The current analysis evaluated the goodness of fit through the use of mixed models that evaluated the immediate threshold shift (TS0) following exposure and the permanent threshold shift (PTS) evaluated approximately 4 weeks following exposure. The threshold shifts were evaluated using six different outcome variables: categorical classification for a 25 dB shift in hearing (permanent and temporary); categorical classification for a 15 dB shift in hearing (permanent and temporary); and as a continuous variable for threshold shift (permanent and temporary). Three explanatory variables were considered with respect to each exposure criterion: the exposure criterion, frequency, and baseline threshold.

Goodness-of-Fit: Generally, the statistical analysis demonstrated that LAeq8hr provided the best fit to the threshold shift data for both the permanent and temporary outcomes. The Pfander and Smoorenburg models generally demonstrated the second and third best fits. The Mil-Std 1474D typically had the poorest fit. Goodness-of-fit was judged using the Akaike and Bayesian information criteria. In a separate analysis, the threshold shift data were fit at the individual frequencies against the HIs using a logistic model and the threshold shift as a continuous variable. In these fits, the LAeq8hr was also demonstrated to have the best fit as demonstrated by the Coefficient of Determination, r^2.

Discrimination: Discrimination was tested by analyzing the Receiver Operator Characteristic (ROC) curves for each HI and the threshold shift outcomes. In this sort of analysis greater area

under the ROC curve (AUC) implies a greater ability to predict whether or not hearing loss will occur in the chinchilla. The discrimination results depended on the outcome variable. For the categorical permanent threshold shifts (25 dB and 15 dB) the Unwarned AHAAH provided the best discrimination. No statistically significant difference was observed between the Warned AHAAH and the LAeq8hr, however, both methods were significantly better discrimination than the Smoorenburg, Pfander and MIL-STD 1474D. For the categorical temporary threshold of 25 dB the Unwarned AHAAH, Warned AHAAH, and LAeq8hr indices were better than all the rest, but did not differ significantly from each other. For the categorical outcome of a 15 dB temporary shift, the LAeq8hr index was not significantly different from the Unwarned AHAAH, but better than all the rest. The Unwarned AHAAH was better than three of the rest.

Conclusions: The purpose of the interagency agreement between NIOSH and US Army Aeromedical Research Laboratories was to investigate the ability of the several hazard indices to fit the chinchilla data. The LAeq8hr index provided the best fit to the data for all outcome variables, with the Pfander and Smoorenburg indices second and third except in the case of the continuous outcome for permanent threshold shift. In the case of the continuous permanent threshold shift, the LAeq8hr index provided the best fit and the Unwarned AHAAH model had the second best fit. While the Unwarned AHAAH model exhibited better discrimination, the Warned AHAAH model did not exhibit significantly better discrimination than the LAeq8hr index.

List of Abbreviations and Acronyms

AHAAH…………………………………Auditory Hazard Assessment Algorithm for Human

AHU.. Auditory Hazard Unit

AIC.. Akaike Information Criterion

AUC ....Area Under the ROC Curve

BASE ....Baseline Auditory Evoked Potential Threshold

BIC... Bayesian Information Criterion

dBA...A-weighted Sound Pressure Level,

dBPTS..Permanent Threshold Shift as a continuous variable

dBTS0 .... Temporary Threshold Shift as a continuous variable

gllamm.............................Procedure to fit generalized linear mixed models in STATA

LAeq8hr ....A-weighted 8-hour Equivalent Level

L_P...

.... Level calculated with Pfander Criterion

L_{pk} ....Peak Sound Pressure Level re $20x10^{-6}$ Pa

L_S ....Exposure Level calculated with the Smoorenburg Criterion

MIL-STD-1474D .... Military Standard 1474D

NIOSH ....National Institute for Occupational Safety and Health

p_0 .... Reference Pressure $20x10^{-6}$ Pascals

PTS.. Permanent Threshold Shift *ptscat15*.................. Binary variable for Permanent Threshold Shift with 15-dB or more shift

ptscat25.................... Binary variable for Permanent Threshold Shift with 25-dB or more shift

ROC ....Receiver Operating Characteristic Curves

SELA... A-weighted Sound Exposure Level

T_D .... D-duration of the reverberent envelope

TS0... Temporary Threshold Shift at Time

ts0cat15 Binary variable for Temporary Threshold Shift with 15-dB or more shift

ts0cat25....................Binary variable for Temporary Threshold Shift with 25-dB or more shift

TSMAX.. Maximum Temporary Threshold Shift

I. Introduction

A. Background

For more than 50 years, the US Army has conducted and sponsored research designed to assess the risk of hearing loss due to exposure to high-level noise from weapons and weapon systems. As a subset of these exposures, research has focused on developing a better understanding of how various parameters of impulsive noise exposure affect hearing. The ability to assess the hazard of impulsive noise exposures is critical for: (a) protection of the war fighter (b) development of weapon systems, and (c) the implementation of hearing conservation programs (i.e., selection of hearing protection devices).

Since approximately 1980, the U.S. Army Medical Research and Materiel Command has funded a series of investigations into the effects of impulsive noise exposures on hearing using chinchillas as a surrogate animal model for human exposures (Hamernik et al. 1998a, 1998b). These exposures have systematically investigated the intensity (peak level), the spectrum, number and temporal spacing of impulses, the development of an isohazard spectral weighting function, the effects of reverberation and the effects of impulse peak versus energy[1]. The data resulting from these exposures include one impulse waveform for each noise exposure type, the temporary (compound) and permanent threshold shift of the auditory evoked potential at several frequencies and the quantitative estimate of inner and outer hair cell counts. These data may provide valuable insight into the effects of impulsive noise exposures and

---

[1] The detailed references are summarized in the Hamernik et al. contract reports 1998a and 1998b

methods to characterize the relative hazard of the exposures and the effect on the auditory mechanism in a species with hearing capabilities similar to humans.

The U.S. Army Medical Research and Material Command, Aeromedical Research Laboratory (USAARL) entered into an interagency agreement (08-19-09M1, MIPR9J07586218) with the National Institute for Occupational Safety and Health (NIOSH) for the purpose of applying several damage risk criteria to the impulse noise exposure data in order to evaluate whether these criteria provide reasonable predictors of the hearing loss observed in the chinchilla model.

USAARL provided NIOSH with the 50 acoustic waveforms used in 137 exposure conditions where the number of impulses and the inter-stimulus interval (ISI) were varied. The temporary threshold shifts, permanent threshold shifts and histology for 905 animals were provided in a Microsoft Access database. Separate tables defined the stimulus, exposure conditions, audiometric assessments and histological evaluations. Audiometry was conducted for baseline hearing thresholds (BASE) at 125, 250, 500, 1000, 2000, 3000, 4000, 6000, 8000, and 11200 Hz. All animals had thresholds measured immediately following exposure (TS0) and at several post exposure times in order to establish the maximum temporary threshold shift (TSMAX) at 500, 2000, and 8000 Hz. Permanent threshold shift (PTS) was evaluated at 500, 1000, 2000, 4000, 8000 and 11200 Hz for most animals. PTS data at 11200 Hz were not collected for 212 animals. Hearing thresholds were measured using the auditory evoked potential measured from a pair of electrodes chronically implanted in the inferior colliculus of brainstem (signal) and the dura of the cortex (reference). A third electrode (typically a surface or subcutaneous electrode) provided the ground for the differential measurement of the evoked potential.

The impulsive noise exposures were analyzed with the MIL-STD-1474D (1997), LAeq8hr (DTAT, 1983; Dancer 2003) AHAAH (Price and Kalb, 1991; Price 2007a, 2007b), Pfander (1982) and Smoorenburg (1982, 1992) criteria.

B. Previous Analysis of Chinchilla Data

The Auditory Research Laboratory of the State University of New York at Plattsburg, New York has previously published several papers on the effects of impulse noise exposure in chinchilla for the purpose of developing improved hazard criteria[2]. Primarily, two analytical approaches were used: nonlinear regression of the threshold shifts without regard to the particular exposure and isohazard analysis which tried to identify exposures that should have produced similar shifts in hearing based upon the level and number of impulses. Both approaches showed that the spectral weighting function provided the best fit to the data. Specifically, the P-weighting function and its variants (P1, P2 or R) are similar to A-weighting except that more of the low and high frequency energy is removed below 1 kHz and above 10 kHz, respectively. The variant forms treat the mid frequencies with some emphasis or de-emphasis of the energy that might be reminiscent of the transfer function of the chinchilla's pinna (Murphy and Davis, 1997; Song and Kim, 2008).

Chan (2005) used the SUNY/USAARL chinchilla dataset to develop a human impulse noise injury model for unprotected ears. Chan applied the A-weighted Sound Exposure Level (SELA) to estimate the probability that an ear would be injured immediately or permanently following a given exposure. While the model is designed to provide estimates of exposures for

---

[2] See Hamernik et al. 1998a and 1998b for a detailed listing of contract reports and journal articles.

humans in the form of a risk based on chinchilla exposures, it was nonetheless a novel approach to dealing with the wide range of variability observed in the exposure effects.

This analysis seeks to determine which metric best describes the data in a manner similar to the Hamernik et al.'s investigation (1998a, 1998b). The US Army Research Laboratory has developed a model of the response of the human ear (Auditory Hazard Assessment Algorithm for Human, AHAAH) that uses a recording of a noise and processes the noise through an electro-acoustic equivalent model of the human ear (Price and Kalb, 1991; Price 2007a; 2007b). NIOSH investigators have used the AHAAH model to analyze a wide range of impulsive noises from field studies of gun shot noise, the Albuquerque Blast Overpressure Walkup Study (Murphy et al., 2009) and now the Chinchilla Blast Overpressure data. Our approach to the analysis was to apply three different statistical models and to compare several damage risk criteria by evaluating goodness of fit and discrimination using the different statistical models with each of the criteria.

II. Method

The impulse waveforms were analyzed using LAeq8hr, MIL-STD 1474D, Unwarned AHAAH model, Warned AHAAH model, Pfander and Smoorenburg hazard indices (HIs). According to the Chinchillas Blast Wave Exposure Study Protocol, each of the 905 chinchillas was exposed to one of 137 different exposure conditions. The exposure evaluations provide a wide range of exposure conditions, numbers of impulses and interpeak intervals that may be related to hearing loss and cochlear sensory cell loss. The main objective of this research was to determine the best indicator of the amount of hazard associated with an impulse noise exposure. The first effort was to evaluate each of the waveforms for the HIs and perform a regression of the TS0 and PTS against the HIs. Table 1 describes the different impulsive sources. Table 2 describes the association between the sources and the exposure groups. Table 3 describes the evaluation of the impulse waveforms as evaluated by the different HIs. In Table 3, the number of impulses (1, 10, 100) and the interpeak interval (6, 60, 600 or 3 seconds) are given.

A. Hazard Indices

1. MIL-STD 1474D

MIL-STD 1474D is a Department of Defense Design Criteria Standard (1997) that provides specific noise limits and related requirements to equipment designers and manufacturers. These limits should not be exceeded if the materiel is to be acceptable and are intended to cover typical operational conditions. The limits evolved from considerations of hearing damage-risk, speech intelligibility, aural detection, state-of-the-art noise reduction, and government legislation. The upper limit for MIL-STD-1474D with single hearing protection is 177 dB for very short B-durations (about 1 ms).

The MIL-STD-1474D effective exposure levels were determined for each waveform as follows:

$$L_m = L_{pk} + 6.64\log(T_B / 200) + 5\log(N), \qquad (1)$$

where L_{pk} is the peak sound pressure level in dB SPL, T_B is the B-duration in milliseconds and N is the number of impulses. L_{pk} was the maximum positive peak pressure in the waveform. B-duration was derived from an exponentially decaying function fit to the waveform envelope determined by the magnitude of the complex Hilbert transform of the pressure signal (See

Figure 1. Definition of impulse noise duration (Smoorenburg, 1992)). For the shock tube data, a B-duration cannot necessarily be estimated because there was no reverberant energy due to reflections in the room. Regardless, the duration of the impulse waveform that was within 20 decibels of the peak was used for the B-duration. The Hilbert Transform effectively phase shifts the waveform 90 degrees in the time domain and the magnitude yields the amplitude envelope of an arbitrary wave (Zechmann, 2009). The MIL-STD 1474D exposures ranged from 140.1 dB for a 3350 Hz narrow band impact noise of a nominal 124 dB peak and 100 impulses to 181.7 dB for a conventional shock tube in a reverberant environment with a nominal peak pressure of 160 dB and 100 impulses (Patterson et al., 1993; Ahroon et al., 1996). The range of exposures covered approximately 40 dB.

2. LAeq8hr

LAeq8hr is the A-weighted acoustic energy delivered to the ear for an equivalent eight-hour exposure. The A-weighting curve is an approximation of the equal loudness perception curve for pure tones relative to a reference of 40 dB sound pressure level at 1000 Hz. The inverse of the A-weighting curve also provides a model of the transfer function of the outer and middle ear for the human. The chinchilla has a similar outer/middle ear transfer function as that of the human. Thus the frequency range and the dynamic range for sensitivity are quite close to that of humans. However, it is understood that chinchilla tend to be more sensitive to hearing loss than humans due to a better impedance matching between air and the cochlea. The French Committee on Weapons Noises advocated the use of A-weighted energy in the form of LAeq8hr as a damage risk criterion for unprotected ears with a limit of 85 dB (DTAT, 1983; Dancer, 2003).

LAeq8hr integrates the energy of an impulse and equates the result to an equivalent amount of energy for an A-weighted 8-hour exposure to a continuous noise. The pressure-time waveform is filtered first in the time domain with an A-weighting filter and then the energy is integrated and adjusted for duration and the number of impulses as follows:

$$L_{Aeq8hr} = 10\log\left(\frac{1}{t_2 - t_1}\int_{t_1}^{t_2}\frac{p_A^2(t)}{p_0^2}dt\right) + 10\log\left(\frac{t_2 - t_1}{T_{8hr}}\right) + 10\log(N), \qquad (2)$$

where L_{Aeq8hr} is the equivalent 8-hour, A-weighted sound pressure level (dB), p_0 is the reference pressure level (20 µPa), $p_A(t)$ is the A-weighted pressure time-waveform in Pascals, t_1 is the start time of the impulse event (secs), t_2 is the end time of the impulse (secs), T_{8hr} is the equivalent time in seconds (28,800 secs) and N is the Number of impulse events (Earshen, 2003, Zechmann, 2009). Hamernik et al.'s (1998a, b) analysis considered A-weighted Sound Exposure Level (SELA), which differs from LAeq8hr by replacing T_{8hr} in the second term of Eq. (2) with $T_{1\ sec}$, thus reporting a an equivalent energy for one second of exposure. In this analysis, that amounts to adding a constant to Hamernik et al.'s results which would shift the curves by 10log($T_{1\ sec}$ / T_{8hr}) = -44.6 dB. Exposures ranged from 58.9 dBA for a single spark-gap impulse (peak level of 150 dB) to 105.1 dBA for 100 impulses due to a fast acting valve (peak level of 160 dB) (Ahroon et al., 1996).

3. AHAAH Model

The AHAAH model is an electro-acoustic model of the ear designed to approximate the response properties of the ear and reproduce the measured transfer functions from free-field to the stapes and then into basilar membrane displacements (Price and Kalb, 1991; Price 2007a, 2007b). The basilar membrane response is modeled coarsely by a 23-element network transmission line that correspond to 1/3rd octave band intervals. An estimate of Auditory Hazard Units (AHU) is calculated at each location by squaring the peak amplitude of each upward displacement of the basilar membrane (in microns) and summing them for the analysis interval. The maximum AHU at any of the 23 segments is defined as the auditory hazard of an exposure. When an impulse exceeds a predefined threshold (108 dB SPL in this analysis), the stiffness of the middle model is increased reducing the transmitted energy to the cochlear portion of the model. In the AHAAH model, the Warned condition presumes that the middle ear reflex response has already been activated prior to the arrival of the impulse. The Unwarned AHAAH model allows the impulse waveform to activate the middle ear reflex response. Dynamically, this effect is seen to be an asymptotic increase of the stiffness over a time of about 200 milliseconds commencing at the start of the peak impulse.

The AHAAH model Version 1.1 was used to process the digitized waveforms (http://www.arl.army.mil/www/default.cfm?Action=31&Page=343). Both the warned and unwarned AHUs were computed for each waveform. For each exposure cell, the average AHU for all waveforms was determined. To estimate the total for any given exposure cell, the AHU dose for the waveform is multiplied by the number of impulses the animal experienced. The auditory hazard units were computed for the Warned and unwarned exposure conditions which are summarized in Tables 3. The minimum HIs for the Unwarned and Warned AHAAH conditions were 36.6 and 4.0 AHU, respectively, produced by a conventional shock tube in a nonreverberant environment. The maximum Unwarned AHAAH was 162982 AHU for a 2450 Hz narrow band impact with a peak level of 144 dB and 100 impulses, whereas the maximum Warned HI was 73228 for the fast-acting valve with a peak level of 160 dB and 100 impulses.

4. Pfander (1980)

The Pfander effective exposure level, L_P, is calculated as follows:

$$L_P = L_{pk} + 10\log(T_C) + 10\log(N), \quad (3)$$

where L_{pk} is the peak pressure, and T_C is the C-duration (the integrated time in milliseconds where the absolute amplitude of the waveform is within 10 dB of the peak pressure) and the trading ratio for impulses is 10 log(*N*) (Pfander, 1980). Chan et al., (2001) examined the goodness-of-fit of the Albuquerque Blast Overpressure study data with a modified Pfander HI adding 15-dB peak reduction of a hearing protector. While Chan found LAeq8 to be a better functional fit than Pfander, the Pfander criterion yielded a better fit than the current MIL-STD 1474D. Hamernik et al. examined a Pfander effective exposure level in their report as well (Peak level with C-duration and N). The Pfander HI has been included in this report to provide an historical link to the two reports. The Pfander exposure levels ranged from 138.6 dB produced by a single spark gap impulse (150 dB peak pressure) to 184.5 dB for 100 impulses from a conventional shock tube (160 dB peak pressure) in a reverberant environment.

5. Smoorenburg (1982)

Similarly, the Smoorenburg effective exposure level, L_S, is calculated as follows:

$$L_S = L_{pk} + 10\log(T_D) + 10\log(N), \quad (4)$$

where L_{pk} is the peak sound pressure level, T_D is the D-duration in milliseconds and *N* is the number of impulses (Smoorenburg, 1982). The D-durations was calculated according to the procedure described above for the B-duration. The magnitude of the Hilbert Transform provided the amplitude envelope and the D-duration is then the period of time where the envelope is within 10 dB of the peak sound pressure level. The Smoorenburg exposure levels ranged from 140.8 for a single impulse from the spark gap (150 dB peak pressure level) to 184.5 for the 100 impulses of conventional shock tube in a nonreverberant room (155 dB peak pressure level). Surprisingly, the 160 dB stimulus was not evaluated as having the maximum L_S because the T_D was considerably shorter than the T_D for the 155 dB impulse.

B. Statistical Analysis

1. Nonlinear Curve Fit

Each of the Hazard Indices provided an evaluation of the exposure for a given stimulus and exposure group. In the case of those groups where the animals were exposed at different interpeak intervals (IPI), the evaluation of the hazard yields the same estimate regardless of the IPI. The TS0 and PTS results should provide some sense of the organization of the outcome with increasing estimated hazard. That is, if the exposure criterion is higher, then the animals exposed should exhibit greater tendency of temporary and permanent effects. One way to get a sense of the tendency was to perform a nonlinear curve fit of the TS0 or PTS data against the HIs. In this case, the following functional form was applied:

$$y = \frac{A}{1 + e^{-b(x - x_0)}}, \tag{5}$$

where y is the threshold shift outcome (TS0 or PTS in decibels), x is the exposure criterion (Mil-Std 1474D, LAeq8hr, Unwarned AHAAH, Warned AHAAH, Pfander or Smoorenburg), A adjusts the magnitude of the average threshold shift at the highest levels, b adjusts the slope of the exponential increase and x_0 adjusts the curve to the right or left. Equation was transformed to a linear function and a linear regression was applied,

$$\ln\left(\frac{A}{y - y0} - 1\right) = -bx + bx_0, \quad (6)$$

where A was set at the maximum value for the threshold shift for the permanent or temporary threshold shifts, y_0 is set at a value below the minimum threshold measured for the group of animals.

The fits were independently performed for each frequency of TS0 and PTS data against each of the hazard indices using the Matlab `fit()` function. A coefficient of determination, r^2, varies between 0 and 1 and can be interpreted as the proportion of variation explained by the predictor variable.

2. Linear Mixed Models

The goal in analyzing the chinchilla data was to ascertain which exposure criterion best characterized the effect of noise exposure on threshold shift. Effects due to the type of noise exposure and the frequency of auditory testing were considered in the analysis. The general linear mixed model was used (Laird and Ware, 1982) with temporary or permanent threshold shift as the outcome variable and noise exposure (one of the noise hazard indices), type of exposure, and frequency as the explanatory variables. The correlations of measurements taken at different frequencies on the same chinchilla were incorporated by treating subject (chinchilla) as a random variable in a mixed effects model. The exposure group was also considered a random variable and frequency a fixed variable. The models for the different noise hazard indices were evaluated and goodness-of-fit were judged using the Akaike Information Criterion. The models using different indices were not nested, so likelihood ratio tests were not used.

Six exposure criteria were evaluated: MIL-STD 1474D, LAeq8hr, Unwarned AHAAH, Warned AHAAH, Pfander, and Smoorenburg. The exposure criteria were incorporated into generalized linear mixed models (Fitzmaurice et al., 2004) and linear mixed models (Laird & Ware, 1982) as explanatory variables and the models were compared using the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). The AIC and BIC are model-selection criteria in which

$$\mathrm{AIC} = -2\ln(likelihood) + 2p \tag{7}$$

$$\mathrm{BIC} = -2\ln(likelihood) + p\ln(N) \tag{8}$$

where p = number of parameters and N = sample size. In the case of repeated measures, N may indicate the number of samples or the number of subjects; we have used the number of subjects – the same convention as currently used in PROC MIXED in SAS (though Stata was used for the analysis). The larger the likelihood, the better the fit of a model, thus smaller values of $-2\ln(likelihood)$ indicate a better fit of the model to the data; the terms $2p$ and $p\ln(N)$ are penalties imposed for increasing the number of parameters (Long, 1997, p. 109). Since the groups of models being compared had the same p and N, AIC and BIC could be used for comparing model fit. For each set of six models based on the six exposure criteria, we determined the order of the AIC's and BIC's. The order of the AIC's and BIC's were the same for each set.

The outcomes of interest included four binary variables and two continuous variables. Given the previous human data sets evaluated by Chan et al., (2001), the binary outcome variable, *ptscat25*, equaled 1 if the permanent threshold shift (dBPTS) was greater than or equal to 25 and equaled 0 if dBPTS < 25. The binary outcome, *ptscat15*, equaled 1 if 15 ≤ dBPTS and 0 if dBPTS < 15. The binary outcome, *ts0cat25*, was set equal to 1 if the temporary threshold shift (dBTS0) was greater than or equal to 25 and set equal to 0 if dBTS0 < 25. The final binary outcome variable, *ts0cat15*, was set equal to 1 if 15 $\leq$ dBTS0 and set equal to 0 if dBTS0 < 15. The continuous outcome variables were dBPTS and dBTS0.

The binary data were modeled using generalized linear mixed models, as implemented in the Stata command *gllamm* (StataCorp, 2007). The models were all three-level logistic random intercept models (Rabe-Hesketh & Skrondal, 2008, p. 446). In all models subject (individual chinchilla) and exposure code were treated as random effects, with subject nested within exposure code.

Correlation within an exposure group is accounted for by treating subject as a random variable. Correlation within repeated measures of a single animal is expected to be higher than the correlation between animals in the same exposure group. Correlation between animals in the same exposure groups is modeled by allowing exposure group to be a random variable.

The models differ in outcome variables and in fixed effects. The outcome variables (four binary and two continuous) have been described above. The three general types of models in regard to fixed effects are 1) those with just exposure criterion; 2) those with exposure criterion + frequency; and 3) those with exposure criterion + baseline threshold. For example, the form of the model with just the exposure criterion for a fixed effect is:

$$\ln\left\{\frac{\Pr(y_{hij}=1\,|\,bi_j,b_j)}{\Pr(y_{hij}=0\,|\,b_{ij},b_j)}\right\}=\beta_0+\beta_1 x_{hij}+bi_j+b_j \qquad (9)$$

where x_{hij} = value of exposure criterion for occasion h on chinchilla i for exposure code j;

β_0 = fixed intercept component;

β_1 = coefficient of fixed effect of exposure criterion;

$h = 1, \ldots, n_i$ = number of occasions measurements taken on chinchilla *i;*

b_{ij} = random intercept component for chinchilla *i* in exposure code *j;*

b_j = random intercept component for exposure *j*.

In performing the numerical integration required in *gllamm*, Gauss-Hermite quadrature (the default) was used. Initially the default number of integration points, 8, was used, but fairly substantial changes in the values of AIC and BIC appeared when 20 integration points were used. The calculations were re-run a third time using 25 integration points and only small changes were found in comparison with the results using 20 integration points (mean of absolute changes = 1.104; standard deviation of absolute changes = 1.134); the ranks of the AIC and BIC did not change in going from 20 to 25 integration points. Accuracy in estimation should increase with increasing number of integration points (Rabe-Hesketh & Skrondal, 2008, p. 67), thus the results are reported using 25[*] integration points.

The continuous outcomes, permanent threshold shift (dBPTS) and temporary threshold shift (dBTS0), were modeled with linear mixed models, using the Stata command *xtmixed*

---

[*] The drawback to increasing the number of integration points is the increased time needed for calculation. For example, to run the generalized linear mixed model with 25 integration points, with just the exposure criterion MIL STD 1474D for the fixed effect and *ptscat25* for the outcome variable, took 48 min 49 sec. This was using the 64-bit version of Stata 10 with 10 of the 16 available GB of memory allotted to Stata on an Intel Xeon processor running at 3.72 GHz.

(StataCorp, 2007). As with the above models, subject and exposure code were treated as random variables with subject nested within exposure code.

In addition to determining which exposure criterion provided the best model for the data we also sought to determine what effect, if any, frequency and baseline threshold had in predicting the outcome of interest. Normally a problem of this nature might be approached with all three fixed effects initially in the model (exposure criterion + frequency + baseline threshold) and various model reductions, transformations, interactions, and other possible effects of one explanatory variable on another would be considered. However, for this study a comparison of just the exposure criteria was of paramount importance. Therefore we examined all 36 models with just exposure criterion as the fixed effect and then models with frequency or baseline threshold added to exposure criteria.

3. Receiver Operating Characteristic Modeling of Discrimination

ROC (receiver operating characteristic) curves were used to compare six sound exposure criteria, using hearing data obtained from chinchillas. The basic goal was to determine how well a given exposure criterion predicted loss of hearing, and which one did it best. ROC curves were used for the following reason. Though not used here, one might consider using a classification table (Hosmer & Lemeshow 2000, pp. 156-160) to evaluate the exposure criteria, as shown in Table 14.

In Table 14, the sensitivity would be $\left(\frac{d}{b+d}\right)*100\%$ and the specificity would be $\left(\frac{a}{a+c}\right)*100\%$. Table 15 shows an example of subjects classified as being in hazardous or safe sound conditions relative to those that did or did not experience a hearing loss due to the exposure. A major drawback to such an approach is that the numbers in the table depend on how an event is classified. The outcome is binary, i.e., present or absent, and the predictor (exposure criterion) is continuous. Since the exposure criterion is continuous, a certain level or cutoff must be chosen to predict presence of a characteristic. The numbers in the cells (a, b, c, and d) will depend on the chosen cutoff. Thus one could generate many different classification tables, using different cutoffs.

A solution to this problem is to use an ROC curve, which basically uses a series of cutoffs. The procedure is to plot the true positive rate (*sensitivity*) vs. false positive rate (1 - *specificity*) for a series of cutoffs, as illustrated in Figure 17. The larger the area under the curve, the better a marker will be. Thus one can compare markers by comparing the area under the curve (AUC), also referred to as Harrell's (1996) C-index.

ROC analysis is commonly used in medical research to compare different methods of discrimination. The general idea is that an AUC of 0.5 would represent no better than random assignment to either of the two possible categories and an AUC of 1.0 would represent perfect prediction. Table 16 provides an interpretation for values of AUC ranging from 0.5 to 1.0 as suggested by Hosmer & Lemeshow (2000, p. 162). While this rule of thumb was developed for independent observations; it should be a useful guide with respect to clustered data.

In calculating the AUC's for the different exposure criteria for chinchillas one must recognize the correlated nature of the data. Calculation of AUC is a well-established procedure for independent observations (see, for example, Agresti, 2002). However, the chinchilla data are not independent observations because of the repeated threshold measures at several frequencies on each animal. For this reason the nonparametric method developed by

Obuchowski (1997) for clustered data was used. In this case the cluster consists of the observations taken from a single chinchilla.

The outcomes of interest included four binary variables. The binary outcome variable, *ptscat25*, equaled 1 if the permanent threshold shift (dBPTS) was greater than or equal to 25 and equaled 0 if dBPTS < 25. The binary outcome, *ptscat15*, equaled 1 if 15 $\leq$ dBPTS and 0 if dBPTS < 15. The binary outcome, *ts0cat25*, was set equal to 1 if the temporary threshold shift (dBTS0) was greater than or equal to 25 and set equal to 0 if dBTS0 < 25. The final binary outcome variable, *ts0cat15*, was set equal to 1 if 15 $\leq$ dBTS0 and set equal to 0 if dBTS0 < 15. The software used to implement Obuchowski's (1997) method was the Stata command *somersd* (Newson, 2001).

III. Results

A. Waveform Evaluation

1. Spectral Comparison of AHAAH and A-weighted Analysis

The AHAAH analysis allows one to model the response of the cochlea to a particular impulse and determine the location(s) of damage. The AHAAH model first propagates the waveform from the outer ear (at the free-field, concha or ear canal opening positions) to the middle ear and to the stapes footplate. Because the AHAAH model utilizes nonlinearity in the middle ear response by dynamically increasing the stiffness over the course of a few hundred milliseconds and by limiting the displacement of the stapes footplate to mimic the annular ligament suspension, the propagation must be performed in the time domain to solve for the stapes displacement as a function of time. The cochlea is included in the time domain solution as a lumped impedance element.

Once the stapes displacement is solved, then a WKB solution is applied in the frequency domain to estimate the response of the basilar membrane. The WKB solution is appropriate since the cochlea is modeled as a linear transmission line. The AHAAH model divides the cochlea into 23 segments of the basilar membrane and maps the mass, resistance and stiffness to values that are physiologically representative of the human cochlea. While it is well-known that the dynamics of the basilar membrane are nonlinear for low input levels, these nonlinearities are compressive and are effectively masked by the linear behavior of the cochlea at high input levels above about 80 dB SPL. Thus accurate integration of the basilar membrane response should be largely unaffected by the low-level nonlinearities when the input levels are several orders of magnitude greater than 80 dB SPL.

The AHAAH model provides a temporary hazard file (`temp.haz`) which reports the frequencies and AHU's associated with each segment. For each of the 50 stimuli, the hazard file was saved and stored in an array to allow visualization of the basilar membrane response. The AHUs are plotted as a series of spectra after transforming the AHU to a decibel form:

$$\mathrm{AHU(dB)} = 10\log\left(\frac{\mathrm{AHU}}{10^{-5}}\right) \qquad (10)$$

where AHU is the hazard at any basilar membrane location and 10^{-5} represents the minimum quantity output from the AHAAH model. The particular choice of a reference value (10^{-5}

AHU) will shift the relative magnitude of the AHU in decibels by a constant. The suggested daily dose for the AHAAH model is 500 AHU.[3]

The Unwarned AHAAH spectra are plotted in Figure 2. The relative spectra of the exposures demonstrate that the AHAAH model differentiates the exposures based upon the spectral response of the cochlear model. The time-integrated response of the model effectively provides a frequency band analysis of the predicted damage.

The Warned AHAAH spectra are plotted in Figure 3. In this case, the middle ear muscles are activated prior to the arrival of the impulse waveform. Comparing the Unwarned and Warned responses using the same Z-axis scale and color scaling, the Unwarned responses exhibit greater estimated hazard over the entire basilar membrane. The maximum hazard locations for the narrow band noise exposures (stimulus codes 18-40) are the same for the Unwarned and Warned conditions. While the effect of warning the ear will stiffen the ossicular chain and reduce the high frequency energy entering the ear, for these stimuli the warned ear has essentially the same shape, only less hazard. The shifts in energy that were observed moved the maximum by one site towards the apex (lower frequency). The decibel

---

[3] Another point for consideration is the fact that the AHAAH Model is currently designed for human auditory periphery and the exposures were conducted on chinchillas. Since this analysis applies the human damage risk criteria to this data, the anatomical differences cannot be discussed until the Army Research Lab creates a working chinchilla model. Dr. Kalb has provided various sets of coefficients that might be used; however the simulation results were not physically reasonable. Thus the human AHAAH model is used to evaluate the waveforms and is compared to the evaluations of the waveforms and hazard indices from the other human-based criteria.

change in the AHU value between adjacent sites was less than 1 and often around 0.1 dB or about 20% of the linear AHU.

The A-weighted spectra as provided in the Access database are shown in Figure 4. The frequencies range from 100 to 16000 Hz. In Figure 5and Figure 6, the comparative frequency and level by band information are shown for each stimulus for the A-weighted one-third octave bands (red circles), Unwarned AHAAH (blue squares) and the Warned AHAAH model (black diamonds) evaluations. The frequencies from the AHAAH model range from 600 Hz to 11,600 Hz. The A-weighting filter is somewhat more severe at the low and high frequencies than the AHAAH model for the narrow band impact noises as indicated by the vertical range of the data (stimuli 18-40). For the broad band impulses produced by the acoustic shock tubes (stimuli 1-3, 13-15, 50), spark gap generator (stimuli 10-12) and fast acting valve (stimuli 4-9, 16-18), the AHAAH model and A-weighting filter provide a similar range of data and generally place the peak frequency of damage in the same region of the cochlea. The comparison between A-weighting and AHAAH is useful because they are both intended for use with human. If an AHAAH model for the chinchilla were available, then the predictions of a maximal damage along the basilar membrane could be correlated with the histological data. Examinations of chinchilla histological data with hearing loss is complicated due to the large variance of TTS, PTS and loss of inner and outer hair cells (Zhu et al., 2009). As will be seen, the hearing loss data are highly variable and well correlated with the LAeq8 metric (Hamernik, 1998a, b).

Table 3 reports the numeric evaluation of the waveforms for each of the 137 exposure groups. Since many exposure groups differed only in the number of impulses, "*N*", the terms in Equations 1-4 differentiate the severity of exposure (e.g. a factor of 1, 10 or 100 for the

AHAAH; 0, +10, + 20 dB for the LAeq8, Pfander and Smoorenburg and 0, +5 or +10 dB for the MIL-STD 1474D criteria). As can be seen from Equations 1 through 4, the interpeak interval is not included as a variable. Thus, the exposure groups that were exposed to impulses having different intervals will be evaluated as being the same.

2. Nonlinear Curve Fit

In Figure 7 to Figure 16 TS0 and PTS are plotted to investigate the relation between the exposure outcomes and the exposure metric. As described earlier, the TS0 and PTS data were fit to a linearized logistic curve that adjusts for maximum threshold shift (*A*), minimum threshold shift (y_0) slope of the dose response function (*b*) and the location of the midpoint of the dose function (x_0). These various curves that were fit are meant to illustrate the dose response relation for the different frequencies with the estimated HIs for the six exposure criteria.

In Figure 7 for the 500 Hz TS0 data, the MIL-STD 1474D exhibited the poorest organization and lowest coefficient of determination, r^2, with a value of 0.055. Although the correlation is low, it is statistically significant given the large number of observations (n = 905, $p << 0.001$) of animals. For exposure levels below about 160 dB, the low correlation is evident. Many of the animals in those exposure groups exhibited significant TS0 (>40 dB) for the lowest exposures. For the other metrics, the correlations were greater and the spread of the TS0 data at low and high exposure levels was less than that observed for the MIL-STD metric. LAeq8hr had the highest coefficient of determination; r^2 exhibited the least spread in the middle range of exposures. The trend that can be seen for LAeq8hr Figure 7 compared to the other metrics is that it has fewer low TS0 values for the highest exposure levels. Thus the fit of the curve to the data will tend to be better and the r^2 will be larger.

In Figure 8 and Figure 9, the same general trend is observed for the TS0 data measured at 2000 and 8000 Hz. In both figures the Warned AHAAH had the poorest coefficient of determination of 0.113 and 0.168, respectively for 2000 and 8000 Hz. Similarly the LAeq8hr had the highest r^2 of 0.390 and 0.455 for 2000 and 8000 Hz. Particularly for the LAeq8hr, the low exposure levels exhibited TS0 about 20 dB or less. At the higher exposure levels, animals generally exhibited TS0 more than 20 dB and typically 40 to 60 dB. For the Pfander and Smoorenburg criteria, several of the highest exposure groups had animals that did not exhibit a large TS0. Thus the r^2 was lower than that estimated for LAeq8hr. For the AHAAH model in both unwarned and warned conditions, the curve had lower r^2 than LAeq8hr. As can be seen in the fits, the low exposure levels reached an asymptote that was significantly greater than 0 at the left side of the curve. Because the AHAAH model is defined as a linear quantity representing the summation of the square of the basilar membrane displacement, the metric can never be less than 0. Small, relatively innocuous exposures will yield AHUs of about 0 while more hazardous exposures will yield 200 or more AHUs. If the AHAAH model were logarithmic, then innocuous exposures would be well-separated from the hazardous exposures as illustrated in Figure 10.

The linear versus the logarithmic responses of the AHAAH model are compared in Figure 10 for the TS0 and PTS data at 8000 Hz. The r^2 are shown for both. For the AHAAH model with AHU in dB relative 10^{-5} AHU show in the top row, the r^2 markedly improved from 0.222 to 0.394 for the TS0 data and from 0.185 to 0.239 for the PTS data. When the data are plotted on a linear abscissa (rather than the logarithmic abscissa), the exposure levels are bunched together on the left side of the plot as seen in the lower panels. When the nonlinear curve fit is applied, the functional form allows for the exposure level to be negative. The AHAAH model, however, cannot yield a negative value. Thus, the curve fit intercepts the ordinate at a point that is significantly greater than 0 dB TS0. When plotted on the logarithmic abscissa, this effect becomes apparent and asymptotically approaches a nonzero value of about 30 dB in the middle panels. Fundamentally, the result suggests a problem with using the linear form of the AHAAH model, which for low exposure levels, the discrimination may be poor.

In Figure 11 to Figure 16, the nonlinear curve fits of PTS data against the exposure level generally have the same ranking of the coefficient of determination for the six Hazard Indices. The MIL-STD 1474D consistently exhibited the lowest r^2 ranging from 0.003 to 0.128. For all frequencies, LAeq8hr had the highest r^2 ranging from 0.174 at 500 Hz to 0.268 at 2000 Hz. The unwarned AHAAH model had the second highest r^2 at all frequencies ranging from 0.119 at 500 Hz and 0.198 at 8000 Hz. Because the animal had recovered for several weeks following the exposures, many of the animals exhibited a recovery to near normal thresholds. Generally the plots exhibit a clustering of the PTS data in a band between -10 and +10 dB. An analysis of the baseline data from the chinchilla found that the standard deviation of the baseline was about 5 dB. Significant threshold shifts would occur for PTS greater than 2 standard deviations or about 10 to 12 dB. The same nonzero asymptotic trend for the AHAAH model is observed for the PTS data as it was for the TS0 data. However, the PTS asymptote was about 10 dB and the curve was close to the recovered thresholds for exposures between 36 and 16000 AHU for the unwarned case and between 3 and about 1200 AHU for the warned case.

B. Evaluation of Goodness of Fit

Table 4 through Table 9 present the outcomes of the linear mixed models for the *ptscat25* (Table 4), *ptsct15* (Table 5), *ts0cat25* (Table 6), *ts0cat15* (Table 7), *dBPTS* (Table 8), and *dBTS0* (Table 9). In each table the rank order of the information criteria is given for each of the three different model treatments (Exposure criterion, Exposure criterion + frequency, and Exposure criterion + dBBase). For every outcome variable (*ptscat25*, *ptscat15*, *ts0cat25*, *ts0cat15*, *dBPTS*, and *dBTS0*) the models including LAeq8hr as a predictor provided the best fit. The fits were not only the best, but in all cases the AIC and BIC were substantially better (recall smaller is better) than whatever exposure criterion was second. Raftery (1995) suggests that a difference in BIC's of 10 represents "strong evidence" for preferring one model to another (in this case preferring one exposure criterion over another). As shown in Table 10, most of the differences in BIC are in fact much greater than 10. For example, the BIC for the outcome variable *ptscat25* for LAeq8hr is 63.5 less than that for the warned AHAAH.

Other patterns are apparent regarding the exposure criteria. The Pfander and Smoorenburg criteria are always the second and third best fit, respectively, except when dBPTS was the outcome variable. The unwarned AHAAH criterion was always judged better than the warned AHAAH.

Since the LAeq8hr exposure criterion was judged best for all outcomes and all models, the details of the LAeq8hr analyses are presented in the following tables. Some interesting patterns emerge with respect to frequency (Table 11). For binary outcomes frequency has a significant effect for permanent threshold shifts (*ptscat25* and *ptscat15*), but not for the temporary threshold shifts (*ts0cat25* and *ts0cat15*). For both of the continuous outcomes frequency is significant, but for some reason the effects are reversed (negative for *dBPTS* and positive for *dBTS0*). The results shown in Table 12 are only for LAeq8hr, but the same pattern holds for other exposure criteria as can be seen by using likelihood ratio tests based on the information in Table 4 through Table 9. For example in Table 4, the log likelihood without the frequency (L_0) for Mil-Std 1474D is -1770.5 and the log likelihood with the frequency (L_1) is -

1764.8. This yields $-2\ln\left(\frac{L_0}{L_1}\right) = -2(-1770.5 + 1764.8)$ = 11.4 = X^2. Since the models differ by one parameter (the coefficient for frequency) X^2 is compared to the chi-square distribution with 1 degree of freedom, $X_\alpha(1)$. Letting α=0.001, $X_{0.001}(1)$ = 10.83 which is less than the observed X^2 = 11.4. Therefore, frequency has an effect significant at the 0.001 level. Similar tests, performed for the other exposure criteria for *ptscat25* and *ptscat15*, show a significant impact for frequency. In a similar vein, likelihood ratio tests performed for *ts0cat25* and *ts0cat15* for all exposure criteria reveal no significant effect for frequency.

The impact of baseline threshold is pronounced. As shown in Table 13 for LAeq8hr the impact of baseline threshold is highly significant for all outcome variables. In all cases the effect is negative. The effect of baseline threshold is also highly significant for the other outcome variables, as can be seen by doing likelihood ratio tests from the information in Table 4 through Table 9 in a manner similar to that done above in testing for the effect of frequency.

C. Receiver Operator Characteristic Discrimination Analysis

For the binary outcome *ptscat25* the areas under the ROC curves for the different exposure criteria (based on 6841 observations from 900 chinchillas) are given in Table 17. The differences in the AUC's are given in Table 18 and the summary of the paired comparisons for *ptscat25* are given in Table 19. Two major patterns emerge with respect to *ptscat25*. The first is that the exposure criterion Unwarned AHAAH is superior to all other exposure criteria in its ability to discriminate. Second, the exposure criterion MIL-STD 1474D is worse than all the others.

For the binary outcome *ptscat15* the areas under the ROC curves for the different exposure criteria (based on 6841 observations from 900 chinchillas) are given in Table 20. The differences in the AUC's are given in Table 21 and the summary of the paired comparisons is shown in Table 22. The overall results for *ptscat15* are the same as for *ptscat25*. As before, the two most obvious patterns are that Unwarned AHAAH is superior to all other exposure criteria and MIL-STD 1474D is inferior to all other exposure criteria.

For the binary outcome *ts0cat25* the areas under the ROC curves for the different exposure criteria (based on 4162 observations from 903 chinchillas) are given in Table 23. The differences in AUC's are given in Table 24 and the summary of the paired comparisons is given in Table 25. When *ts0cat25* is the outcome variable three patterns are evident. As before, the MIL-STD 1474D exposure criterion is inferior to all others. Secondly the Pfander and Smoorenburg criteria are not significantly different from each other, yet are inferior to the Unwarned AHAAH, Warned AHAAH, and LAeq8hr criteria. Finally, the Unwarned AHAAH, Warned AHAAH, and LAeq8hr criteria are not significantly different from each other.

For the binary outcome *ts0cat15* the areas under the ROC curves for the different exposure criteria (based on 4162 observations from 903 chinchillas) are given in Table 26. The differences in AUC's are given in Table 27 and the summary of the paired comparisons is given in Table 28. When *ts0cat15* is the outcome variable the results are similar to those for *ts0cat25*, but not exactly the same. Again, MIL-STD 1474D is inferior to all other exposure criteria. Also, as before, the Pfander and Smoorenburg criteria are not significantly different from each other, but are both inferior to the Unwarned AHAAH, Warned AHAAH, and LAeq8hr criteria. The difference with *ts0cat15* is that the LAeq8hr is superior to the Warned

AHAAH. Thus, the LAeq8hr criterion is superior in four out of five comparisons and the Warned AHAAH and Unwarned AHAAH in three out of five.

One might wonder why the comparison of ROC curves favors different criteria for different outcomes (Unwarned AHAAH for two *ptscat25* and *ptscat15*, none for *ts0cat25*, and LAeq8hr for *ts0cat15*), whereas in the previous section the results indicate that, for all of the outcome variables, the LAeq8hr exposure criterion showed the best fit to the data. The answer may lie in the fact that different things were being assessed. In the previous section, calibration (goodness of fit) was being evaluated and in this section discrimination was being assessed. Further, as Hosmer & Lemeshow (2000, pp. 162-163) point out, a model may have a poor fit to the data, but still provide good discrimination. Hosmer & Lemeshow (2000) suggest that a model should be evaluated both in terms of calibration and discrimination.

IV. Discussion

A. LAeq8hr vs. AHAAH model

One of the major objectives of this study was to apply a similar analysis to the chinchilla data as Murphy et al., (2009) applied to the Albuquerque Blast Overpressure walkup study. In the BOP study, the data consisted of identifying the soldier participants that suffered a temporary threshold shift. When the participants exhibited any blast related threshold shifts or sequelae (i.e. petechiae, reddening of the inside of the throat) they were removed from the study or were restricted from participating in higher energy exposures. For the auditory portion, any participant that exhibited a threshold shift of 25 dB or more was restricted from higher energy exposures and was counted as a failure for the particular exposure cell. For participants that suffered a TTS between 15 and 25 dB, they were treated as a conditional failure and were moved to a less energetic exposure with the possibility to progress to higher energies and numbers of shots. In Murphy et al., (2009), the audiometric failure was established for TTS $\geq$ 25 dB and audiometric and conditional failure were established for 15 dB $\leq$ TTS < 25 dB. For the chinchilla analysis, the failures amounted to a categorical classification (*ptscat25*, *ptscat15*, *ts0cat25* and *ts0cat15*) for each exposure cell. A further difference between the BOP analysis and the chinchilla data is the fact that the animals were exposed to only one condition where as the soldiers progressed through a matrix of exposure conditions from low energy and low impulse counts to the highest energies and 100 shots. The BOP study typically had 40 to 60 persons in an exposure cell, whereas the chinchilla study had 6 to 10 animals exposed to a particular condition. Because the present analysis does not distinguish between different interpeak intervals, the number of animals for 10 and 100 impulses was effectively tripled. For those conditions where the interpeak interval was varied (6, 60 or 600 seconds), the effective exposure levels had three times as many animals. Thus from a statistical perspective, the chinchilla data are better segregated and the confounding effects of multiple exposure conditions are absent.

The evaluations of the different exposure criteria for the animals did not produce the results as expected. Ideally, more than one stimulus waveform was desired for each exposure cell. Instead, only 50 waveforms were provided in the database which was cross-referenced to each of the 137 exposure conditions. Thus the variance of the waveforms could not be assessed[4] and the effect of a particular feature in a given waveform upon any exposure criterion could not be determined. Price has emphasized on several occasions that some

[4] In the review of this manuscript, Dr. Hamernik noted, "The variance of the waveforms was *very* small!" From our experience with shock tubes, the ranges peak impulse sound pressure levels are about 2 dB. The ranges peak impulse levels for speakers is about 1 dB.

feature of the waveform (e.g. a wiggle on the decaying slope of the initial impulse) was responsible for a significant portion of the estimated hazard when evaluated with the AHAAH model. However, the waveforms permitted a comparison of the AHAAH spectra with A-weighted spectra. Figure 2 through Figure 6 illustrate the similarity of the basilar membrane response predicted by the AHAAH model with the A-weighted spectra. The narrow band noises evaluated in Stimuli 18 through 40 have a similar spectral location. The general trend for spectral separation appears to be better in the A-weighting approach than it is in the AHAAH model when examined on the decibel scale for the AHAAH model. However, if viewed in the linear AHAAH scale, the localization of the predicted damage would be more pronounced (the logarithmic axis of the plots will compress large differences). The AHAAH model treats the exposure as the summation of the square of the displacement of the basilar membrane, which is effectively the linear version of the energy seen at any particular segment. LAeq8hr is also an energy-base metric, however, it is expressed as a logarithmic quantity. In other words, the two metrics are quite similar, but expressed differently.

Finally, one way that could reconcile the differences between AHAAH and LAeq8hr would be to determine the motion of the stapes and compare the energy and maximum amplitudes derived from each method. The analysis of the cochlear model in the AHAAH could easily be conducted through the use of narrow band filters or through a wavelet model as has been proposed by Zhu et al. (2009). The real difference between the AHAAH and LAeq8hr is the treatment of the middle ear reflex and stapes suspension. The annular ligament nonlinearity is unique to the AHAAH model and suggests limitations for the amount of energy entering the cochlea.

B. Goodness-of-Fit and Discrimination

The calibration analysis of goodness-of-fit demonstrated that the LAeq8hr metric provided the best fit to the data across the six different treatments of the data for temporary and permanent threshold shift categorical and continuous models. Whereas, the ROC analysis demonstrated that the Unwarned AHAAH model had the best discrimination for the permanent threshold shift data while the LAeq8hr had the best discrimination for the temporary threshold shift data. For the analysis of the *ptscat25* and *ptscat15* data (Tables 19 and 22) the difference between the warned AHAAH and LAeq8hr models was not statistically significant. Furthermore, for the *ts0cat25* data (Table 25) the differences between the LAeq8hr and both the warned and unwarned AHAAH models were not statistically significant. Similarly the difference between the LAeq8hr and unwarned AHAAH model was not statistically significant for the *ts0cat15* data. Since these analyses were performed for chinchilla and not for humans, the application of any damage risk criterion must be cross-validated with similar human data from unprotected impulse noise exposures.

From the nonlinear curve fit analysis, the regression that was used with the AHAAH model is flawed because the hazard cannot be less than zero and when used in the linearized equation, yielded a nonzero intercept. Thus the AHAAH metric should be used in a decibel form. This finding is consistent with Patterson and Ahroon (2004) who examined the 95% confidence limits for the MIL-STD 1474D and AHAAH models and with that from Murphy et al., (2009) who found a similar nonzero intercept for the AHAAH models in the BOP analysis. As can be seen from the curves displayed in Figure 10, the logarithmic display is quite similar to that observed for the LAeq8hr presentation. In fact, when nonlinear regression was applied to the decibel form of the AHAAH, the coefficient of determination was considerably improved and comparable to that of the LAeq8hr coefficient. The present analysis focused on evaluating the AHAAH model in the form presented as a replacement for the MIL-STD 1474D.

The Hamernik et al., (1998a, 1998b) analysis of the chinchilla data demonstrated that the P weighting or a variant would improve the fit of the threshold shift data to the exposure criteria. Thus the LAeq8hr could be improved by applying a modified weighting function. However, the form of the weighting function applicable to humans is unclear. A possible future effort could apply different weighting filters to the Albuquerque BOP data to compare goodness-of-fit and discrimination.

The exposures evaluated in this study did not cover the wide range of conditions that might be necessary to cover the parameter space where the AHAAH model's nonlinearity affects the result. In most cases, the difference between Unwarned and Warned AHAAH model were approximately 10 to 15 dB on the logarithmic scale. The exposures likely reflect the effect of the middle ear muscles and probably not the nonlinear stapes suspension. Since the exposures were conducted during a period that preceded the proposed use of the AHAAH model as a damage risk criterion, this situation could not have been foreseen. In Price's studies of cats and impulse exposures, the exposure to 105 mm howitzer impulses and M16 rifle impulses begins to explore the wider parameter space (Price and Wansack, 1987).

Since the LAeq8hr does a better job of discriminating a temporary threshold shift, it should be considered for evaluating battlefield exposures and whether situational awareness will be adversely affected by using a particular weapons system. Although LAeq8hr provides a better fit for the permanent outcomes than the unwarned AHAAH model and warned AHAAH model, the better discrimination of the unwarned AHAAH model suggests it may have utility to predict the long-term effects on hearing due to impulsive exposures. The long-term effects that are observed in this study were collected for an acute noise exposure in an animal model and are not the same as a career's worth of exposure to high-level noise.

According to this study, Hamernik et al., (1998a, 1998b) and Murphy et al., (2009), the current MIL-STD 1474D performs poorly relative to the LAeq8hr and the AHAAH model in predicting hazards associated with impulsive noise exposure. MIL-STD 1474D consistently provided the poorest curve fit, the poorest goodness of fit, and the poorest discrimination. Therefore the Army should consider replacing the MIL-STD 1474D. The use of the LAeq8hr would harmonize the criteria with that used in Europe (DTAT, 1983). The LAeq8hr can be readily measured by off-the-shelf equipment with slight changes in the microphone configurations and preamplifiers. LAeq8hr has a further advantage of relating directly to exposure damage risk criteria for continuous noise and complex noise exposures (ISO, 1990; ISO, 2009). Exposures to impulsive noise may also be harmonized with continuous or complex noise with newer approaches that consider weighting for kurtosis (Davis et al., 2009; Zhao et al., 2009; Goley, 2010; Goley et al., 2011) or other metrics based upon analytic wavelet analysis (Zhu et al., 2009). MIL-STD 1474D relies upon the peak level of waveform and an estimate of the reverberant decay of the waveform. While peak measures seem easy to collect, they are difficult to accurately collect and require a careful effort to correctly orient the microphones and to avoid confounding effects of diffraction at the microphone location. Similarly, the AHAAH model requires as much effort to capture the waveform accurately: avoiding inadequate sampling rates, minimizing diffraction effects, choosing correct filtering to avoid ringing. The position of the microphone relative to the source and any reflective surfaces may significantly affect the evaluation of the AHAAH model. Thus, these factors suggest that the most parsimonious choice would be to use the LAeq8hr. Diffraction effects, filtering, sampling rate and general ability to complete the measurement quickly favor the LAeq8hr criteria.

V. Conclusions and Recommendations

According to the statistical analysis of the chinchilla data using nonlinear curve fit analysis and linear mixed models, LAeq8hr provided the best calibration (goodness of fit) and excellent discrimination for the temporary threshold shift data. The AHAAH model did not yield the best fits to the chinchilla threshold shift data when examined in the statistical modeling. The Unwarned AHAAH model provided the best discrimination for the permanent threshold shift data. The MIL-STD 1474D consistently yielded the poorest goodness of fit and the worst discrimination.

1. The Army should strongly consider replacing the current MIL-STD 1474D with the LAeq8hr metric for evaluation of hazardous noise produced by military equipment and weapons systems.

2. Future research should focus on developing the chinchilla AHAAH model so that a species appropriate AHAAH model might be used to reevaluate this data.

3. The AHAAH model should be reformulated to output a logarithmic exposure level rather than the linear metric currently provided.

4. The LAeq8hr metric might perform even better if a different weighting function (P, P1, P2 or R) were developed appropriate for a human.

5. Future blast exposure studies need to carefully map out the range of potential effects for competing hazard indices and expose sufficiently large numbers of animals to gain statistical power.

VI. References

A. Agresti. Categorical Data Analysis, 2nd Ed. Wiley, New Jersey 2002.

W.A. Ahroon, R.P. Hamernik, and S-F. Lei. The effects of reverberant blast waves on the auditory system. J. Acoust. Soc. Am. 100:2247-2257, 1996.

P.C. Chan, K.C. Ho, K.K. Kan, J.H. Stuhmiller, and M.M. Mayorga. Evaluation of impulse noise criteria using human volunteer data. J. Acoust. Soc. Am. 110:1967-1975, 2001.

P.C. Chan and K.C. Ho. Impulse Noise Injury Model, Technical Report: J0910-07-330, Contract No. M67854-05-D-5110 DO 002, 2005.

R.R.A. Coles, G.A. Garinther, G.C. Hodge, C.R. and Rice. U.S. Army Technical Memorandum 13-67 Criteria for Assessing Hearing Damage Risk from Impulse-Noise Exposure (AMCMS Code 5011.11.84100). Aberdeen Proving Ground, MD. Human Engineering Laboratories, 1967.

A. Dancer, K. Buck, P. Hamery, and G. Parmentier. Hearing protection in the military environment. Noise and Health 5:1-15, 1999.

A. Dancer. LAeq8hr: An effective DRC for Weapon Noises. Presented in NIOSH Symposium on Impulsive Noise, 2003.

R.I. Davis, W. Qiu, and R.P. Hamernik. Role of the kurtosis statistic in evaluating complex noise exposures for the protection of hearing. Ear Hear. 30:628-634, 2009.

Direction Technique des Armements Terrestres. Recommendations on evaluating the possible harmful effects of noise on hearing. AT-83/27/28. (Etablissement Technique de Bourges, Bourges), 1983.

J.J. Earshen. The Noise Manual, 5th Edition. pp. 52-55 and pp. 69-74. (AIHA Press, Fairfax), 2003.

G.M. Fitzmaurice, N.M. Laird, and J.H. Ware. Applied Longitudinal Analysis. (Wiley, New Jersey), 2004.

G.S. Goley. Investigation and Improvement of Occupational and Military Noise Exposure Guidelines: Evaluation of Existing and Modified Noise Exposure metrics Using Historical Animal Data. M.S. Thesis, Mech. Eng. Dept., University of Cincinnati, 2010.

G.S. Goley, W.J. Song, and J.H. Kim. Kurtosis corrected sound pressure level as a noise metric for risk assessment of occupational noises, J. Acoust. Soc. Am. 129(3): 1475-1481, 2011.

R.P. Hamernik, J.H. Patterson Jr., and R.J. Salvi. The effect of impulse intensity and the number of impulses on hearing and cochlear pathology in the chinchilla. J. Acoust. Soc. Am. 81:1118-1129, 1987.

R.P. Hamernik, J.H. Patterson, and W.A. Ahroon. Use Of Animal Data In The Development Of A Human Auditory Hazard Criterion For Impulse Noise (Part 1). Final Technical Report 950342 , Contract No. DAMD17-96-C-6007, 1998.

R.P. Hamernik, J.H. Patterson, and W.A. Ahroon. A Health Hazard Assessment For Blast Overpressure Exposures: Use Of Animal Test Data In The Development Of A Human Auditory Hazard Criterion For Impulse Noise (Part 2). Final Technical Report 950342, Contract No. DAMD17-96-C-6007, 1998.

F.E. Harrell Jr., K.L. Lee, and D.B. Mark. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in Medicine 15:361-387, 1996.

D.W. Hosmer and S. Lemeshow. Applied Logistic Regression, 2nd Ed. Wiley, New York, 2000.

ISO 1999. Acoustics – Determination of occupational noise exposure and estimation of noise-induced hearing impairment. International Organization for Standardization, Geneva, 1990.

ISO 9612. Acoustics – Determination of occupational noise exposure – Engineering method. International Organization for Standardization, Geneva, 2008.

K.D. Kryter, W.D. Ward, J.D. Miller, and D.H. Eldredge. Hazardous exposure to intermittent and steady-state noise. J. Acoust. Soc. Am., 39:451-464, 1966.

N.M. Laird and J.H. Ware. Random-effects models for longitudinal data. Biometrics 38:963-974, 1982.

J S. Long. Regression Models For Categorical And Limited Dependent Variables. Sage Publications, Thousand Oaks, California, 1997.

MIL-STD-1474D. Noise limits from Department of Defense Design Criteria Standard. Updated 2007, 1997.

W.J. Murphy and R.R. Davis. The role of the chinchilla pinna and ear canal in electrophysiological measures of hearing thresholds. J. Acoust. Soc. Am. 103:1951-1956, 1998.

W.J. Murphy, C.A. Kardous, D.C. Byrne, and E.L. Zechmann. Auditory Risk of Hearing Loss due to Gunshot Noise Exposure, National Hearing Conservation Association Savannah GA Feb 16-17, 2007.

W.J. Murphy and R.L. Tubbs. Assessment of noise exposure for an indoor and outdoor firing range. J. Occup. Env. Hyg., 4:688-697, 2007.

W.J. Murphy, A. Khan, and P.B. Shaw. An Analysis of the Blast Overpressure Study Data Comparing Three Exposure Criteria. NIOSH Survey Report: EPHB 309-05h. DHHS-CDC-NIOSH, December, 2009.

R. Newson. Parameters behind "non-parametric" statistics: Kendall's τ_a, Somer's *D*, and median differences. The Stata Journal 1:1-20, 2001.

N.A. Obuchowski. Nonparametric analysis of clustered ROC curve data. Biometrics 53:567-578, 1997.

W. Pan. Akaike's information criterion in generalized estimating equations. Biometrics 57:120-125, 2001.

J.H. Patterson Jr., I.M. Lomba-Gautier, D.L. Curd, R.P. Hamernik, R.J. Salvi, C.E. Hargett Jr., and G. Turrentine. The effect of impulse intensity and the number of impulses on hearing and cochlear pathology in the chinchilla. USAARL Report No. 85-3. 1985.

J.H. Patterson Jr., R.P. Hamernik, C.E. Hargett, and W.A. Ahroon. An isohazard function for impulse noise. J. Acoust. Soc. Am. 93:2860-2869, 1993.

J.H. Patterson and D.L. Johnson. Temporary Threshold Shifts Produced by High Intensity Free-field Impulse Noise in Human Wearing Hearing Protection, USAARL Report 94-46, 1994.

J.H. Patterson and W.A. Ahroon. Evaluation Of An Auditory Hazard Model Using Data From Human Volunteer Studies, USAARL Report 2005-01, 2004.

F. Pfander, H. Bongartz, H. Brinkmann and H. Kietz. Danger of auditory impairment from impulse noise: A comparative study of the CHABA damage-risk criteria and those of the Federal Republic of Germany. J. Acoust. Soc. Am., 67:628-633, 1980.

G.R. Price. Relative Hazard of weapons impulses. J. Acoust. Soc. Am. 73:556-566, 1982.

G.R. Price. Weapon noise exposure of the human ear analyzed with the AHAAH model, Scientific presentation at the U.S. Army Research Laboratory, Human Research and engineering Directorate Aberdeen Proving Ground, MD 21005-5425. Available on the AHAAH Website at http//www.arl.army.mil/hred/AHAAH, 2005.

G.R. Price. Predicting mechanical damage to the organ of Corti. Hear. Res., 226:5-13, 2007a.

G.R. Price. Validation of the auditory hazard assessment algorithm for the human with impulse noise data. J. Acoust. Soc. Am., 122:2786-2802, 2007b.

G.R. Price and J.T. Kalb. Insights into hazard from intense impulses from a mathematical model of the ear. J. Acoust. Soc. Am. 90:219–227, 1991.

G.R. Price and S. Wansack. Hazard from intense midrange impulse. J. Acoust. Soc. Am., 86:2185-2191, 1989.

S. Rabe-Hesketh and A. Skrondal. Multilevel And Longitudinal Modeling Using Stata 2nd Ed. (Stata Press, College Station), 2008.

A.E. Raftery. Bayesian Model Selection In Social Research. In Sociological Methodology 26: 111-163, 1995.

P. B. Shaw. Detailed Comparison of Hearing Exposure Criteria – Using Generalized Linear Mixed Models and Linear Mixed Effects Models of Chinchilla Data. Statistical report for W.J. Murphy, 2009.

G. F. Smoorenburg. Damage risk criteria for impulse noise, in New perspectives on Noise-Induced Hearing Loss. Edited by R. Hamernik, D. Henderson, and R. Salvi. (Raven Press, New York), pp. 471-490, 1982.

G. F. Smoorenburg. Damage Risk for Low-frequency Impulse Noise. in Noise-Induced Hearing Loss, Edited by A. Dancer, D. Henderson, R. Salvi, and R. P. Hamernik. (Mosby Year Book, Saint Louis) pp. 313–324, 1992.

G.F. Smoorenburg. Risk of Hearing Loss from exposure to impulse sounds. in Reconsideration of the Effects of Impulse Noise. Research and Technology Organization AC323 (HFM)-022) TP/17, pp. 1.1-30, 2001.

W.J. Song. Study of Human Auditory System Models and Risk Assessment of Noise-induced Hearing Loss. Ph.D. Dissertation, Mech. Eng. Dept., University of Cincinnati, 2010.

W.J. Song and J. Kim. Comparative study on the responses of human middle ear models. Proceedings of NOISE-CON 2008 117:225-235, 2008.

H.H. Spoendlin. Anatomical Changes Following Various Noise Exposures, Effects of Noise Hearing. Edited by D. Henderson, R.P. Hamernik, D.S. Dosanjh, J.M. Mills, (Raven Press, New York) pp 69-79, 1976.

StataCorp. Stata Statistical Software: Release 10. College Station, TX: StataCorp LP, 2007.

R.L. Tubbs and W.J. Murphy. NIOSH Health Hazard Evaluation Report: HETA #2002-0131-2898 Fort Collins Police Services, Fort Collins, Colorado. DHHS-CDC-NIOSH, March, 2003.

H.Q., Yan, C.A. Kardous, and W.J. Murphy. NIOSH Impulsive Noise Measurement System (NIMS). proceedings of NOISE-CON 2004, Baltimore MD July 12-14, 2004.

E.L. Zechmann. Impulsive Noise Meter. National Institute for Occupational Safety and Health. URL www.mathworks.com/matlabcentral/fileexchange/18095, 2009.

Y-M. Zhao, W. Qiu, Z. Lin, S-S. Chen, X-R. Cheng, R.I. Davis, R.P. Hamernik. Application of the kurtosis statistic to the evaluation of the risk of hearing loss in workers exposed to high-level complex noise, Ear Hear., 31:527-532, 2010.

X-D. Zhu, J.H. Kim, W-J. Song, W.J. Murphy, and S-H. Song. Development of a noise metric for assessment of exposure risk to complex noises. J. Acoust. Soc. Am. 126:703-712, 2009.

VII. Tables

Table 1. Exposures Conditions for chinchillas examined in this study.

Study	Animals with Audiometric Data	Animals with Histological Data	Open Field or Enclosure
Conventional shock tube, nonreverberant	109	109	Open Field
Fast-acting valve (5”) nonreverberant	105	105	Open Field
Fast-acting valve (3.5”) nonreverberant	105	105	Open Field
Spark gap, nonreverberant	104	104	Open Field
Conventional shock tube, reverberant	135	135	Enclosure
Fast-acting valve (3.5”), reverberant	136	136	Enclosure
Narrow band impact	130	130	Open Field
290C driver 146 dB and 138 dB peak SPL	12	12	Open Field
290C driver, High peak wave	36	36	Open Field
290C driver, Low peak wave	18	18	Open Field
290C driver, 131 dB peak SPL	5	5	Open Field
USAARL Conventional shock tube, nonreverberant (unpublished)	10	26	Open Field

Table 2. Description of the types of stimuli used in the exposures

Stimulus Code	Description
1 – 3	Conventional shock tube, nonreverberant
4 – 6	Fast-acting valve (5”), nonreverberant
7 – 9	Fast-acting valve (3.5”), nonreverberant
10 – 12	Spark gap, nonreverberant
13 – 15	Conventional shock tube, reverberant
16 – 18	Fast-acting valve (3.5”), reverberant
19 – 20	260 Hz Narrow-band impact
21 – 23	775 Hz Narrow-band impact
24 – 27	1025 Hz Narrow-band impact
28 – 30	1350 Hz Narrow-band impact
31 – 34	2450 Hz Narrow-band impact
35 – 38	3550 Hz Narrow-band impact
39 – 40	2075 Hz Narrow-band impact
41 – 42	146 dB peak SPL and 138 dB peak SPL
43, 45 & 47	290C driver, High peak wave, USAARL Report 86-7
44, 46, & 48	290C driver, Low peak wave, USAARL Report 86-7
49	290C driver, 131 peak SPL, 100x , USAARL Report 85-3
50	USAARL Conventional shock tube, nonreverberant (unpublished)

Table 3. Evaluation of the exposures for the different hazard indices. The Exposure code, Stimulus code, number of impulses, and interpeak interval are shown for each of the various exposures. The evaluations

of the MIL-STD 1474D, LAeq8hr, AHAAH Unwarned and Warned conditions, Pfander and Smoorenburg hazard indices are given in the subsequent columns.

Exposure Code	Stimulus Code	Number of Impulses	Interpeak Interval (sec)	MIL-STD 1474D T_B (dB)	LAeq8 (dBA)	AHAAH Unwarned (AHU)	AHAAH Warned (AHU)	Pfander T_C (dB)	Smoorenburg T_D (dB)
1	1	1	0	157.9	68.9	36.6	4.0	156.7	159.6
2	1	10	6	162.9	78.9	366.4	39.6	166.7	169.6
3	1	10	60	162.9	78.9	366.4	39.6	166.7	169.6
4	1	10	600	162.9	78.9	366.4	39.6	166.7	169.6
5	1	100	6	167.9	88.9	3663.7	395.9	176.7	179.6
6	1	100	60	167.9	88.9	3663.7	395.9	176.7	179.6
7	1	100	600	167.9	88.9	3663.7	395.9	176.7	179.6
8	2	1	0	162.9	73.9	59.7	7.9	161.7	164.5
9	2	10	6	167.9	83.9	596.5	79.3	171.7	174.5
10	2	10	60	167.9	83.9	596.5	79.3	171.7	174.5
11	2	10	600	167.9	83.9	596.5	79.3	171.7	174.5
12	2	100	6	172.9	93.9	5965.4	792.7	181.7	184.5
13	2	100	60	172.9	93.9	5965.4	792.7	181.7	184.5
14	2	100	600	172.9	93.9	5965.4	792.7	181.7	184.5
15	3	1	0	167.7	73.5	72.5	12.5	150.8	150.8
16	3	10	6	172.7	83.5	725.4	124.5	160.8	160.8
17	3	10	60	172.7	83.5	725.4	124.5	160.8	160.8
18	3	10	600	172.7	83.5	725.4	124.5	160.8	160.8
19	3	100	6	177.7	93.5	7254.2	1245.3	170.8	170.8
20	3	100	60	177.7	93.5	7254.2	1245.3	170.8	170.8
21	3	100	600	177.7	93.5	7254.2	1245.3	170.8	170.8
22	4	1	0	152.3	63.3	74.8	16.0	141.4	142.1
23	4	10	6	157.3	73.3	748.3	160.4	151.4	152.1
24	4	10	60	157.3	73.3	748.3	160.4	151.4	152.1
25	4	10	600	157.3	73.3	748.3	160.4	151.4	152.1
26	4	100	6	162.3	83.3	7483.1	1604.3	161.4	162.1
27	4	100	60	162.3	83.3	7483.1	1604.3	161.4	162.1
28	4	100	600	162.3	83.3	7483.1	1604.3	161.4	162.1
29	5	1	0	157.9	69.9	182.2	37.9	147.3	146.9
30	5	10	6	162.9	79.9	1822.4	378.5	157.3	156.9
31	5	10	60	162.9	79.9	1822.4	378.5	157.3	156.9
32	5	10	600	162.9	79.9	1822.4	378.5	157.3	156.9
33	5	100	6	167.9	89.9	18223.5	3785.3	167.3	166.9
34	5	100	60	167.9	89.9	18223.5	3785.3	167.3	166.9
35	5	100	600	167.9	89.9	18223.5	3785.3	167.3	166.9
36	6	1	0	163.9	73.3	233.6	54.7	152.4	152.6
37	6	10	6	168.9	83.3	2336.0	546.8	162.4	162.6
38	6	10	60	168.9	83.3	2336.0	546.8	162.4	162.6
39	6	10	600	168.9	83.3	2336.0	546.8	162.4	162.6
40	6	100	6	173.9	93.3	23360.1	5468.2	172.4	172.6
41	6	100	60	173.9	93.3	23360.1	5468.2	172.4	172.6
Exposure Code	Stimulus Code	Number of Impulses	Interpeak Interval (sec)	MIL-STD 1474D T_B (dB)	LAeq8 (dBA)	AHAAH Unwarned (AHU)	AHAAH Warned (AHU)	Pfander T_C (dB)	Smoorenburg T_D (dB)
42	6	100	600	173.9	93.3	23360.1	5468.2	172.4	172.6

43	7	1	0	152.0	63.1	74.3	15.8	140.7	141.8
44	7	10	6	157.0	73.1	743.0	158.2	150.7	151.8
45	7	10	60	157.0	73.1	743.0	158.2	150.7	151.8
46	7	10	600	157.0	73.1	743.0	158.2	150.7	151.8
47	7	100	6	162.0	83.1	7429.9	1582.2	160.7	161.8
48	7	100	60	162.0	83.1	7429.9	1582.2	160.7	161.8
49	7	100	600	162.0	83.1	7429.9	1582.2	160.7	161.8
50	8	1	0	158.0	69.1	171.8	35.1	147.4	146.5
51	8	10	6	163.0	79.1	1717.5	350.6	157.4	156.5
52	8	10	60	163.0	79.1	1717.5	350.6	157.4	156.5
53	8	10	600	163.0	79.1	1717.5	350.6	157.4	156.5
54	8	100	6	168.0	89.1	17175.1	3505.8	167.4	166.5
55	8	100	60	168.0	89.1	17175.1	3505.8	167.4	166.5
56	8	100	600	168.0	89.1	17175.1	3505.8	167.4	166.5
57	9	1	0	163.5	73.2	251.7	55.3	152.4	152.1
58	9	10	6	168.5	83.2	2516.6	552.9	162.4	162.1
59	9	10	60	168.5	83.2	2516.6	552.9	162.4	162.1
60	9	10	600	168.5	83.2	2516.6	552.9	162.4	162.1
61	9	100	6	173.5	93.2	25166.1	5529.0	172.4	172.1
62	9	100	60	173.5	93.2	25166.1	5529.0	172.4	172.1
63	9	100	600	173.5	93.2	25166.1	5529.0	172.4	172.1
64	10	1	0	151.2	58.9	47.2	7.4	138.6	140.8
65	10	10	6	156.2	68.9	471.7	74.0	148.6	150.8
66	10	10	60	156.2	68.9	471.7	74.0	148.6	150.8
67	10	10	600	156.2	68.9	471.7	74.0	148.6	150.8
68	10	100	6	161.2	78.9	4716.6	739.9	158.6	160.8
69	10	100	60	161.2	78.9	4716.6	739.9	158.6	160.8
70	10	100	600	161.2	78.9	4716.6	739.9	158.6	160.8
71	11	1	0	155.1	63.6	103.6	16.5	143.4	145.3
72	11	10	6	160.1	73.6	1036.1	165.4	153.4	155.3
73	11	10	60	160.1	73.6	1036.1	165.4	153.4	155.3
74	11	10	600	160.1	73.6	1036.1	165.4	153.4	155.3
75	11	100	6	165.1	83.6	10360.9	1654.3	163.4	165.3
76	11	100	60	165.1	83.6	10360.9	1654.3	163.4	165.3
77	11	100	600	165.1	83.6	10360.9	1654.3	163.4	165.3
78	12	1	0	161.3	68.8	264.4	45.4	148.5	150.8
79	12	10	6	166.3	78.8	2644.0	454.4	158.5	160.8
80	12	10	60	166.3	78.8	2644.0	454.4	158.5	160.8
81	12	10	600	166.3	78.8	2644.0	454.4	158.5	160.8
82	12	100	6	171.3	88.8	26439.6	4543.5	168.5	170.8
83	12	100	60	171.3	88.8	26439.6	4543.5	168.5	170.8
84	12	100	600	171.3	88.8	26439.6	4543.5	168.5	170.8
85	13	1	0	161.6	70.5	427.9	136.9	155.7	154.4
86	13	10	60	166.6	80.5	4279.1	1368.7	165.7	164.4
87	13	100	60	171.6	90.5	42790.8	13686.7	175.7	174.4
88	14	1	0	162.0	75.6	695.2	233.4	161.3	158.9

Exposure Code	Stimulus Code	Number of Impulses	Interpeak Interval (sec)	MIL-STD 1474D T_B (dB)	LAeq8 (dBA)	AHAAH Unwarned (AHU)	AHAAH Warned (AHU)	Pfander T_C (dB)	Smoorenburg T_D (dB)
89	14	10	60	167.0	85.6	6951.5	2333.9	171.3	168.9

90	14	100	60	172.0	95.6	69515.3	23339.0	181.3	178.9
91	15	1	0	171.7	81.9	1348.2	620.9	164.5	164.4
92	15	10	60	176.7	91.9	13481.7	6209.2	174.5	174.4
93	15	100	60	181.7	101.9	134816.9	62091.7	184.5	184.4
94	16	1	0	154.4	73.0	481.9	105.5	145.5	146.7
95	16	10	60	159.4	83.0	4818.8	1054.9	155.5	156.7
96	16	100	60	164.4	93.0	48188.2	10549.1	165.5	166.7
97	17	1	0	158.9	78.2	867.2	282.4	154.0	151.4
98	17	10	60	163.9	88.2	8672.0	2823.6	164.0	161.4
99	17	100	60	168.9	98.2	86720.3	28235.8	174.0	171.4
100	18	1	0	164.9	85.3	1582.8	732.3	164.4	156.5
101	18	10	60	169.9	95.3	15827.6	7322.8	174.4	166.5
102	18	100	60	174.9	105.3	158276.1	73227.8	184.4	176.5
103	19	100	3	156.5	80.0	1888.4	88.4	165.4	168.6
104	20	100	3	163.5	87.0	3882.9	315.7	172.4	175.6
105	21	100	3	150.6	79.0	4858.3	193.1	158.5	161.8
106	22	100	3	155.5	84.0	8885.7	506.9	163.5	166.6
107	23	100	3	160.6	89.0	13989.3	1169.9	168.5	171.8
108	24	100	3	144.9	75.4	5751.3	221.0	152.9	154.2
109	25	100	3	149.8	80.2	10665.2	584.0	157.8	159.1
110	26	100	3	154.8	85.1	18731.0	1394.7	162.7	164.1
111	27	100	3	159.0	89.1	28207.2	2574.2	166.9	168.2
112	28	100	3	145.0	75.8	7980.1	393.9	153.1	156.5
113	29	100	3	149.9	80.7	15754.7	1018.1	157.9	161.1
114	30	100	3	154.8	85.5	28329.3	2383.9	162.8	166.0
115	31	100	3	145.2	77.3	15378.7	1517.9	154.5	157.2
116	32	100	3	150.6	82.0	38150.0	3925.8	159.6	162.1
117	33	100	3	154.8	86.8	84518.4	9618.6	164.1	163.5
118	34	100	3	159.7	91.7	162982.2	21072.8	169.0	168.4
119	35	100	3	140.1	71.3	4922.3	779.0	148.6	149.1
120	36	100	3	145.1	76.3	15121.9	2376.9	153.7	154.1
121	37	100	3	150.5	81.0	39947.1	5867.8	158.4	159.0
122	38	100	3	154.4	85.5	93469.5	13435.6	162.4	163.6
123	39	100	3	149.3	80.7	16216.4	1371.7	158.2	160.8
124	40	100	3	154.2	85.6	35451.9	3464.5	163.1	165.7
125	41	100	3	159.0	86.8	15944.6	1967.5	164.3	165.4
126	42	100	3	150.8	78.7	5654.6	504.6	156.1	157.3
127	43	100	3	159.9	87.8	17768.2	2265.4	165.2	166.4
128	44	100	3	157.0	87.5	22307.9	4271.1	165.6	170.5
129	45	100	3	152.1	79.9	6696.8	628.1	157.4	158.5
130	46	100	3	149.1	79.6	5775.4	893.9	157.6	162.6
131	47	100	3	148.1	76.0	3822.4	303.9	153.5	154.6
132	48	100	3	145.2	75.7	3565.1	385.0	153.7	158.8
133	49	100	3	144.0	71.9	2104.6	132.8	149.4	150.5
134	45	10	3	147.1	69.9	669.7	62.8	147.4	148.5
135	43	1	3	149.9	67.8	177.7	22.7	145.2	146.4
Exposure Code	Stimulus Code	Number of Impulses	Interpeak Interval (sec)	MIL-STD 1474D T_B (dB)	LAeq8 (dBA)	AHAAH Unwarned (AHU)	AHAAH Warned (AHU)	Pfander T_C (dB)	Smoorenburg T_D (dB)
136	43	10	3	154.9	77.8	1776.8	226.5	155.2	156.4

137	50	12	20	165.8	83.3	581.8	125.4	167.4	172.1

Table 4. Results of modeling binary outcome *ptscat25* as function of different exposure criteria, frequency, and baseline threshold with generalized linear mixed models. The six models of a given type (e.g. logit() = xB = exposure criterion) are compared using AIC and BIC (smaller is better in both cases) and ranked. Thus for all three types of models the LAeq8hr has the best AIC and BIC and the Warned AHAAH the worst (milstd = MIL-STD 1474D, un_aha=Unwarned AHAAH, wa_aha=Warned AHAAH, laeq8=LAeq8hr, dBBase= baseline threshold, and freq=frequency).

Outcome Variable	Model (xb)	AIC	Rank of IC			BIC (N = 900)	Log likelihood
ptscat25	milstd + freq	3539.6	5			3563.6	-1764.8
ptscat25	milstd	3548.9		5		3568.1	-1770.5
ptscat25	milstd + dBBase	3352.6			5	3376.6	-1671.3
ptscat25	un_aha + freq	3524.0	4			3548.0	-1757.0
ptscat25	un_aha	3532.4		4		3551.6	-1762.2
ptscat25	un_aha + dBBase	3340.5			4	3364.5	-1665.3
ptscat25	wa_aha + freq	3539.9	6			3563.9	-1764.9
ptscat25	wa_aha	3549.2		6		3568.4	-1770.6
ptscat25	wa_aha + dBBase	3355.9			6	3379.9	-1672.9
ptscat25	laeq8 + freq	3476.6	1			3500.6	-1733.3
ptscat25	laeq8	3485.7		1		3504.9	-1738.8
ptscat25	laeq8+ dBBase	3289.4			1	3313.4	-1639.7
ptscat25	pfander + freq	3516.0	2			3540.0	-1753.0
ptscat25	pfander	3524.8		2		3544.0	-1758.4
ptscat25	pfander + dBBase	3330.7			2	3354.7	-1660.3
ptscat25	smoorenburg + freq	3522.3	3			3546.3	-1756.1
ptscat25	smoorenburg	3530.8		3		3550.0	-1761.4
ptscat25	smoorenburg + dBBase	3334.2			3	3358.2	-1662.1

Table 5. Results of modeling binary outcome *ptscat15* as function of different exposure criteria, frequency, and baseline threshold with generalized linear mixed models. The six models of a given type (e.g. logit() = xB = exposure criterion) are compared using AIC and BIC (smaller is better in both cases) and ranked. Thus for all three types of models the LAeq8hr has the best AIC and BIC and the MIL-STD 1474D the worst (milstd = MIL-STD 1474D, un_aha=Unwarned AHAAH, wa_aha=Warned AHAAH, laeq8=LAeq8hr, dBBase= baseline threshold, and freq=frequency).

Outcome Variable	Model (xb)	AIC	Rank of IC			BIC (N = 900)	Log likelihood
ptscat15	milstd + freq	4455.3	6			4479.3	-2222.7
ptscat15	milstd	4500.6		6		4519.8	-2246.3
ptscat15	milstd + dBBase	4345.1			6	4369.1	-2167.6
ptscat15	un_aha + freq	4436.9	4			4460.9	-2213.4
ptscat15	un_aha	4480.4		4		4499.6	-2236.2
ptscat15	un_aha + dBBase	4328.1			4	4352.2	-2159.1
ptscat15	wa_aha + freq	4452.0	5			4476.1	-2221.0
ptscat15	wa_aha	4496.8		5		4516.0	-2244.4
ptscat15	wa_aha + dBBase	4343.3			5	4367.3	-2166.7
ptscat15	laeq8 + freq	4382.3	1			4406.3	-2186.2
ptscat15	laeq8	4427.1		1		4446.4	-2209.6
ptscat15	laeq8+ dBBase	4272.0			1	4296.0	-2131.0
ptscat15	pfander + freq	4424.9	2			4448.9	-2207.5
ptscat15	pfander	4469.2		2		4488.4	-2230.6
ptscat15	pfander + dBBase	4313.5			2	4337.5	-2151.7
ptscat15	smoorenburg + freq	4430.4	3			4454.4	-2210.2
ptscat15	smoorenburg	4475.1		3		4494.3	-2233.6
ptscat15	smoorenburg + dBBase	4318.8			3	4342.9	-2154.4

Table 6. Results of modeling binary outcome *ts0cat25* as function of different exposure criteria, frequency, and baseline threshold with generalized linear mixed models. The six models of a given type (e.g. logit() = xB = exposure criterion) are compared using AIC and BIC (smaller is better in both cases) and ranked. Thus for all three types of models the LAeq8hr has the best AIC and BIC and the Warned AHAAH the worst.

Outcome Variable	**Model (xb)**	**AIC**	**Rank of IC**			**BIC (N = 903)**	**Log likelihood**
ts0cat25	milstd + freq	3360.1	5			3384.2	-1675.070
ts0cat25	milstd	3358.2		5		3377.5	-1675.116
ts0cat25	milstd + dBBase	3203.1			4	3227.1	-1596.552
ts0cat25	un_aha + freq	3346.4	4			3370.5	-1668.213
ts0cat25	un_aha	3344.5		4		3363.7	-1668.229
ts0cat25	un_aha + dBBase	3204.0			5	3228.0	-1596.998
ts0cat25	wa_aha + freq	3362.0	6			3386.0	-1675.980
ts0cat25	wa_aha	3360.0		6		3379.2	-1676.001
ts0cat25	wa_aha + dBBase	3214.5			6	3238.5	-1602.241
ts0cat25	laeq8 + freq	3269.0	1			3293.0	-1629.333
ts0cat25	laeq8	3266.8		1		3286.0	-1629.375
ts0cat25	laeq8+ dBBase	3117.4			1	3141.4	-1553.685
ts0cat25	pfander + freq	3316.7	2			3340.7	-1653.336
ts0cat25	pfander	3314.8		2		3334.0	-1653.375
ts0cat25	pfander + dBBase	3163.4			2	3187.4	-1576.696
ts0cat25	smoorenburg + freq	3323.3	3			3347.3	-1656.652
ts0cat25	smoorenburg	3321.4		3		3340.6	-1656.681
ts0cat25	smoorenburg + dBBase	3171.0			3	3195.1	-1580.519

Table 7. Results of modeling binary outcome *ts0cat15* as function of different exposure criteria, frequency, and baseline threshold with generalized linear mixed models. The six models of a given type (e.g. logit() = xB = exposure criterion) are compared using AIC and BIC (smaller is better in both cases) and ranked. Thus for all three types of models the LAeq8hr has the best AIC and BIC and the MIL-STD 1474D the worst for two of the model types (exposure criterion and exposure criterion + frequency) and Warned AHAAH the worst for exposure criterion + baseline threshold.

Outcome Variable	Model (xb)	AIC	Rank of IC			BIC (N = 903)	Log likelihood
ts0cat15	milstd + freq	3156.6	6			3180.6	-1573.279
ts0cat15	milstd	3155.1		6		3174.4	-1573.570
ts0cat15	milstd + dBBase	2995.4			5	3019.5	-1492.720
ts0cat15	un_aha + freq	3131.6	4			3155.6	-1560.784
ts0cat15	un_aha	3130.4		4		3149.6	-1561.177
ts0cat15	un_aha + dBBase	2989.3			4	3013.3	-1489.646
ts0cat15	wa_aha + freq	3156.1	5			3180.1	-1573.054
ts0cat15	wa_aha	3154.8		5		3174.0	-1573.401
ts0cat15	wa_aha + dBBase	3005.8			6	3029.8	-1497.906
ts0cat15	laeq8 + freq	3063.7	1			3087.7	-1526.846
ts0cat15	laeq8	3062.3		1		3081.5	-1527.143
ts0cat15	laeq8+ dBBase	2910.8			1	2934.8	-1450.384
ts0cat15	pfander + freq	3116.3	2			3140.3	-1553.133
ts0cat15	pfander	3114.9		2		3134.1	-1553.438
ts0cat15	pfander + dBBase	2960.3			2	2984.3	-1475.135
ts0cat15	smoorenburg + freq	3124.0	3			3148.0	-1556.979
ts0cat15	smoorenburg	3122.6		3		3141.8	-1557.310
ts0cat15	smoorenburg + dBBase	2970.3			3	2994.3	-1480.133

Table 8. Results of modeling continuous outcome *dBPTS* as function of different exposure criteria, frequency, and baseline threshold with linear mixed models. The six models of a given type (e.g. y = xB = exposure criterion) are compared using AIC and BIC (smaller is better in both cases) and ranked. Thus for all three types of models the LAeq8hr has the best AIC and BIC and the MIL-STD 1474D has the worst.

Outcome Variable	Model (xb)	AIC	Rank of IC			BIC (N = 900)	Log likelihood
dBPTS	milstd + freq	51280.0	6			51308.8	-25634.0
dBPTS	milstd	51284.7		6		51308.8	-25637.4
dBPTS	milstd + dBBase	50712.0			6	50740.8	-25350.0
dBPTS	un_aha + freq	51236.6	2			51265.4	-25612.3
dBPTS	un_aha	51241.0		2		51265.0	-25615.5
dBPTS	un_aha + dBBase	50675.0			2	50703.8	-25331.5
dBPTS	wa_aha + freq	51263.4	5			51292.2	-25625.7
dBPTS	wa_aha	51268.0		5		51292.0	-25629.0
dBPTS	wa_aha + dBBase	50699.4			5	50728.2	-25343.7
dBPTS	laeq8 + freq	51214.8	1			51243.6	-25601.4
dBPTS	laeq8	51219.5		1		51243.6	-25604.8
dBPTS	laeq8+ dBBase	50646.9			1	50675.7	-25317.4
dBPTS	pfander + freq	51253.1	3			51281.9	-25620.5
dBPTS	pfander	51257.7		3		51281.7	-25623.8
dBPTS	pfander + dBBase	50685.4			3	50714.2	-25336.7
dBPTS	smoorenburg + freq	51258.8	4			51287.6	-25623.4
dBPTS	smoorenburg	51263.3		4		51287.3	-25626.6
dBPTS	smoorenburg + dBBase	50691.6			4	50720.4	-25339.8

Table 9. Results of modeling continuous outcome *dBTS0* as function of different exposure criteria, frequency, and baseline threshold with linear mixed models. The six models of a given type (e.g. y = xB = exposure criterion) are compared using AIC and BIC (smaller is better in both cases) and ranked. Thus for all three types of models the LAeq8hr has the best AIC and BIC and the Warned AHAAH has the worst.

Outcome Variable	Model (xb)	AIC	Rank of IC			BIC (N = 903)	Log likelihood
dBTS0	milstd + freq	35060.9	5			35089.8	-17524.5
dBTS0	milstd	35076.5		5		35100.5	-17533.3
dBTS0	milstd + dBBase	34416.8			4	34445.7	-17202.4
dBTS0	un_aha + freq	35055.5	4			35084.3	-17521.7
dBTS0	un_aha	35071.9		4		35095.9	-17530.9
dBTS0	un_aha + dBBase	34436.8			5	34465.7	-17212.4
dBTS0	wa_aha + freq	35068.4	6			35097.2	-17528.2
dBTS0	wa_aha	35084.6		6		35108.6	-17537.3
dBTS0	wa_aha + dBBase	34444.6			6	34473.5	-17216.3
dBTS0	laeq8 + freq	34956.3	1			34985.2	-17472.2
dBTS0	laeq8	34972.2		1		34996.2	-17481.1
dBTS0	laeq8+ dBBase	34329.6			1	34358.4	-17158.8
dBTS0	pfander + freq	35021.2	2			35050.0	-17504.6
dBTS0	pfander	35037.1		2		35061.1	-17513.6
dBTS0	pfander + dBBase	34389.9			2	34418.7	-17189.0
dBTS0	smoorenburg + freq	35031.8	3			35060.7	-17509.9
dBTS0	smoorenburg	35047.9		3		35072.0	-17519.0
dBTS0	smoorenburg + dBBase	34402.1			3	34430.9	-17195.0

Table 10. Amount by which BIC of other exposure criteria exceeded that of LAeq8hr.

	Outcome Variable					
Exposure Criterion	*ptscat25*	*ptscat15*	*ts0cat25*	*ts0cat15*	*dBPTS*	*dBTS0*
MIL-STD 1474D	63.2	73.4	91.5	92.9	65.2	104.3
Unwarned AHAAH	46.7	53.2	77.7	68.1	21.4	99.7
Warned AHAAH	63.5	69.6	93.2	92.5	48.4	112.4
Pfander	39.1	42.0	48.0	52.6	38.1	64.9
Smoorenburg	45.1	47.9	54.6	60.3	43.7	75.8

Table 11. Results of models for which *LAeq8hr* was the only fixed effect. For outcomes that were binary (*ptscat15, ptscat15, ts0cat25, ts0cat15)* generalized linear mixed models were used and for continuous outcomes (*dBPTS* and *dBTS0*) linear mixed effects models were used.

Outcome variable	**Effect**	**Coefficient**	**Std. Err.**	**z**	**p-value**	**95% Conf. Interval**	
ptscat25	**LAeq8hr**	.3486002	.0316757	11.01	0.000	.2865168	.4106835
	intercept	-48.83682	4.239301	-11.52	0.000	-57.1457	-40.52794
ptscat15	**LAeq8hr**	.3253777	.0296733	10.97	0.000	.2672191	.3835364
	intercept	-43.7374	3.852533	-11.35	0.000	-51.28822	-36.18657
ts0cat25	**LAeq8hr**	.2964071	.0256609	11.55	0.000	.2461127	.3467015
	intercept	-36.62919	3.242241	-11.30	0.000	-42.98386	-30.27451
ts0cat15	**LAeq8hr**	.2674114	.0231191	11.57	0.000	.2220987	.3127241
	intercept	-31.6551	2.877887	-11.00	0.000	-37.29565	-26.01454
dBPTS	**LAeq8hr**	.9417714	.0954629	9.87	0.000	.7546677	1.128875
	intercept	-107.2838	12.11803	-8.85	0.000	-131.0347	-83.53287
dBTS0	**LAeq8hr**	2.038576	.1331129	15.31	0.000	1.777679	2.299472
	intercept	-220.516	16.89194	-13.05	0.000	-253.6236	-187.4084

Table 12. Results of models with *LAeq8hr* and *frequency* as fixed effects. For outcomes that were binary (*ptscat15, ptscat15, ts0cat25, ts0cat15)* generalized linear mixed models were used and for continuous outcomes (*dBPTS* and *dBTS0*) linear mixed effects models were used.

Outcome variable	Effect	Coefficient	Std. Err.	z	p-value	95% Conf. Interval	
ptscat25	LAeq8hr	.3491688	.0315312	11.07	0.000	.2873688	.4109689
	frequency	-.0000366	.0000111	-3.31	0.001	-.0000584	-.0000149
	intercept	-48.73959	4.229455	-11.52	0.000	-57.02917	-40.45001
ptscat15	LAeq8hr	.3277724	.0311993	10.51	0.000	.2666229	.388922
	frequency	-.000067	$9.99x10^{-06}$	-6.70	0.000	-.0000866	-.0000474
	intercept	-43.73716	4.073216	-10.74	0.000	-51.72051	-35.7538
ts0cat25	LAeq8hr	.2964185	.0256518	11.56	0.000	.2461419	.346695
	frequency	$-5.13x10^{-06}$	.0000177	-0.29	0.771	-.0000397	.0000295
	intercept	-36.61408	3.241335	-11.30	0.000	-42.96698	-30.26118
ts0cat15	LAeq8hr	.2675465	.0231525	11.56	0.000	.2221683	.3129246
	frequency	.0000138	.0000179	0.77	0.441	-.0000213	.0000488
	intercept	-31.71535	2.883366	-11.00	0.000	-37.36664	-26.06405
dBPTS	LAeq8hr	.9423257	.0952078	9.90	0.000	.7557218	1.12893
	frequency	-.0000578	.0000222	-2.60	0.009	-.0001013	-.0000143
	intercept	-107.0627	12.08602	-8.86	0.000	-130.7509	-83.37456
dBTS0	LAeq8hr	2.037361	.1333166	15.28	0.000	1.776065	2.298657
	frequency	.0003054	.0000721	4.24	0.000	.0001641	.0004467
	intercept	-221.3562	16.91888	-13.08	0.000	-254.5166	-188.1958

Table 13. Results of models with *LAeq8hr* and *baseline threshold* as fixed effects. For outcomes that were binary (*ptscat15, ptscat15, ts0cat25, ts0cat15)* generalized linear mixed models were used and for continuous outcomes (*dBPTS* and *dBTS0*) linear mixed effects models were used.

Outcome variable	Effect	Coefficient	Std. Err.	z	p-value	95% Conf. Interval	
ptscat25	**LAeq8hr**	.3707636	.0282721	13.11	0.000	.3153514	.4261758
	base.	-.0888227	.006847	-12.97	0.000	-.1022425	-.0754029
	intercept	-51.024	3.781652	-13.49	0.000	-58.4359	-43.6121
ptscat15	**LAeq8hr**	.3408505	.0322112	10.58	0.000	.2777178	.4039832
	base.	-.0671	.0056658	-11.84	0.000	-.0782047	-.0559953
	intercept	-45.17202	4.197758	-10.76	0.000	-53.39948	-36.94457
ts0cat25	**LAeq8hr**	.3259738	.0280233	11.63	0.000	.2710492	.3808984
	base.	-.0949637	.0083795	-11.33	0.000	-.1113873	-.0785401
	intercept	-39.19139	3.514821	-11.15	0.000	-46.08031	-32.30247
ts0cat15	**LAeq8hr**	.2912397	.0255154	11.41	0.000	.2412304	.341249
	base.	-.0939851	.0081807	-11.49	0.000	-.1100189	-.0779513
	intercept	-33.40947	3.162125	-10.57	0.000	-39.60713	-27.21182
dBPTS	**LAeq8hr**	.9457887	.0925308	10.22	0.000	.7644316	1.127146
	base.	-.2760543	.0113238	-24.38	0.000	-.2982485	-.2538601
	intercept	-105.0793	11.74697	-8.95	0.000	-128.103	-82.05569
dBTS0	**LAeq8hr**	2.043131	.1326489	15.40	0.000	1.783144	2.303119
	base.	-.7443166	.0280935	-26.49	0.000	-.7993789	-.6892544
	intercept	-212.5219	16.83579	-12.62	0.000	245.5195	-179.5244

Table 14. Classification table for which Y=0 indicates that a characteristic is absent and Y=1 that a characteristic is present.

Classified as:	**Observed as:**	
	Y=0 (absent)	**Y=1 (present)**
Y=0 (absent)	a	b
Y=1 (present)	c	d

Table 15. Hypothetical classification table for which binary outcome (safe or hazardous) is classified on the basis of an exposure criterion and on the basis of whether or not hearing loss actually occurred.

Classified as:	**Observed as:**	
	safe	**hazardous**
safe	a	b
hazardous	c	d

Table 16. **Interpretation of values of AUC (Harrell's C) ranging from 0.5 to 1.0, as suggested by Hosmer & Lemeshow (2000, p. 162).**

AUC (Harrell's C)	**Level of discrimination**
AUC=0.5	No discrimination
$0.7 \leq$ AUC < 0.8	Acceptable discrimination
$0.8 \leq$ AUC < 0.9	Excellent discrimination
$0.9 \leq$ AUC	Outstanding discrimination

Table 17. **AUC for each of the six exposure criteria when *ptscat25* is the outcome variable.**

Exposure Criterion	**Harrell's C (AUC)**	**Jackknife Std. Error**	**95% Conf. Interval**
MIL-STD 1474D	.6047235	.0208645	(.5638297, .6456172)
Unwarned AHAAH	.8194065	.0123776	(.7951468, .8436662)
Warned AHAAH	.7990534	.0129193	(.773732, .8243748)
LAeq8hr	.7857181	.013629	(.7590058, .8124304)
Pfander	.7155696	.0151623	(.6858522, .7452871)
Smoorenburg	.7125368	.0149769	(.6831827, .7418909)

Table 18. **Difference in AUC's for different exposure criteria when *ptscat25* is outcome variable.**

Difference in Exposure Criteria	Difference in Harrell's C	Std. Err.	z	p-value (P>|z|)	95% Conf. Interval
MIL-STD 1474D – Unwarned AHAAH	-.2147	.0235	-9.13	0.000	(-.2607, -.1686)
MIL-STD 1474D – Warned AHAAH Error! Bookmark not defined.	-.1943	.0206	-9.45	0.000	(-.2346, -.1540)
MIL-STD 1474D – LAeq8hr	-.1810	.0158	-11.43	0.000	(-.2120, -.1500)
MIL-STD 1474D – Pfander	-.1108	.0141	-7.86	0.000	(-.1385, -.0832)
MIL-STD 1474D – Smoorenburg	-.1078	.0168	-6.42	0.000	(-.1407, -.0749)
Unwarned AHAAH – Warned AHAAH	.0204	.0058	3.51	0.000	(.0090, .0317)
Unwarned AHAAH – LAeq8hr	.0337	.0112	3.02	0.003	(.0118, .0556)
Unwarned AHAAH – Pfander	.1038	.0146	7.10	0.000	(.0752, .1325)
Unwarned AHAAH – Smoorenburg	.1069	.0146	7.30	0.000	(.0782, .1356)
Warned AHAAH – LAeq8hr	.0133	.0095	1.40	0.162	(-.0054, .0320)
Warned AHAAH – Pfander	.0835	.0126	6.61	0.000	(.0587, .1082)
Warned AHAAH – Smoorenburg	.0865	.0136	6.36	0.000	(.0599, .1132)
LAeq8hr – Pfander	.0701	.0063	11.18	0.000	(.0579, .0824)
LAeq8hr – Smoorenburg	.0732	.0081	8.98	0.000	(.0572, .0891)
Pfander – Smoorenburg	.0030	.0051	0.60	0.549	(-.0069, .0129)

Table 19. **Summary of paired comparisons of AUC for different exposure criteria for outcome variable *ptscat25*. The criterion judged better at discrimination for each pair is indicated by the letter (M=MIL-STD 1474D, U=Unwarned AHAAH, W=Warned AHAAH, L=LAeq8hr, P=Pfander, S=Smoorenburg, O=no significant difference).**

	MIL-STD 1474D	Unwarned AHAAH	Warned AHAAH	LAeq8hr	Pfander	Smoorenburg
MIL-STD 1474D		U	W	L	P	S
Unwarned AHAAH	U		U	U	U	U
Warned AHAAH	W	U		O	W	W
LAeq8hr	L	U	O		L	L
Pfander	P	U	W	L		O
Smoorenburg	S	U	W	L	O	

Table 20. **AUC for each of the six exposure criteria when *ptscat15* is the outcome variable.**

Exposure Criterion	**Harrell's C (AUC)**	**Jackknife Std. Error**	**95% Conf. Interval**
MIL-STD 1474D	.5872589	.018581	(.5508408, .6236769)
Unwarned AHAAH	.8024483	.0120821	(.7787679, .8261287)
Warned AHAAH	.7807624	.0126106	(.7560462, .8054786)
LAeq8hr	.7651092	.0132269	(.7391848, .7910335)
Pfander	.6944177	.0147122	(.6655823, .723253)
Smoorenburg	.6949487	.0146263	(.6662817, .7236156)

Table 21. **Difference in AUC's for different exposure criteria when *ptscat15* is outcome variable.**

Difference in Exposure Criteria	**Difference in Harrell's C**	**Std. Err.**	**z**	**p-value (P>|z|)**	**95% Conf. Interval**
MIL-STD 1474D – Unwarned AHAAH	-.2152	.0214	-10.05	0.000	(-.2572, -.1732)
MIL-STD 1474D – Warned AHAAH	-.1935	.0189	-10.24	0.000	(-.2305, -.1565)
MIL-STD 1474D – LAeq8hr	-.1779	.0141	-12.62	0.000	(-.2055, -.1502)
MIL-STD 1474D – Pfander	-.1072	.0127	-8.43	0.000	(-.1321, -.0823)
MIL-STD 1474D – Smoorenburg	-.1077	.0152	-7.10	0.000	(-.1374, -.0780)
Unwarned AHAAH – Warned AHAAH	.0217	.0056	3.84	0.000	(.0106, .0327)
Unwarned AHAAH – LAeq8hr	.0373	.0107	3.47	0.001	(.0163, .0584)
Unwarned AHAAH – Pfander	.1080	.0142	7.63	0.000	(.0803, .1358)
Unwarned AHAAH – Smoorenburg	.1075	.0142	7.55	0.000	(.0796, .1354)
Warned AHAAH – LAeq8hr	.0157	.0093	1.68	0.093	(-.0026, .0339)
Warned AHAAH – Pfander	.0863	.0123	6.99	0.000	(.0621, .1105)
Warned AHAAH – Smoorenburg	.0858	.0133	6.44	0.000	(.0597, .1119)
LAeq8hr – Pfander	.0707	.0060	11.74	0.000	(.0589, .0825)
LAeq8hr – Smoorenburg	.0702	.0078	8.98	0.000	(.0549, .0855)
Pfander – Smoorenburg	-.0005	.0048	-0.11	0.912	(-.0099, .0088)

Table 22. **Summary of paired comparisons of AUC for different exposure criteria for outcome variable *ptscat15*. The criterion judged better at discrimination for each pair is indicated by the letter (M=MIL-STD 1474D, U=Unwarned AHAAH, W=Warned AHAAH, L=LAeq8hr, P=Pfander, S=Smoorenburg, O=no significant difference).**

	MIL-STD 1474D	Unwarned AHAAH	Warned AHAAH	LAeq8hr	Pfander	Smoorenburg
MIL-STD 1474D		U	W	L	P	S
Unwarned AHAAH	U		U	U	U	U
Warned AHAAH	W	U		O	W	W
LAeq8hr	L	U	O		L	L
Pfander	P	U	W	L		O
Smoorenburg	S	U	W	L	O	

Table 23. AUC for each of the six exposure criteria when *ts0cat25* is the outcome variable.

Exposure Criterion	Harrell's C (AUC)	Jackknife Std. Error	95% Conf. Interval
MIL-STD 1474D	.6323472	.0177433	(.597571, .6671234)
Unwarned AHAAH	.7816304	.0142462	(.7537085, .8095524)
Warned AHAAH	.7814936	.0135726	(.7548918, .8080954)
LAeq8hr	.8025834	.0124967	(.7780903, .8270765)
Pfander	.7324229	.015368	(.7023022, .7625437)
Smoorenburg	.7328792	.0160538	(.7014143, .7643441)

Table 24. **Difference in AUC's for different exposure criteria when *ts0cat25* is outcome variable.**

Difference in Exposure Criteria	Difference in Harrell's C	Std. Err.	z	p-value (P>|z|)	95% Conf. Interval
MIL-STD 1474D – Unwarned AHAAH	-.1493	.0248	-6.02	0.000	(-.1979, -.1007)
MIL-STD 1474D – Warned AHAAH Error! Bookmark not defined.	-.1491	.0220	-6.78	0.000	(-.1922, -.1060)
MIL-STD 1474D – LAeq8hr	-.1702	.0146	-11.65	0.000	(-.1989, -.1416)
MIL-STD 1474D – Pfander	-.1001	.0127	-7.87	0.000	(-.1250, -.0752)
MIL-STD 1474D – Smoorenburg	-.1005	.0154	-6.51	0.000	(-.1308, -.0703)
Unwarned AHAAH – Warned AHAAH	.0001	.0066	0.02	0.984	(-.0128, .0131)
Unwarned AHAAH – LAeq8hr	-.0210	.0134	-1.56	0.119	(-.0473, .0054)
Unwarned AHAAH – Pfander	.0492	.0184	2.67	0.008	(.0131, .0853)
Unwarned AHAAH – Smoorenburg	.0488	.0186	2.62	0.009	(.0123, .0852)
Warned AHAAH – LAeq8hr	-.0211	.0120	-1.76	0.078	(-.0445, .0024)
Warned AHAAH – Pfander	.0491	.0168	2.91	0.004	(.0161, .0821)
Warned AHAAH – Smoorenburg	.0486	.0179	2.72	0.007	(.0136, .0837)
LAeq8hr – Pfander	.0702	.0073	9.57	0.000	(.0558, .0845)
LAeq8hr – Smoorenburg	.0697	.0092	7.55	0.000	(.0516, .0878)
Pfander – Smoorenburg	-.0005	.0051	-0.09	0.929	(-.0105, .0096)

Table 25. **Summary of paired comparisons of AUC for different exposure criteria for outcome variable *ts0cat25*. The criterion judged better at discrimination for each pair is indicated by the letter (M=MIL-STD 1474D, U=Unwarned AHAAH, W=Warned AHAAH, L=LAeq8hr, P=Pfander, S=Smoorenburg, O=no significant difference).**

	MIL-STD 1474D	**Unwarned AHAAH**	**Warned AHAAH**	**LAeq8hr**	**Pfander**	**Smoorenburg**
MIL-STD 1474D		U	W	L	P	S
Unwarned AHAAH	U		O	O	U	U
Warned AHAAH	W	O		O	W	W
LAeq8hr	L	O	O		L	L
Pfander	P	U	W	L		O
Smoorenburg	S	U	W	L	O	

Table 26. AUC for each of the six exposure criteria when *ts0cat15* is the outcome variable.

Exposure Criterion	**Harrell's C (AUC)**	**Jackknife Std. Error**	**95% Conf. Interval**
MIL-STD 1474D	.6296356	.0179684	(.5944181, .664853)
Unwarned AHAAH	.7765193	.0153349	(.7464634, .8065751)
Warned AHAAH	.7759702	.0140847	(.7483648, .8035757)
LAeq8hr	.8039006	.0119457	(.7804875, .8273138)
Pfander	.7331744	.0152786	(.7032288, .7631199)
Smoorenburg	.7339939	.0162066	(.7022296, .7657583)

Table 27. **Difference in AUC's for different exposure criteria when *ts0cat15* is outcome variable.**

Difference in Exposure Criteria	**Difference in Harrell's C**	**Std. Err.**	**z**	**p-value (P>|z|)**	**95% Conf. Interval**
MIL-STD 1474D – Unwarned AHAAH	-.1469	.0273	-5.38	0.000	(-.2004, -.0933)
MIL-STD 1474D – Warned AHAAH	-.1463	.0244	-5.99	0.000	(-.1942, -.0985)
MIL-STD 1474D – LAeq8hr	-.1743	.0161	-10.85	0.000	(-.2057, -.1428)
MIL-STD 1474D – Pfander	-.1035	.0138	-7.49	0.000	(-.1306, -.0764)
MIL-STD 1474D – Smoorenburg	-.1044	.0169	-6.18	0.000	(-.1374, -.0713)
Unwarned AHAAH – Warned AHAAH	.0005	.0069	0.08	0.937	(-.0131, .0142)
Unwarned AHAAH – LAeq8hr	-.0274	.0147	-1.86	0.063	(-.0562, .0015)
Unwarned AHAAH – Pfander	.0433	.0202	2.15	0.032	(.0038, .0829)
Unwarned AHAAH – Smoorenburg	.0425	.0202	2.10	0.035	(.0029, .0821)
Warned AHAAH – LAeq8hr	-.0279	.0132	-2.12	0.034	(-.0537, -.0021)
Warned AHAAH – Pfander	.0428	.0184	2.33	0.020	(.0068, .0788)
Warned AHAAH – Smoorenburg	.0420	.0193	2.18	0.030	(.0042, .0798)
LAeq8hr – Pfander	.0707	.0079	8.96	0.000	(.0553, .0862)
LAeq8hr – Smoorenburg	.0699	.0097	7.23	0.000	(.0510, .0889)
Pfander – Smoorenburg	-.0008	.0055	-0.15	0.881	(-.0115, .0099)

Table 28. Summary of paired comparisons of AUC for different exposure criteria for outcome variable *ts0cat15*. The criterion judged better at discrimination for each pair is indicated by the letter (M=MIL-STD 1474D, U=Unwarned AHAAH, W=Warned AHAAH, L=LAeq8hr, P=Pfander, S=Smoorenburg, O=no significant difference).

	MIL-STD 1474D	Unwarned AHAAH	Warned AHAAH	LAeq8hr	Pfander	Smoorenburg
MIL-STD 1474D		U	W	L	P	S
Unwarned AHAAH	U		O	O	U	U
Warned AHAAH	W	O		L	W	W
LAeq8hr	L	O	L		L	L
Pfander	P	U	W	L		O
Smoorenburg	S	U	W	L	O	

VIII. Figures

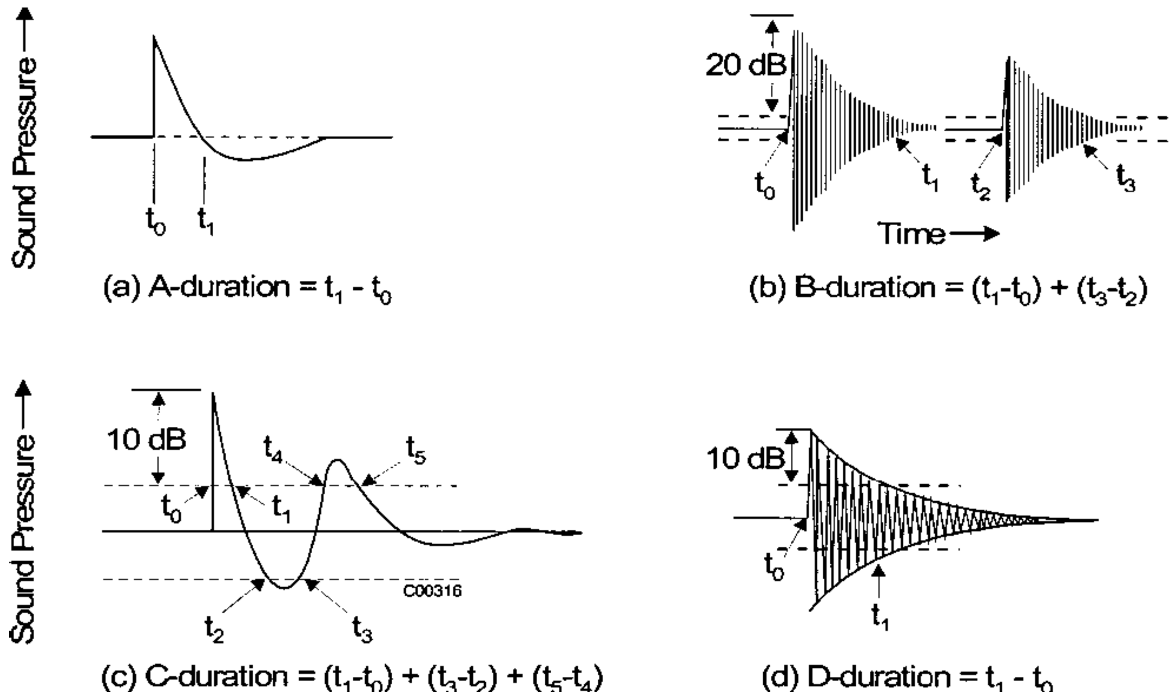

Figure 1. Definition of impulse noise duration (Smoorenburg, 1992)

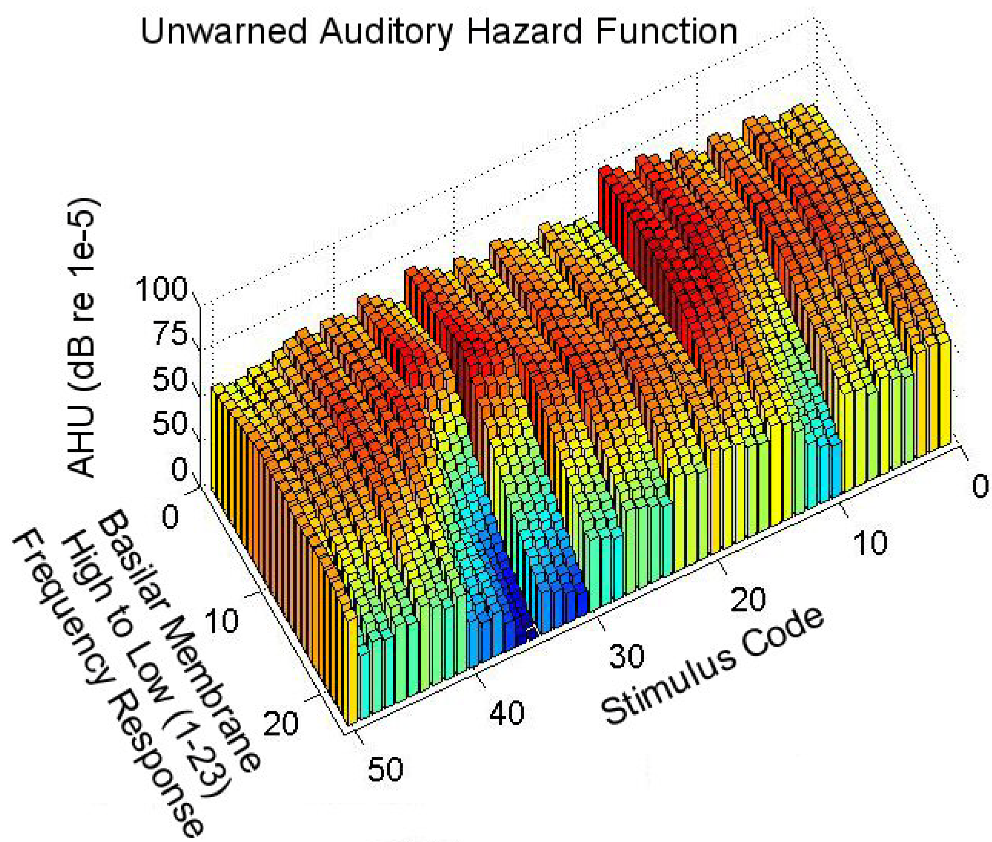

Figure 2. The Basilar membrane response from the AHAAH model for the unwarned condition. The basilar membrane sections are numbered 1 to 23 and correspond to High to Low frequencies. The stimulus codes are given. For stimuli 18 to 40, a series of increasingly high frequency narrow band noises produce greater hazard at the higher frequency segments.

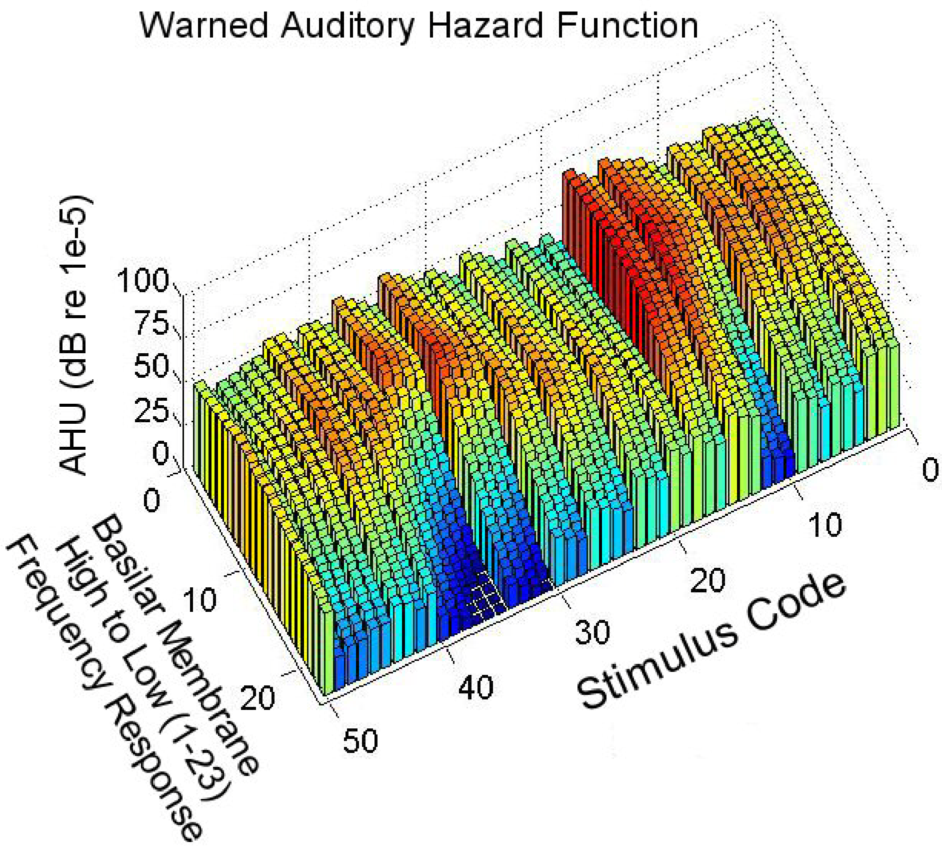

Figure 3. The Basilar membrane response from the AHAAH model for the warned condition. The basilar membrane sections are numbered 1 to 23 and correspond to High to Low frequencies. The stimulus codes are given. For stimuli 18 to 40, a series of increasingly high frequency narrow band noises produce greater hazard at the higher frequency segments. For the segments that are blue, the model response was at its minimum. Generally the warned model has a lower overall response than the unwarned response.

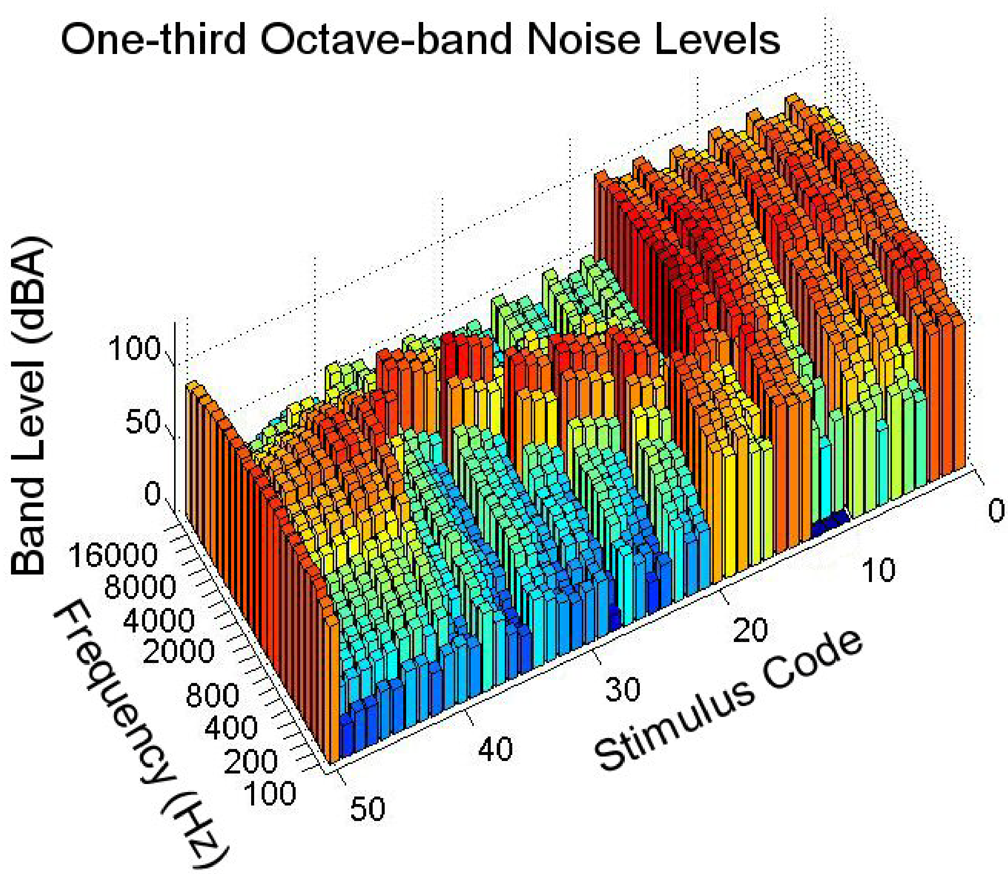

Figure 4. The one-third octave band analysis of the 50 stimulus waveforms in A-weighted Sound Exposure Level (dB SELA). For stimuli 18 to 40, a series of increasingly high frequency narrow band noises exhibit increasing energy at higher frequencies. At stimuli 10-12, the lowest level of the analysis yielded 0 dB in the 100 Hz band. The values for these waveforms were taken directly from the Hamernik et al. DTIC report and the Microsoft ACCESS database.

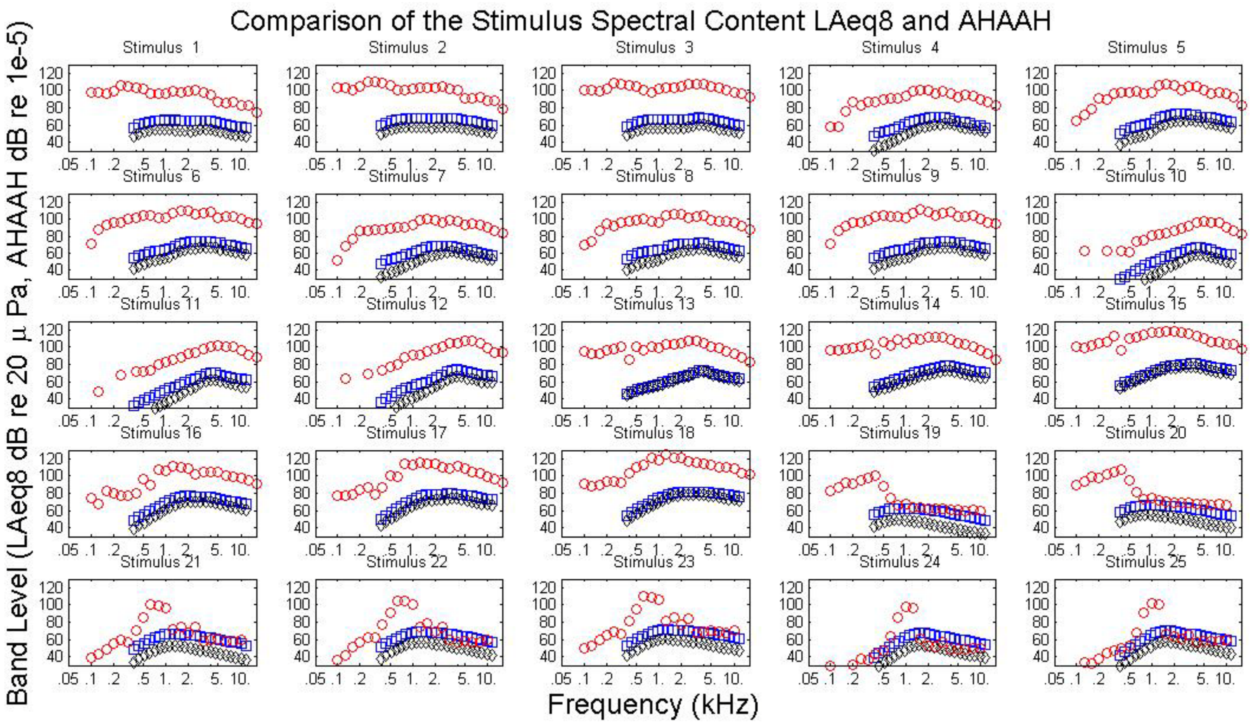

Figure 5. Comparison of A-weighted impulse spectra (red circles), Unwarned AHAAH spectra (blue squares) and Warned AHAAH spectra (black diamonds). The stimuli 1 through 25 are shown in this figure and can be compared relative to one another in Figure 2, Figure 3, and Figure 4.

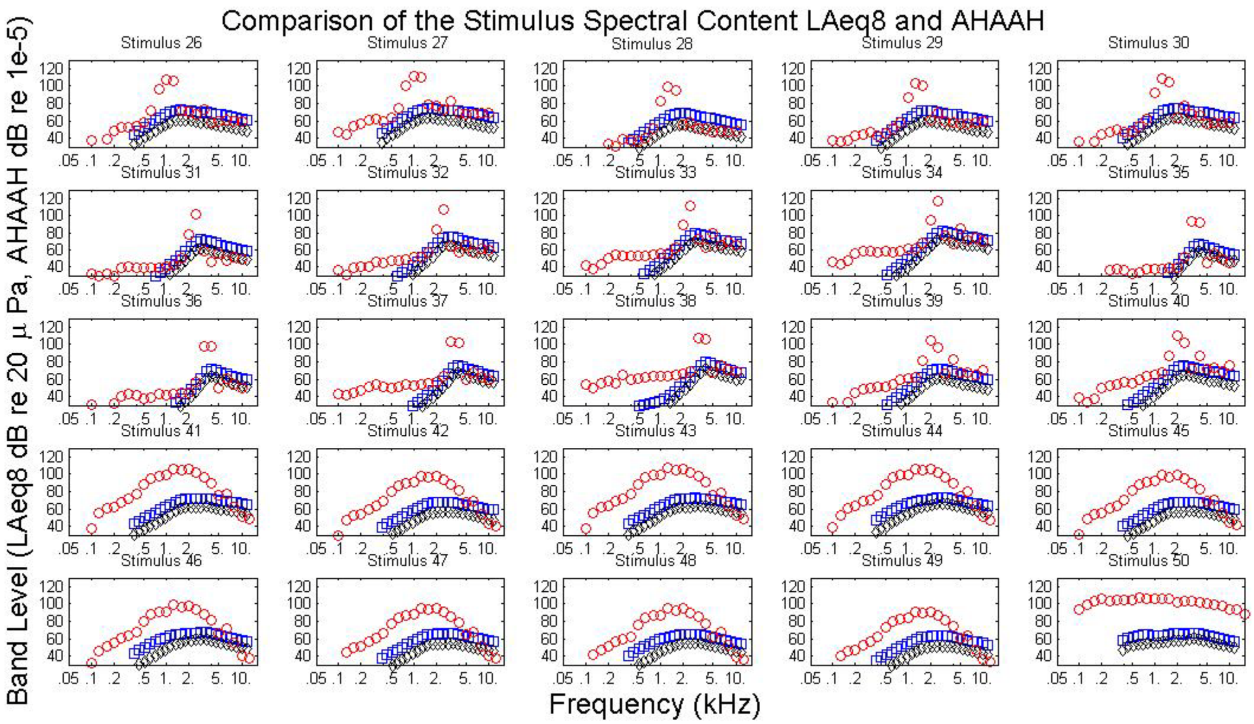

Figure 6. Comparison of A-weighted impulse spectra (red circles), Unwarned AHAAH spectra (blue squares) and Warned AHAAH spectra (black diamonds). The stimuli 26 through 50 are shown in this figure and can be compared relative to one another in Figure 2, Figure 3, and Figure 4.

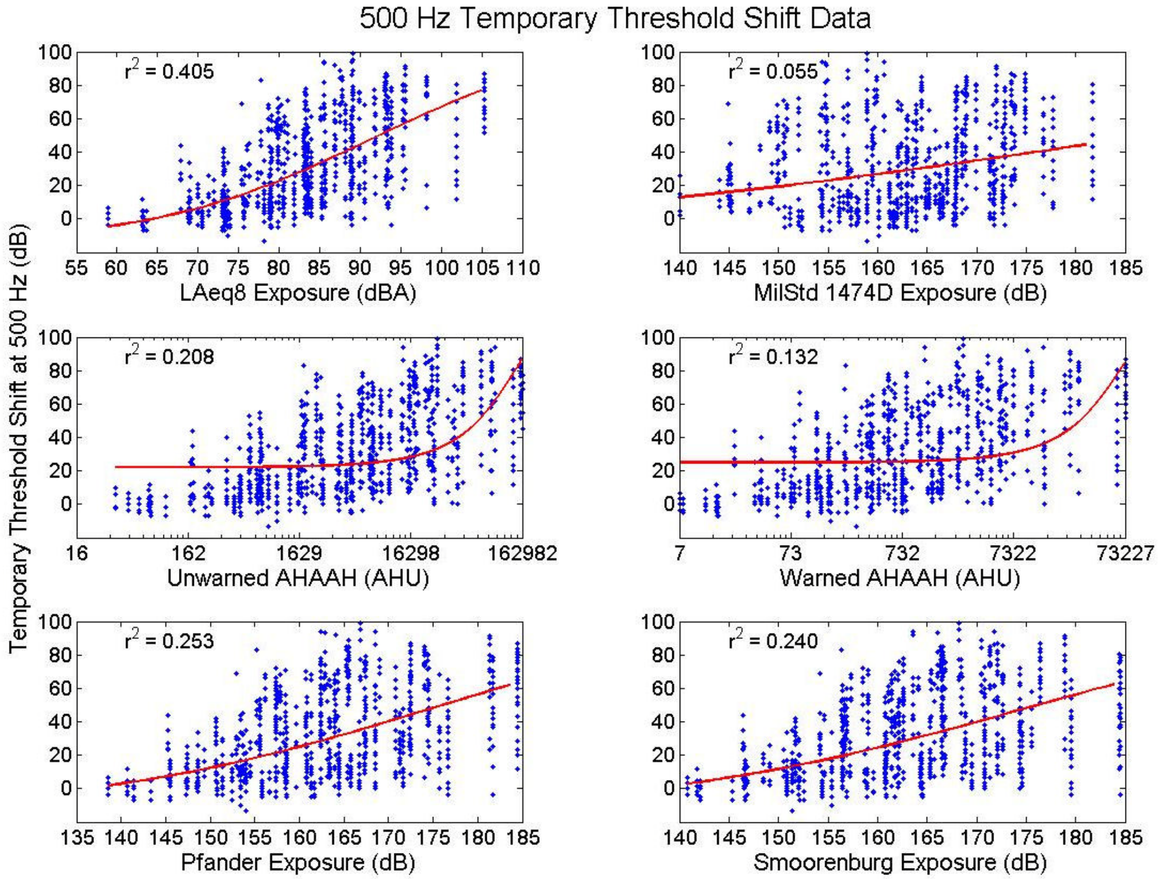

Figure 7. The 500 Hz Temporary Threshold Shift data 1 hour after exposure plotted against the six hazard indices. LAeq8hr exhibits the best organization of the data as indicated by the r^2 coefficient of determination. MilStd 1474D exhibits the poorest organization. The Unwarned and Warned AHAAH models exhibit a nonzero fit of the curve at the left side of the plot as a result of the linear character of the AHU statistic.

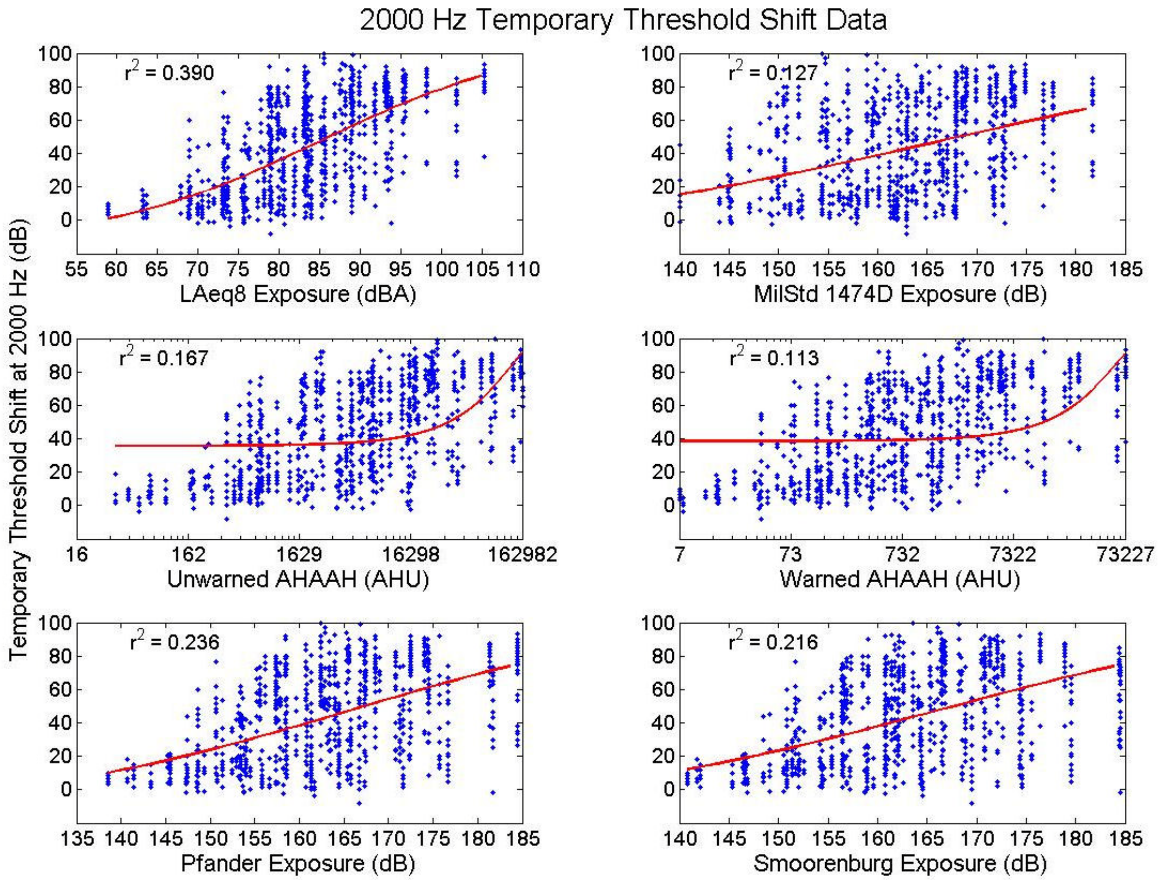

Figure 8. The 2000 Hz Temporary Threshold Shift data 1 hour after exposure plotted against the six hazard indices. LAeq8hr exhibits the best organization of the data as indicated by the r^2 coefficient of determination. The Warned AHAAH model exhibits the poorest organization. The Unwarned and Warned AHAAH models exhibit a nonzero fit of the curve at the left side of the plot as a result of the linear character of the AHU statistic.

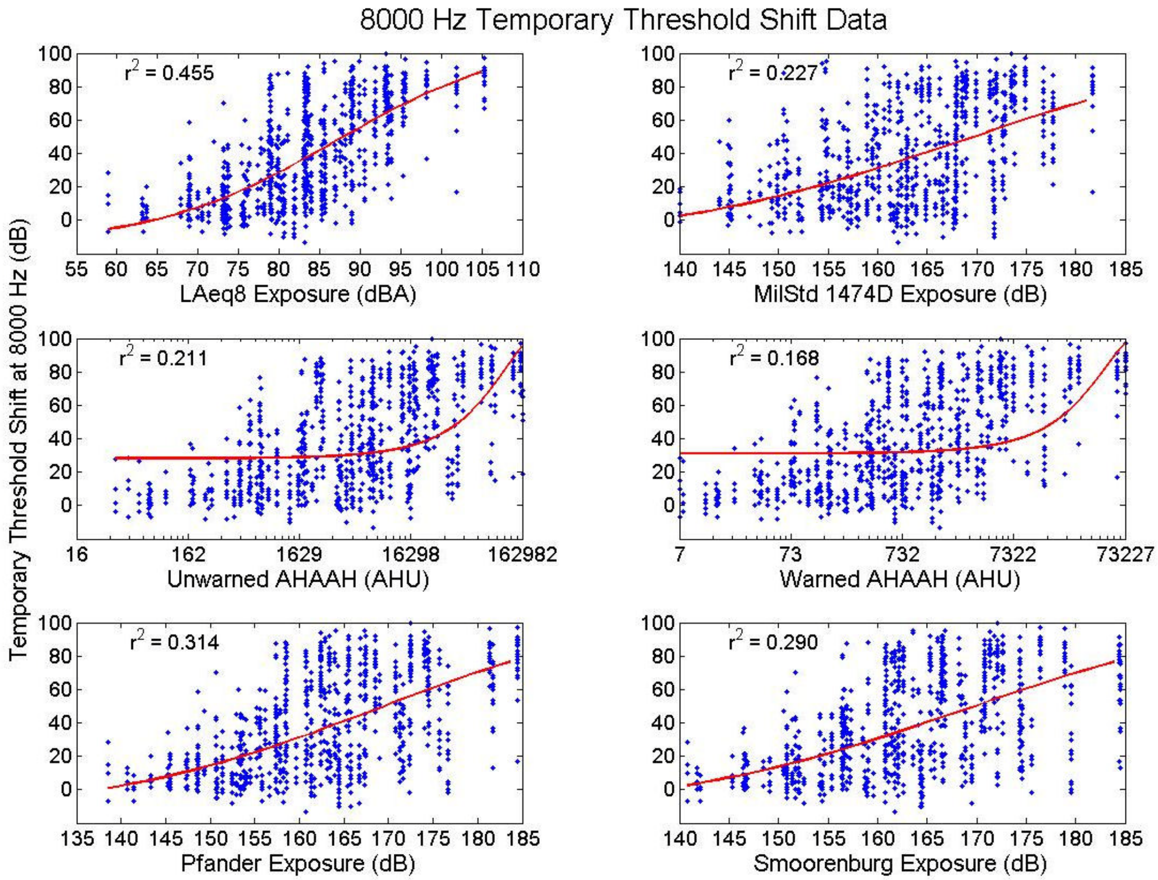

Figure 9. The 8000 Hz Temporary Threshold Shift data 1 hour after exposure plotted against the six hazard indices. LAeq8hr exhibits the best organization of the data as indicated by the r^2 coefficient of determination. The Warned AHAAH exhibits the poorest organization. The Unwarned and Warned AHAAH models exhibit a nonzero fit of the curve at the left side of the plot as a result of the linear character of the AHU statistic.

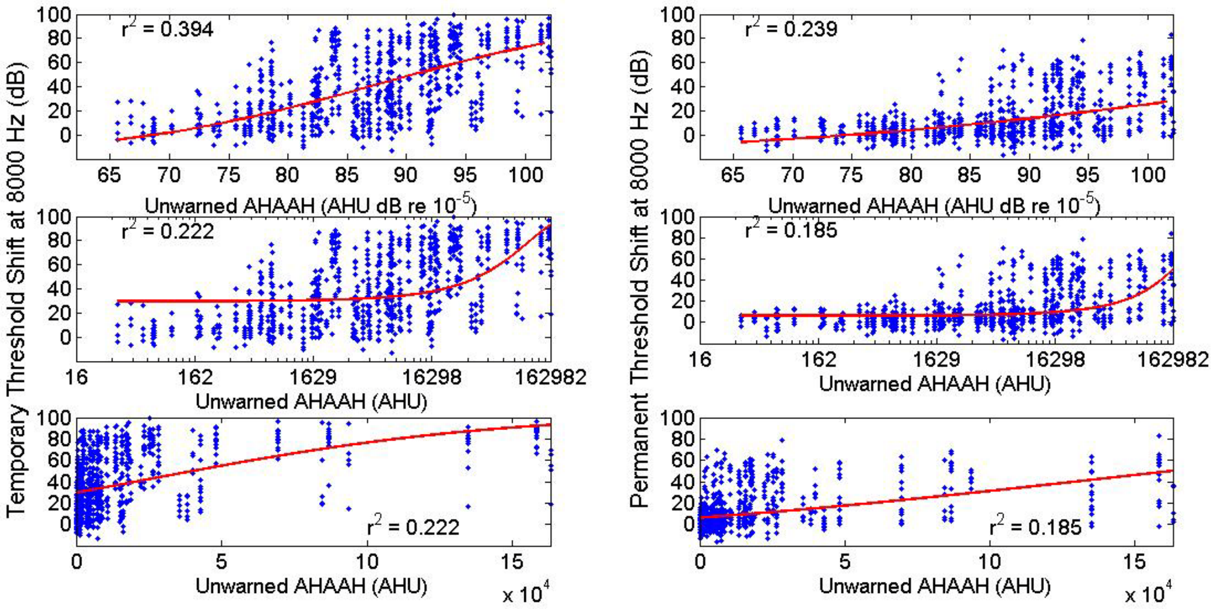

Figure 10. Comparison of the nonlinear fits for logarithmic and linear Unwarned AHAAH model data for the temporary (TS0) and permanent (PTS) threshold shift data at 8000 Hz. The left panels exhibit the TS0 data plotted against the AHU in dB (re 10^{-5} AHU) at the top panel, linear AHU on a log abscissa in the middle panel and linear AHU on a linear abscissa in the lower panel. The right panels exhibit the PTS data plotted against the AHU in dB (re 10^{-5} AHU) at the top panel, linear AHU on a log abscissa in the middle panel and linear AHU on a linear abscissa in the lower panel. The coefficient of determination for the lower two rows will be identical for the TS0 and PTS plots, since only the plot axis was changed.

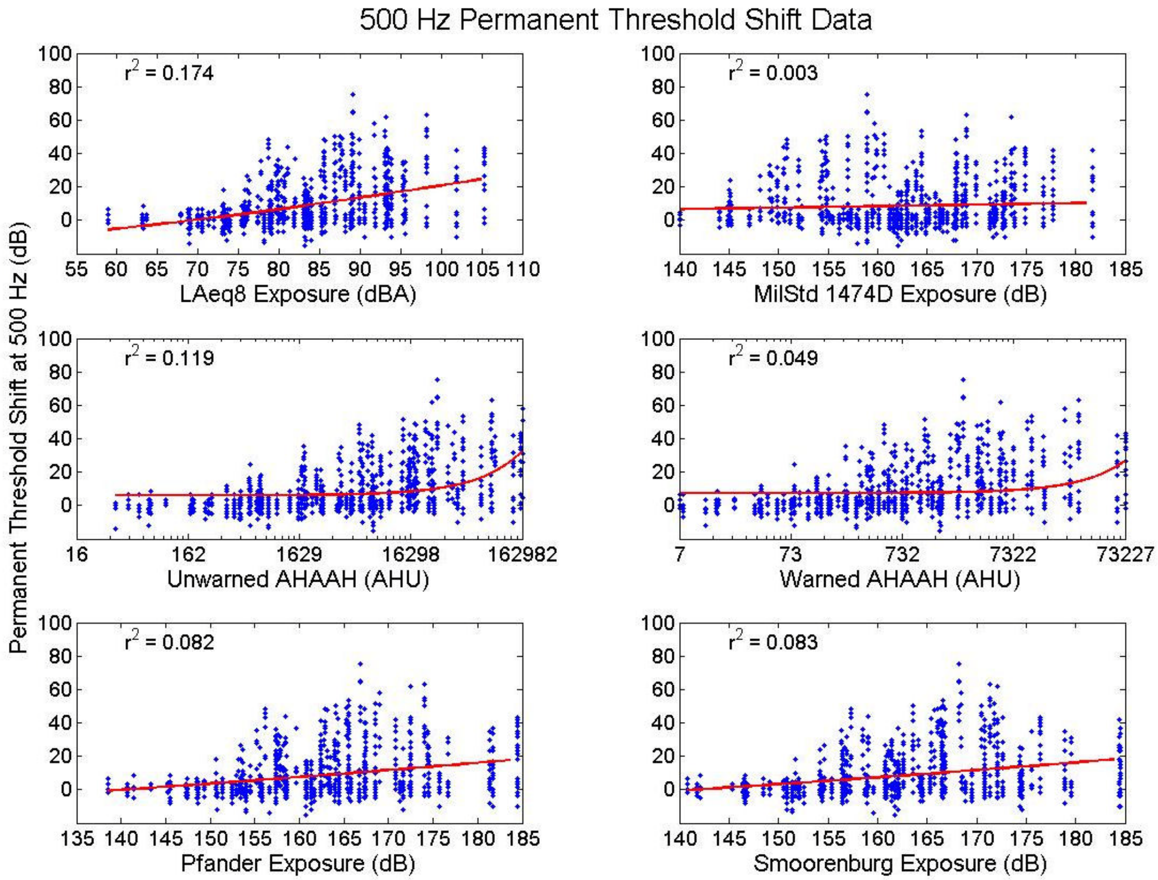

Figure 11. The 500 Hz Permanent Threshold Shift data approximately 4 weeks after exposure plotted against the six hazard indices. LAeq8hr exhibits the best organization of the data as indicated by the r^2 coefficient of determination. MilStd 1474D exhibits the poorest organization.

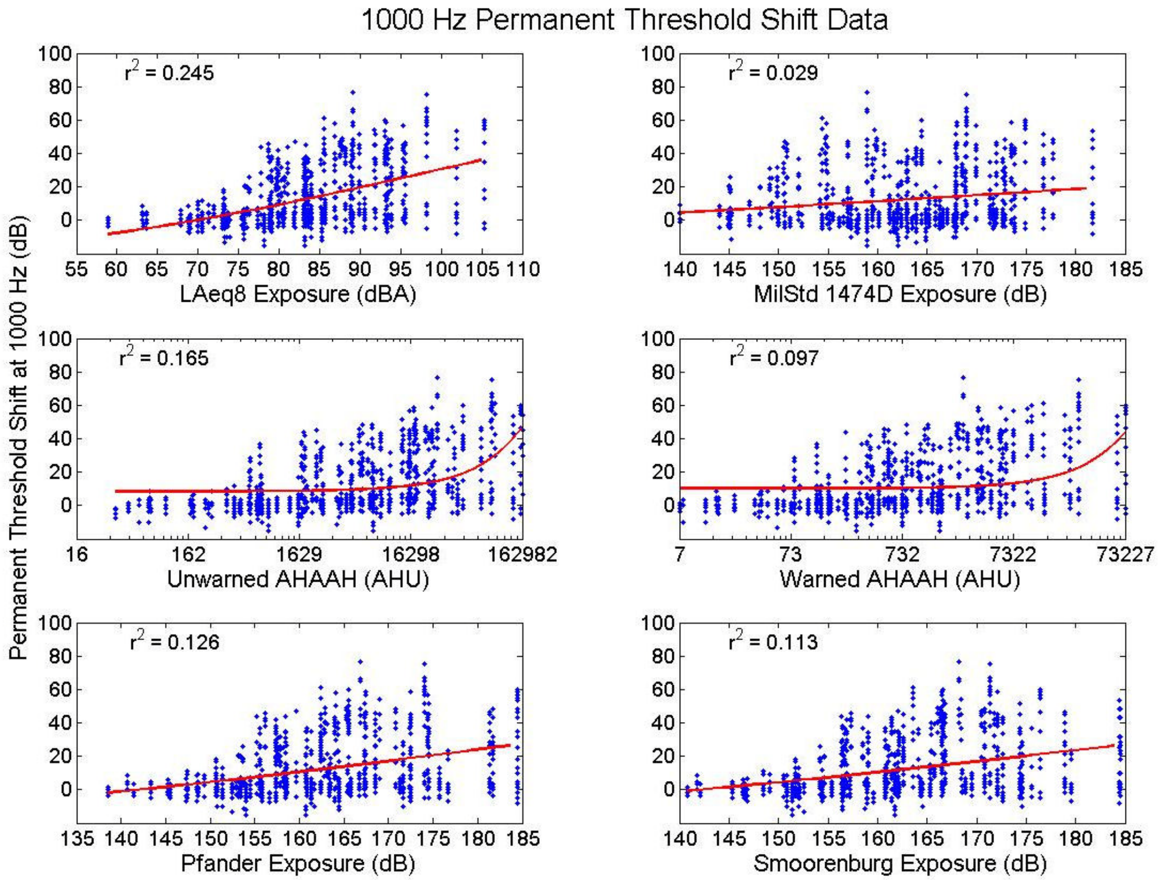

Figure 12. The 1000 Hz Permanent Threshold Shift data approximately 4 weeks after exposure plotted against the six hazard indices. LAeq8hr exhibits the best organization of the data as indicated by the r^2 coefficient of determination. MilStd 1474D exhibits the poorest organization.

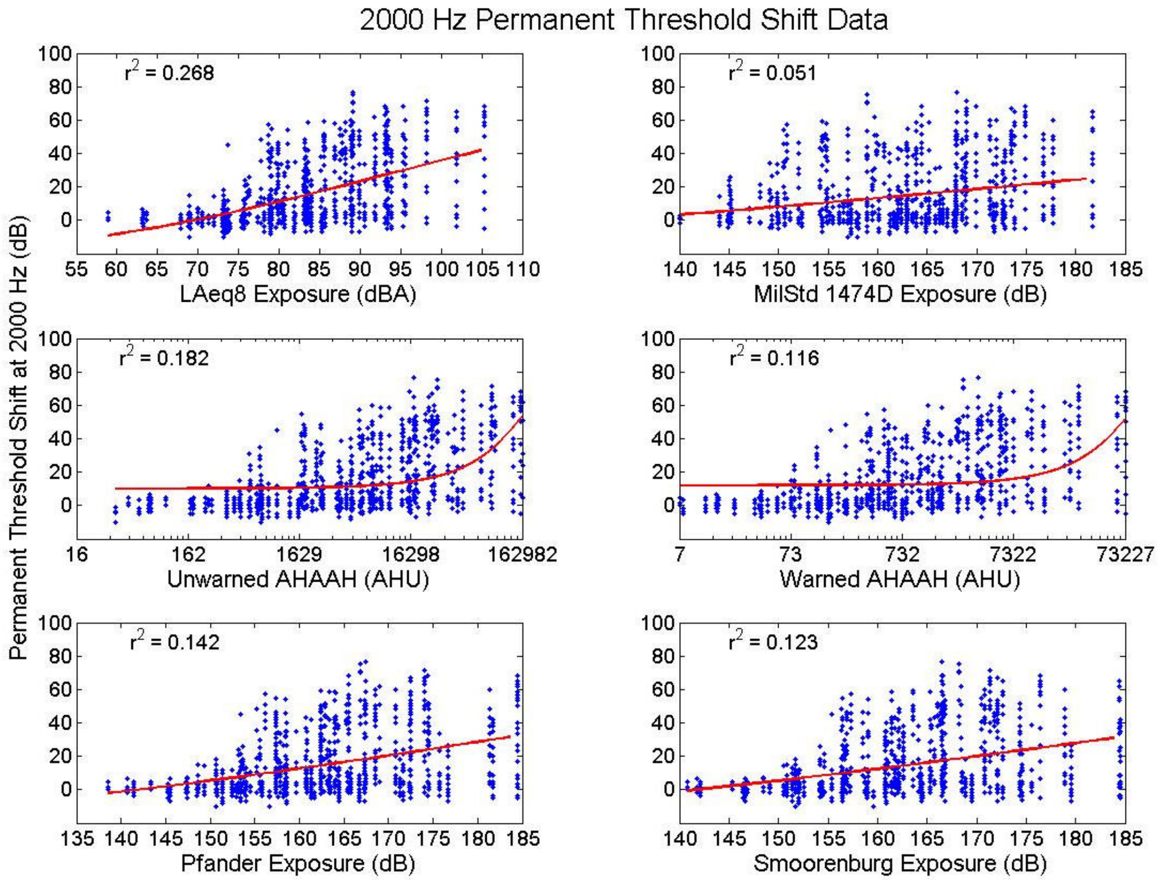

Figure 13. The 2000 Hz Permanent Threshold Shift data approximately 4 weeks after exposure plotted against the six hazard indices. LAeq8hr exhibits the best organization of the data as indicated by the r^2 coefficient of determination. MilStd 1474D exhibits the poorest organization.

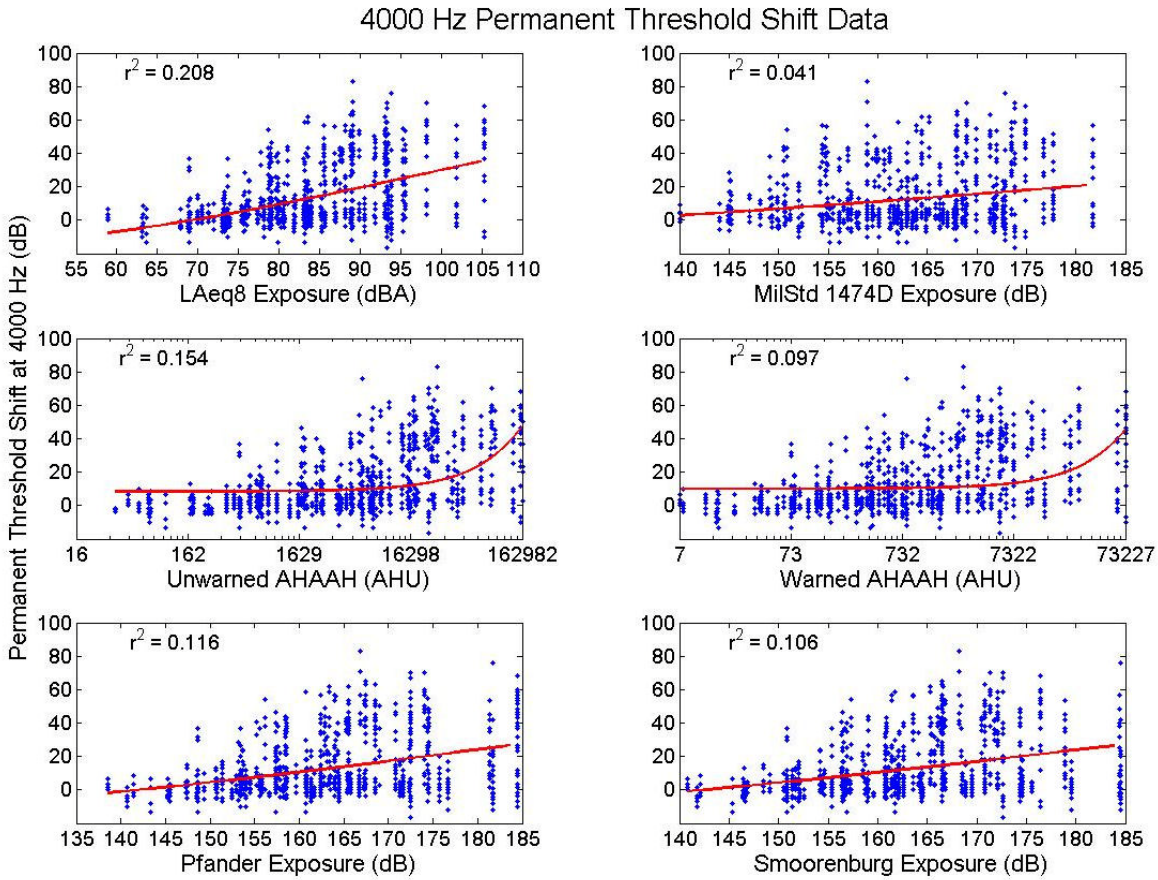

Figure 14. The 4000 Hz Permanent Threshold Shift data approximately 4 weeks after exposure plotted against the six hazard indices. LAeq8hr exhibits the best organization of the data as indicated by the r^2 coefficient of determination. MilStd 1474D exhibits the poorest organization.

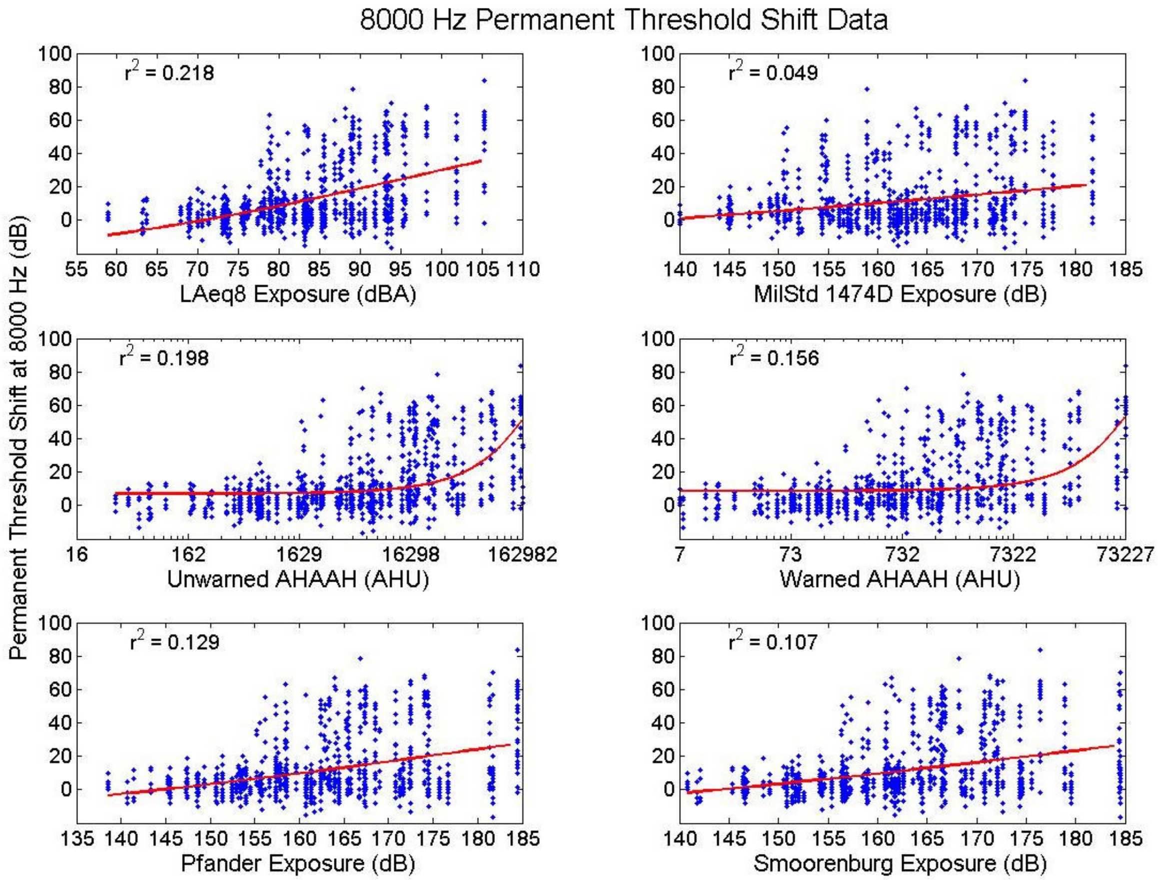

Figure 15. The 8000 Hz Permanent Threshold Shift data approximately 4 weeks after exposure plotted against the six hazard indices. LAeq8hr exhibits the best organization of the data as indicated by the r^2 coefficient of determination. MilStd 1474D exhibits the poorest organization.

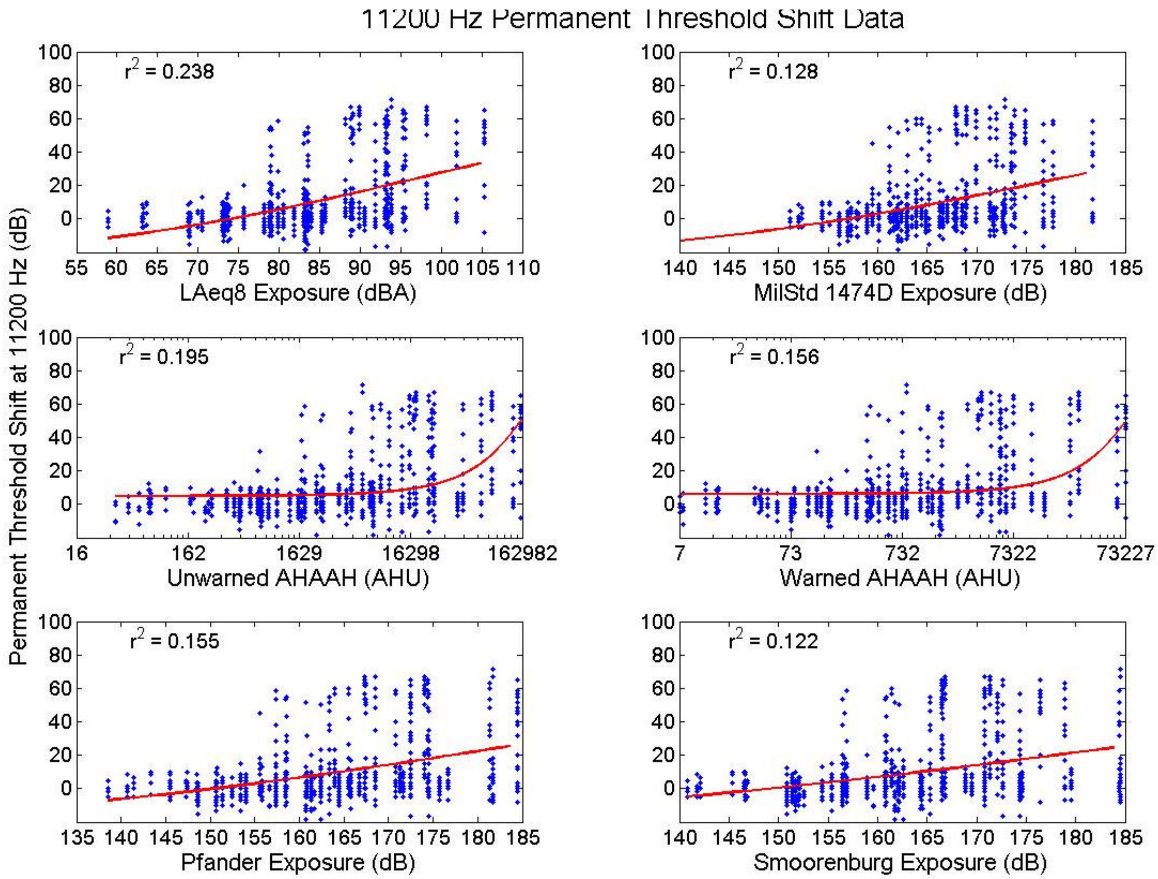

Figure 16. The 11200 Hz Permanent Threshold Shift data approximately 4 weeks after exposure plotted against the six hazard indices. LAeq8hr exhibits the best organization of the data as indicated by the r^2 coefficient of determination. Smoorenburg criteria exhibits the poorest organization.

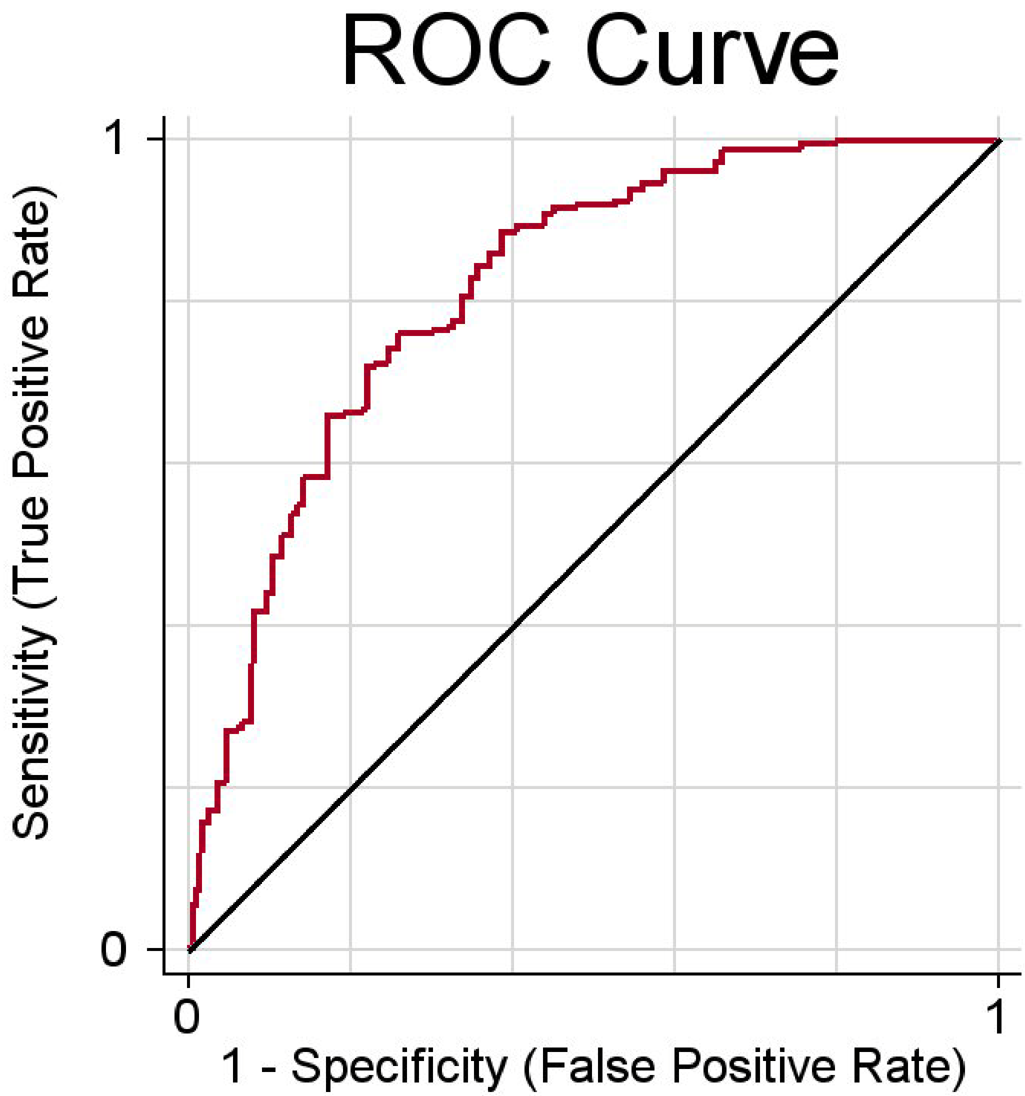

Figure 17. Example Receiver Operating Characteristic (ROC) curve that illustrates the area under the curve (AUC) for increased specificity and sensitivity. Larger AUC indicates the classification scheme does a better job in a discrimination task.