This Survey Report and any recommendations made herein are for the specific facility evaluated and may not be universally applicable. Any recommendations made are not to be considered as final statements of NIOSH policy or of any agency or individual involved. Additional NIOSH Survey Reports are available at http://www.cdc.gov/niosh/surveyreports.

An Analysis of the Blast Overpressure Study Data Comparing Three Exposure Criteria

REPORT WRITTEN BY:

William J. Murphy, Ph.D.

Amir Khan

Peter B. Shaw, Ph.D.

REPORT DATE:

December 3, 2009

REPORT NUMBER:

EPHB 309-05h

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.

An Analysis of the Blast Overpressure Study Data Comparing Three Exposure Criteria

Table of Contents

Executive Summary

The Blast Overpressure (BOP) data from the U.S. Army Albuquerque studies were analyzed with population-average (generalized estimating equations – GEE) and subject-specific (generalized linear mixed models – GLMM) logit models using five different noise exposure criteria as independent variables: MIL-STD-1474D for the free-field condition, the A-weighted equivalent 8-hour level (L_{Aeq8hr}) for the free-field and protected conditions, Auditory Hazard Assessment Algorithm for the Humans (AHAAH) warned and unwarned status of the middle ear reflex for the protected condition. The purpose of the current analysis was to determine which noise exposure criterion best characterized the BOP injury data. The noise exposure criteria that yielded the best fit as judged by the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for the GLMM and the quasi-likelihood under the independence model criterion (QIC) for the GEE models was the free-field L_{Aeq8hr}, followed by the MIL-STD-1474D, protected L_{Aeq8hr}, unwarned AHAAH, and warned AHAAH. Generally, the free-field data for L_{Aeq8hr} provided a better fit to the BOP injury data compared to the protected data.

I. Introduction

From 1989 through 1994 at the Kirtland Air Force Base Test Site, in Albuquerque, New Mexico, the U.S. Army Medical Research and Materiel Command conducted a series of studies to determine the safe limits of exposure in which volunteer soldier participants were exposed to high intensity free-field impulse noises such as that produced by large caliber military weapons (Johnson, 1994; Johnson, 1997; Patterson and Johnson, 1994). Discharge of large caliber weapons produces a compressive concussion wave also referred to as blast overpressure (BOP) which propagates away from the source of the detonation. Military personnel frequently experience blast overpressure during peacetime training and combat operations for specialties that include but are not limited to infantry, artillery, armored cavalry and special operations.

A. Albuquerque Exposure Design

The amplitude, spectral content and number of impulses were varied systematically during Albuquerque BOP exposures. Seven different weights of explosive charge were used to create seven different peak sound pressure levels in the exposures. The number of impulses was selected from 6, 12, 25, 50 and 100 shots at each peak impulse level. The distance from the charge to the soldier participants (1, 3 or 5 meters) varied the spectral content. Waveforms recorded from the free-field pressure gauges and microphones under hearing protectors are shown in Figure 1. At the 1-meter distance, the duration of the first positive pressure portion of the waveform (A-duration) was about 0.9 ms. At the 3 and 5-meter distances, the A-durations were about 1.6 and 3.0 ms, respectively. As the A-duration increases, the low-frequency spectral content increases.

Several different types of hearing protector devices were evaluated in the Albuquerque BOP studies: RACAL talk-through ear muffs, RACAL modified talk-through ear muffs, E•A•R® Classic™ foam plugs and perforated plugs (e.g., Combat Arms earplug prototype and Fort Rucker plug). However in this analysis, only the free-field and under-the-muff data for the RACAL modified earmuffs were evaluated.

The exposures of this study were designed to prevent participants from incurring any permanent thresholds shift (PTS) during testing. Since temporary threshold shifts (TTS) in excess of 40 dB were likely to produce a PTS, the following post-exposure criteria were used by the original investigators: Audiometric failures occurred for a TTS of 25 dB or greater; Conditional failures occurred for TTS of between 15 and 25 dB (Kryter, 1966; Johnson 1998). Participants were tested twice each day prior to being exposed to the blast overpressures and were tested at +2 minutes, 20 minutes, 1 hour and at 2 hour intervals as necessary to identify the resolution of TTS. The thresholds at 2 minutes were used to determine whether the participant incurred an audiometric or conditional failure.

Following an audiometric failure, the participant was not allowed to progress to more energetic peak exposure levels. The next exposure was decreased by two levels and the number of impulses increased. Conceivably, participants could be exposed at a less energetic peak level with increasing numbers of impulses. However, the participants who exhibited audiometric failure tended to traverse a path of decreasing peak level with increasing numbers of shots or they withdrew from the study all together. Participants were free to withdraw from the study at any time. Following a conditional failure (TTS between 15 and 25 dB), the participants were exposed at the next lower peak level and the number of impulses was increased. If conditionally failed participants passed the lower level with increased impulses they were permitted to be exposed at progressively more energetic levels and greater numbers of impulses.

B. Previous Analysis of BOP Results

Several previous analyses of the BOP data have been performed. Patterson and Johnson (1994) evaluated the failure rates and identified a critical free-field level above which the exposures produced six or more audiometric failures for the RACAL modified and unmodified earmuffs. Patterson and Johnson reported critical levels of about 184 to 187 dB peak SPL for exposures of 100 impulses. Their analysis did not consider all exposures in a single regression model; rather they separated the three exposure distances and considered the performance of the group of participants assuming that those who failed were counted as failures at a higher energy level and/or higher number of impulses.

Chan et al. (2001) evaluated several candidate noise exposure criteria by applying logistic regression to the observed failure rates from the BOP data. The candidate noise exposure criteria were MIL-STD-1474D, free-field[1] A-weighted 8-hour Equivalent Level (L_{Aeq8hr}), C-duration Damage-risk Criteria (DRC) (Pfander et al. 1980), D-duration DRC (Smoorenburg, 1982) and a best-fit model. The analysis of Chan et al. (2001) took into account the fact that repeated measures were taken on each participant over different levels and numbers of impulses. Their analysis estimated the 95% confidence interval for the 5% failure rate, $L_{(95,95)}$, for each metric and compared them to the recommended exposure limits. Chan found that the BOP data yielded an $L_{(95,95)}$ of 185.3 dB for the MIL-STD-1474D evaluation of the free-field data and 112.7 dB for the L_{Aeq8hr} evaluation. These values are above the 177-dB design limit for MIL-STD-1474D and the 85 dB level for unprotected ears. Chan assumed that the RACAL earmuff used in the study provided about 15 dB additional protection which increases the 85-dB limit to 100 dB. Chan et al. further provided a measure of the quality of the fit of the metric to facilitate comparison of the various metrics. Chan developed a modification of the MIL-STD-1474D by varying the peak limit level and the coefficients modifying the B-duration and accumulation terms to create a best-fit model for the exposure results.

Patterson and Ahroon (2004) undertook to evaluate the Auditory Hazard Assessment Algorithm for Human (AHAAH) network transmission line model of the auditory periphery developed by the U.S. Army Research Laboratory (Price and Kalb, 1991; Price 2007a; Price 2007b). The AHAAH model estimates a single exposure metric based upon the recording location (free-field, at the entrance of the ear canal or at the tympanic membrane) and upon the status of the middle ear muscles (not activated/unwarned or activated/warned). During warned conditions, the model underestimated the hazards and during unwarned conditions the model overestimated the hazards. Patterson and Ahroon reported a lack of a systematic relationship between the calculated Auditory Risk Units (ARU) and the confidence level that a TTS would occur in the most sensitive percentage of the exposed participants.

---

[1] The term free-field in acoustics typically means that the sound field is created in the absence of reflections or an anechoic environment. For the purpose of this paper, free-field is used to mean the acoustic environment that is not measured in the ear. Reflections were present due to the multiple participants and equipment in the exposure area. The waveforms were time-windowed to remove reflections that would have arrived at the microphone at a later time.

Price (2007b) applied the AHAAH noise exposure criteria to the BOP data and audiometric failure rates. The waveforms were analyzed using the AHAAH model to determine the auditory hazard of each exposure cell. Using the threshold of 500 ARUs, Price performed an analysis of the predictiveness of the different noise exposure criteria for their ability to identify safe and hazardous exposures. Price made the assumption that an exposure was observed to be hazardous when the confidence level fell below 5% (six audiometric failures). The relative performance of the exposure criteria was assessed by calculating the percentage of correct predicted /observed classifications. According to Price, the AHAAH model correctly predicted the hazards in 95% of cases, while the MIL-STD-1474D correctly predicted in 42% of the cases and A-weighted energy was correct in predicting hazards in 25% of the cases. Classification performed at a single damage threshold yields only a partial picture of the predictiveness of an exposure criterion as will be shown later.

The background for the different interpretations is crucial to the conclusions that will be reported in this paper. This report will evaluate the exposures for both free-field and protected waveforms measured during the BOP exposures. The primary purpose of this study was to evaluate five competing noise exposure criteria with respect to the goodness-of-fit and the ability to discriminate exposure outcomes. The noise exposure criteria models include: free-field conditions for the MIL-STD-1474D, and L_{Aeq8hr} and the protected conditions for the L_{Aeq8hr} and the AHAAH model for the warned and unwarned middle ear status.

II. Method

For the analysis of the blast overpressure studies, the U.S. Army Medical Research and Materiel Command provided the U.S. National Institute for Occupational Safety and Health (NIOSH) with copies of the archived contractor reports and the U.S. Army Research Lab, Aberdeen Proving Ground provided NIOSH with the digital waveforms. The exposure facility permitted six participants to be exposed at a time. During any particular impulse event, the experimenters collected measurements under-the-muff for three exposed ears, measurements under-the-muff for three non-exposed head-shadowed ears and two free-field measurements. The recordings were only a representative sampling of the entire series of blast exposures. Additionally, some data were missing from the overall data provided to NIOSH. For the 1-meter Level 7 exposure, the protected waveforms were missing. For the 3-meter Level 1, both the protected and free-field waveforms were missing. For the remaining exposure cells, the average and median number of waveforms was 15; the minimum and maximum numbers of waveforms were 3 and 32 recordings. In some cases, protected waveform data were not available from the subject population; however, the investigators had participated in the exposures to determine the effect of angle of incidence using the same weight explosive charges. The waveforms collected at the same levels and angle of incidence as the actual exposures were only used to fill in missing entries.

The progression of every participant through the exposure matrix was diagrammed by the original investigators (Johnson, 1994). Temporary threshold shifts (TTS) were used to determine whether that person progressed to more energetic exposures or not. The TTS at each frequency was calculated by subtracting the average baseline audiogram from the post exposure audiograms for each participant. When the threshold shift at any audiometric frequency (125, 250, 500, 1000, 2000, 3000, 4000, 6000, or 8000 Hz) was between 15 and 25 dB, the soldier was considered to be a conditional failure at the level producing the threshold shift. The subsequent exposure was reduced one level and the number of impulses was incremented. When the TTS exceeded 25 dB at any audiometric frequency, the subsequent exposure was reduced by two levels and the numbers of impulses were incremented. Participants with an audiometric failure were prohibited from being exposed at the same or higher peak impulse level. They were allowed to receive exposure at lower peak impulse levels.

The failure rate for each exposure cell in the test matrix was computed using the actual (not presumed) observed audiometric (or audiometric and conditional) failures divided by the total number of participants within each cell of the exposure matrix. In cases where the principal investigator decided to remove a participant from an exposure or a participant withdrew from the study, that person was counted as an incomplete exposure and the data were not used toward estimating failure. If a participant exhibited a conditional failure and later passed at the same impulse peak level and number of impulses, the conditional failure was cancelled and then counted as a pass. The observed injury data for the blast overpressure studies are summarized in Tables 1 through 4. Each exposure cell has four numbers representing passes, conditional failures, audiometric failures and incomplete tests.

A. Calculating Noise Exposure Criteria

1. MIL-STD-1474D

MIL-STD-1474D (1997) is a Department of Defense Design Criteria Standard that provides specific noise limits and related requirements to weapons system designers and manufacturers.

These limits are intended to cover typical operational conditions, and the limits must not be exceeded if the materiel is to be acceptable for procurement. 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 values were determined for the free-field waveforms using the following formula,

$$L_m = L_p + 6.64\log\left(\frac{T_b}{200}\right) + 5\log(N) \tag{1}$$

The peak sound pressure level (L_p in dB SPL) was identified for the maximum positive pressure. The B-duration (T_b in milliseconds) was determined from a fit of the waveform envelope derived from the magnitude of the complex Hilbert transform of the pressure signal. The number of impulses (N = 6, 12, 25, 50 or 100) was used to estimate the MIL-STD-1474D level for each exposure cell.

2. A-weighted 8-hour Equivalent Level

L_{Aeq8hr} 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 French Committee on Weapons Noises advocated the use of A-weighted energy in the form of L_{Aeq8hr} as a noise exposure criterion for unprotected ears with a limit of 85 dBA (Direction Technique de Armements Terrestres, 1983). L_{Aeq8hr} 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 first convolved with a digital A-weighting finite impulse response filter and then squared, integrated and adjusted for duration and the number of impulses as follows:

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

where L_{Aeq8hr} is the equivalent 8-hour, A-weighted sound pressure level (dBA), p_0 is the reference pressure level (20 µPa), $p_A(t)$ is the A-weighted pressure time-waveform in Pascals, t_1 is 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).

3. The 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). The basilar membrane response is modeled coarsely by a 23-element network transmission line that corresponds to 1/3rd octave band intervals. An estimate of Auditory Risk Units (ARU) 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 ARU at any of the 23 segments is defined as the auditory hazard of an exposure.

Version 1.1 of the AHAAH model was downloaded from the Army Research Laboratory and used to process the pressure-time waveforms. Both the warned and unwarned ARUs were computed for each waveform. For each exposure cell, the average ARU for all waveforms was determined. The average ARU was multiplied by the number of impulses within an exposure cell to estimate the overall exposure. The auditory risk units were computed for the warned exposure conditions for 1-meter; 3-meter and 5-meter exposure conditions with the modified earmuff and also for the 5-meter exposure condition with the unmodified earmuff – as summarized in Tables 5 through 8.

B. Statistical Analysis

The performance of different noise exposure criteria was assessed in two ways: goodness-of-fit (calibration) and discrimination (Hosmer and Lemeshow, 2000, p. 163). Goodness-of-fit or calibration addresses the issue of how well the models fit the data. Discrimination addresses the matter of how well the model predicts failures.

1. Goodness-of-fit of Different Noise Exposure Criteria

The relationships of the hazard indices MIL-STD-1474D, L_{Aeq8hr} (free-field and protected), and AHAAH (warned and unwarned) to threshold shifts in hearing were modeled using subject-specific and population-average models. The subject-specific models were generalized linear mixed models (GLMM) (Fitzmaurice et al., 2004), and the population-average models were generalized estimating equations (GEE) (Liang and Zeger, 1986). The two dependent variables of interest were an outcome variable representing just audiometric failures and an outcome variable for which failure was deemed to occur if either an audiometric or conditional failure occurred. The data consisted of all the measurements from modified earmuffs; all of the data from participants who suffered auditory (or conditional) failures were included in the analysis.

For the subject-specific model a logistic GLMM was utilized such that

$$\ln\left\{\frac{P\left(Y_{ij}=1\,|\,b_i\right)}{P\left(Y_{ij}=0\,|\,b_i\right)}\right\}=b_i+\beta_0+\beta_1 X_{ij}\,, \tag{3}$$

is a random-intercept model (Fitzmaurice et al., 2004, p. 330) in which Y_{ij} is a binary random variable (equal to 0 or 1), i represents the subject, j represents the exposure series (1, 3, or 5-meters), P is the probability of a failure, β_0 and β_1 are the fitting coefficients for the logistic regression, X_{ij} is the independent variable and b_i is the subject-specific random component of the intercept. For example, Y_{ij} could be the dependent variable representing audiometric failure (fail=1 and pass=0) and X_{ij} could represent the value of the MIL-STD-1474D index. The GLMM's were fit using *gllamm* in Stata (Rabe-Hesketh and Skrondal, 2005).

For the population-average models a GEE model was used with the logit link:

$$\ln\left(\frac{\mu_{ij}}{1-\mu_{ij}}\right)=\beta_0+\beta_1 X_{ij} \tag{4}$$

(Fitzmaurice et al., 2004, p. 298) where $\mu_{ij} = E[Y_{ij}]$ is the expectation value of the Y_{ij}, Y_{ij} is the outcome binary random variable representing either audiometric or audiometric and conditional failures, and X_{ij} is the independent variable of interest. The logit function provides the relationship between the linear predictor and the mean of the distribution function, in this case the binomial or multinomial distribution. The coefficients β_0 and β_1 were estimated using *xtgee* in Stata. The failure rate is:

$$\hat{f}=\frac{1}{1+e^{-\left(\hat{\beta}_0+\hat{\beta}_1 x_{ij}\right)}}\,. \tag{5}$$

As shown by Chan et al. (2001), the variance of the quantity $\hat{\beta}_0+\hat{\beta}_1 L$, σ^2, depends on the level of the noise exposure metric, L, and is equal to:

$$\sigma^2 \;=\; \sigma_0^2 \;+\; 2\sigma_{01}^2 L \;+\; \sigma_1^2 L^2\,. \tag{6}$$

In Eq. 6, σ_0^2 is the variance of β_0, σ_1^2 is the variance of β_1, and σ_{01}^2 is the covariance of β_0 and β_1. The estimates of these parameters, $\hat{\sigma}_0^2$, $\hat{\sigma}_1^2$, and $\hat{\sigma}_{01}^2$, are obtained from the covariance matrix after estimating the GEE model with the use of the empirical or “sandwich” estimator (Fitzmaurice et al., 2004). Thus the standard deviation becomes:

$$\hat{\sigma} = \sqrt{\sigma_0^2 + 2\sigma_{01}^2 L + \sigma_1^2 L^2}\ , \tag{7}$$

and the 95% confidence limits are calculated as:

$$\left(\frac{1}{1+e^{-\left(\hat{\beta}_0+\hat{\beta}_1 x_{ij} - 1.96\hat{\sigma}\right)}}, \frac{1}{1+e^{-\left(\hat{\beta}_0+\hat{\beta}_1 x_{ij} + 1.96\hat{\sigma}\right)}} \right). \tag{8}$$

The graphs of the GEE models below also include failure rates calculated from the data (See Figures 2 – 11). Each open symbol represents the failure rate for a unique combination of distance, number of blasts, and level. For example, for the 1-meter Level 5-100 shots, six of the seventeen participants tested had either a conditional or audiometric failure. Thus the observed failure rate for that treatment combination was 0.3529. In order to calculate the $L_{(95,95)}$, the level of the index for which the upper confidence limit was equal to 0.05 was determined (Chan et al., 2001). Since $\hat{\sigma}$ depends on the level of the index, an iterative procedure to determine $L_{(95,95)}$ was used. First, the value of the index for which the mean failure rate is equal to 0.05 was calculated. To do so, Equations 4 and 5 are solved for an index value corresponding to a failure rate of 0.05 to obtain,

$$L_{95} = \frac{\left(\ln\left(\frac{0.05}{0.95}\right) - \hat{\beta}_0 \right)}{\hat{\beta}_1}. \tag{9}$$

Using L_{95}, the first estimate of the standard error, $\hat{\sigma}^{(1)}$, is obtained:

$$\hat{\sigma}^{(1)} = \sqrt{\hat{\sigma}_0^2 + 2\hat{\sigma}_{01}^2 L_{95} + \hat{\sigma}_1^2 L_{95}^2}\ . \tag{10}$$

Thus the next iteration of $L_{(95,95)}$ yields:

$$L_{(95,95)}^{(2)} = \frac{\left(\ln\left(\frac{0.05}{0.95}\right) - \hat{\beta}_0 - 1.96\hat{\sigma}^{(1)} \right)}{\hat{\beta}_1}. \tag{11}$$

The next iteration yields:

$$\hat{\sigma}^{(2)} = \sqrt{\hat{\sigma}_0^2 + 2\hat{\sigma}_{01}^2 L_{(95,95)}^{(2)} + \hat{\sigma}_1^2 \left(L_{(95,95)}^{2}\right)^2} \tag{12}$$

and

$$L_{(95,95)}^{(3)} = \frac{\left(\ln\left(\frac{0.05}{0.95} \right) - \hat{\beta}_0 - 1.96\hat{\sigma}^{(2)} \right)}{\hat{\beta}_1} \quad . \tag{13}$$

The iterations continue until the difference between $L_{(95,95)}^{(n-1)}$ and $L_{(95,95)}^{(n)}$ is less than 0.01, at which point the procedure is considered to have converged and $L_{(95,95)} = L_{(95,95)}^{(n)}$. The values of L_{95} and $L_{(95,95)}$ for different indices and outcome variables are shown in Table 9.

The fit of different models was assessed. Generalized linear mixed models – being likelihood-based – were compared using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) indices. The GEE models, being quasi-likelihood-based, were compared using the quasi-likelihood under the independence model criterion (QIC) (Pan, 2001) with an independence working correlation matrix. The AIC and BIC indices were obtained using *gllamm* in Stata and the QIC index was calculated using programs from Cui (2007). For the AIC, BIC, and QIC the smaller the value of the index, the better is the fit of the model.

Goodness-of-fit tests for population-average models were also performed using an extended Hosmer-Lemeshow method described in Horton et al. (1999) and Hardin and Hilbe (2003). The tests were performed using Wald tests, as in Hardin and Hilbe (2003). Although Horton et al. (1999) suggest the use of score tests, Hardin and Hilbe (2003) note that the results of score and Wald tests should not differ greatly for large data sets. In performing these tests for just audiometric failures, the observations were divided into three groups of equal size arranged in order of increasing risk. As noted in Hardin and Hilbe (2003, p. 168), any number of groups may be used. For this data set the use of a larger number of groups either resulted in lack of convergence or in such small *p*-values for all indices that the Goodness-of-fit test provided no useful basis for comparing the indices. Thus the following model was used:

$$\ln\left(\frac{\mu_{ij}}{1-\mu_{ij}} \right) = \beta_0 + \beta_1 X_{ij} + \gamma_1 I_1 + \gamma_2 I_2 \tag{14}$$

where $I_k = 1$ if there is a failure in category k and $I_k = 0$ if not, for $k = 1, 2$. (Only two indicator variables are needed to establish membership in one of the three categories.) A Wald test was used to determine if $\gamma_1 = \gamma_2 = 0$. Lastly, a model with four categories of risk was employed. Such a model requires three indicator variables, as in the following:

$$\ln\left(\frac{\mu_{ij}}{1-\mu_{ij}}\right) = \beta_0 + \beta_1 X_{ij} + \gamma_1 I_1 + \gamma_2 I_2 + \gamma_3 I_3 \qquad (15)$$

2. Discrimination of Different Noise Exposure Criteria

One of the important issues in regard to Price's work is his use of a two-by-two contingency table to demonstrate the value of his model in discriminating between individuals that will exhibit hearing failures (Price, 2007b). Such two-by-two contingency tables depend on the level at which a failure is assumed to occur. Clearly one can create a series of two-by-two contingency tables by assuming failure at different exposure thresholds. The ability of a variable to successfully predict an outcome is the basis for the receiver operator characteristic (ROC) curves. ROC analysis has been used widely in the medical field to assess the usefulness of different variables in predicting binary outcomes. The area under the ROC curve (AUC), also referred to as Harrell's (1996) C-index, is a measure of the variable's ability to predict. An AUC of 0.5 would represent a useless predictor and an AUC of 1.0 a perfect predictor. The suggestion has been made that the AUC may be viewed as follows (Hosmer and Lemeshow, 2000, p. 162):

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

This rule of thumb was presented in the context of independent observations, but as a rule, it is assumed to be of value with respect to clustered data.

In order to compare the discriminatory ability of different noise exposure criteria, AUCs have been calculated and compared for the various noise exposure criteria. This procedure is a well-established analysis for independent observations (see, for example, Agresti, 2002). However, the BOP data with its repeated measures on each individual are not independent observations. For this reason the nonparametric method of Obuchowski (1997) for clustered data was used. In this case the cluster consists of the observations taken from a single participant.

Observations taken in pre-specified cells [cells for which shots = 6 and/or level = 6 or (level = 5 and shots = 100)] and those taken on an *ad hoc* basis (those taken after audiometric or conditional failures) were analyzed separately, because of the different manner in which participants were chosen for these two situations. The software used to implement Obuchowski's (1997) method was the Stata procedure *somersd* (Newson, 2001). ROC curves were plotted using the Stata procedure *roccurve* (Pepe et al., 2008).

III. Results

A. Audiometric Data

The analyzed waveform data using MIL-STD-1474D, L_{Aeq8hr} and the AHAAH model and the statistical analyses are presented in this section. The temporary threshold shifts were coded and the numerical evaluations of the exposures were entered into the statistical model(s) described above.

Overall, participants wearing the unmodified and modified earmuff during the 5-meter exposures experienced the lowest incidence of audiometric or conditional failure while those participants wearing the modified earmuff in the 1-meter and 3-meter blast overpressure studies experienced higher incidence of TTS.

Table 1 summarizes the TTS data for the 1-meter free-field exposures with the modified earmuff. Sixty-five participants were exposed to impulses starting at Level 1-6 shots and 27 completed exposures at Level 6-100 shots. In total, 22 conditional failures and 17 audiometric failures were observed with the majority of the failures occurring at noise exposure Level 6. Twenty-four (37%) participants completed Level 6-100 with no audiometric or conditional failure. Due to the number of audiometric failures (Levels 5-6, 7-6, 6-12, 6-25, 5-50, and 6-100), exposures at 6-100 were terminated and the remaining participants were exposed at Level 5-100.

Table 2 summarizes the TTS data for the 3-meter free-field exposures with the modified earmuff. Sixty-eight participants were exposed to impulses starting at Level 1-6 shots and 27 completed exposures at Level 6-100 shots. In total, 17 conditional failures and 11 audiometric failures occurred, with majority of the failures occurring at noise exposure Level 6. Twenty-five (37%) participants completed Level 6-100 with no audiometric or conditional failure. Due to the number of audiometric failures (Levels 5-6, 6-6, 6-12, 6-25, 5-50, 6-50, and 6-100) exposures at 6-100 were terminated and the remaining participants were exposed at Level 5-100.

Table 3 summarizes the TTS data for the 5-meter free-field exposures with the modified earmuff. Fifty-nine participants were exposed to impulses starting at Level 1-6 and fifty-four completed exposures at Level 6-100 shots. In total, 7 conditional failures were observed. Only two audiometric failures were observed in the study and those occurred at Level 6-100. Fifty-two (88%) participants completed the exposure sequence with no audiometric or conditional failures.

Table 4 summarizes the TTS data from the 5-meter free-field exposures with the unmodified muff. While this analysis does not include the unmodified muffs, the information is included for completeness. Sixty-two participants were exposed at Level 1-6 and 39 participants (63%) completed Level 6-100. A single conditional failure was observed at Level 6-6. No audiometric failures were observed.

B. Waveform Data

Tables 5 through 8 summarize the evaluations of the acoustic waveforms according to the various noise exposure criteria. The first column indicates the exposure level in the matrix while the second column indicates the number of free-field waveforms evaluated to estimate the MIL-STD-1474D and unprotected L_{Aeq8hr} exposure metrics reported in the third and fourth columns. The fifth column reports the number of protected waveforms used to estimate the protected L_{Aeq8hr}, AHAAH warned and unwarned metrics in the sixth through eighth columns, respectively. The microphone underneath the protector was positioned in the concha of the participant's ear. Thus when the AHAAH model was evaluated, the insertion point of the waveform was "2" corresponding to the opening of the ear canal. The comparison between free-field and protected waveforms is complicated due to the presence of the modified earmuff. The 5-meter exposures with the unmodified muff are presented only for informational purposes but are excluded from the statistical analysis in the following section. Preliminary analyses for audiometric failures only using the 5-meter unmodified earmuffs yielded estimates of $L_{(95,95)}$ of 190 dB, 106 dBA and 1810 ARUs for the MIL-STD-1474D, L_{Aeq8hr} and AHAAH metrics, respectively. The analyses reported later excluding the 5-meter unmodified muffs showed slightly higher values of $L_{(95,95)}$. Since the free-field data will be compared to the protected data, the inclusion of the intact earmuff would confound the results. Each value given in the table represents the average of all the waveforms for that exposure distance and level. Furthermore, the average has been converted to the value for a single impulse. In this manner, the reader can easily convert the table to estimate the total exposure for the appropriate number of impulses. For MIL-STD-1474D, one would add 5 log(N) to the value in the table to determine the total exposure. For AHAAH, the total exposure is determined by N ARU. For L_{Aeq8hr}, one would add 10 log(N) to determine total exposure.

Table 5 summarizes the results for the 1-meter exposures. MIL-STD-1474D exposures ranged from Level 1, 171.8 to Level 6 188.9 dB. As noted before, the Level 7 waveforms were not available; however, Chan et al. (2001) reported the MIL-STD-1474D for 1-meter Level 7-6 shots as 194.1 dB. Applying a linear regression to the data from Chan et al. and our waveform data, the Level 7 single shot MIL-STD-1474D level is estimated to be 190.2 dB. For the L_{Aeq8hr} estimates, the average single impulse levels ranged from 96.2 to 110.1 dBA. Similarly the L_{Aeq8hr} data could be estimated for the free-field exposure at Level 7-6 shots. In Chan et al. Figure 7, the L_{Aeq8hr} for Level 7, 6 shots was about 119.7 dBA which would yield a single shot L_{Aeq8hr} of 112 dBA. Comparing to Level 6, a single shot was 110.1 dBA from our evaluation. Impulses increased by about 3 dB per exposure level. Even though the Level 7 free-field peak sound pressure level can be estimated, the results from that exposure cell were not included in the regression. For the protected L_{Aeq8hr}, the single shot exposure estimates ranged between 86.0 and 94.0 dBA and the energy of the impulses increased at each level except between Levels 4 and 5, where the energy decreased by 0.2 dB from 91.9 to 91.7 dBA. The difference is within the statistical error of the dataset. For the AHAAH model, the warned evaluation yielded between 17 and 70 ARU while the unwarned evaluation yielded between 97 and 293 ARU. Level 4 was found to be the most hazardous in both the warned and unwarned evaluations with the AHAAH model. In contrast, the Level 7 exposures were estimated to be the most hazardous according to both L_{Aeq8hr} and MIL-STD-1474D.

Table 6 summarizes the results for the 3-meter exposures. MIL-STD-1474D levels ranged from 172.2 dB to 188.4 dB for Levels 2 through 7 for a single impulse. As noted previously, Level 1 waveforms were not available for either the protected or free-field conditions. Using Chan et al. (2001), the values for the free-field MIL-STD-1474D and L_{Aeq8hr} for Level 1 six shots are 173.7 dB and 100.8 dBA which equate to 169.8 dB and 93.0 dBA, respectively for single shot free-field. The protected L_{Aeq8hr} ranged between 83.7 and 92.7 dBA and exhibited a monotonic increase in energy as a function of level. Between Levels 5 and 6, no increase in the energy was observed. For the AHAAH model, the warned evaluation yielded between 29 and 37 ARU and the unwarned evaluation yielded between 156 and 186 ARU. Level 3 was evaluated to be the most hazardous by the AHAAH model and level 4 was slightly less hazardous. In contrast, Level 7 was found to be the most hazardous according to MIL-STD-1474D and the L_{Aeq8hr} protected and free-field conditions.

Table 7 summarizes the results for the 5-meter exposures in which the participants wore the modified earmuff. From Level 1 to 7, MIL-STD-1474D values increased from 165 dB to 183 dB for a single shot impulse. L_{Aeq8hr} increased by about 3 dBA for each level increment for both the free-field and protected exposures. The warned AHAAH model increased slightly from 18.6 ARU to 37.7 ARU. The unwarned AHAAH model exhibited a disorganized response for Levels 1 to 4 and then increased from 88.7 ARU at Level 4 to a maximum of 118.4 ARU at Level 7. Across all DRCs the exposure at Level 7 was evaluated as being the most hazardous.

Table 8 summarizes the results for the 5-meter exposures with the unmodified earmuffs. The free-field data yielded the same MIL-STD-1474D and L_{Aeq8hr} results as the modified earmuffs in Table 7 because the same weight charges were used during the exposure. The L_{Aeq8hr} and AHAAH protected evaluations yield different results from the modified earmuff. Where the modified muff exhibited an inconsistent response for the AHAAH model, the warned and unwarned responses exhibit a monotonic growth in hazard from Level 1 (1.3 and 13.7 ARU) to Level 7 (25.7 and 109.7 ARU). The L_{Aeq8hr} exposures increased monotonically from 72.6 dBA to 92.3 dBA.

C. Statistical Analysis

1. Results related to 5% failure thresholds

Statistical analyses of the blast overpressure data are summarized in Table 9 and Table 10. The statistical model described above was applied to each evaluation metric and the failure rates without propagation of failure to other exposure cells. In Table 9, the estimates of the 5% failure threshold and the 95% lower confidence limit (L_{95} and $L_{(95,95)}$) are reported for the case where only audiometric failures are considered and the other case where both audiometric and conditional failures are considered. For the audiometric failures only, Chan et al. (2001) reported for MIL-STD-1474D L_{95} = 195.8 and $L_{(95,95)}$ = 193.5 dB[2]. The present analysis

---

[2] Chan's use of $L_{(95,50)}$ represents the estimate of the 5% failure threshold, and is equivalent to L_{95}.

found L_{95} = 196.3 dB and $L_{(95,95)}$ = 193.5 dB for MIL-STD-1474D. For L_{Aeq8hr}, Chan et al. (2001) reported L_{95} = 126.6 dBA and $L_{(95,95)}$ = 124.1dBA. This analysis estimated 126.0 and 123.8 dBA for L_{95} and $L_{(95,95)}$, respectively. The results exhibited excellent agreement with Chan. For the protected L_{Aeq8hr}, L_{95} was 113.1 dBA and $L_{(95,95)}$ was 109.5 dBA. The results for the AHAAH model were L_{95} = 13,064 ARU and $L_{(95,95)}$ = 10,109 ARU with the unwarned condition and L_{95} = 3,247 ARU and $L_{(95,95)}$ = 2,480 ARU with the warned model.

When both audiometric and conditional failures are considered, the failure thresholds were lower than when only audiometric failures were analyzed. For MIL-STD-1474D, L_{95} = 188.7 dB and $L_{(95,95)}$ = 185.1 dB. For the Free-field L_{Aeq8hr}, L_{95} = 118.8 dBA and $L_{(95,95)}$ = 115.4 dBA. These values also compare favorably with Chan et al.'s presumed failure results for MIL-STD-1474D of $L_{(95,95)}$ = 185.3 dB and $L_{(95,95)}$ = 112.7 dBA for L_{Aeq8hr}. The protected L_{Aeq8hr} analysis yielded L_{95} = 104.6 dBA and $L_{(95,95)}$ = 101.4 dBA. The results for the AHAAH model were L_{95} = 5,710 ARU and $L_{(95,95)}$ = 3,053 ARU with the unwarned model and L_{95} = 1,378 ARU and $L_{(95,95)}$ = 719 ARU with the warned model.

2. Goodness-of-fit

Some of the models fit the data better than others. The independent variable that yielded the best fit as judged by the AIC and BIC for the generalized linear mixed models and the QIC for the GEE models was the free-field L_{Aeq8hr}, followed by the MIL-STD-1474D. The poorest fit was obtained with the warned AHAAH according to the AIC, BIC, and QIC; the results for the unwarned AHAAH were similar to those for the warned.

The results using the Horton et al. (1999) extended Hosmer-Lemeshow tests are shown in Tables 11 to 13. Tables 11 and 12 use three categories of risk to provide evidence that both of the L_{Aeq8hr} indices and the MIL-STD-1474D index exhibit a good fit ($p > 0.05$) and the AHAAH indices do not ($p < 0.05$). The result using a model with four categories of risk, shown in Table 13, indicate a good fit for the L_{Aeq8hr} free-field and protected models, but not the other three. The differences between the Tables 12 and 13 are probably due to the selection of different cutpoints (points separating the categories). The selection of different cutpoints is known to affect the results of Hosmer-Lemeshow type tests in regular logistic regression (Hosmer et al., 1997). The salient point here, though, is that the results for Table 12 and Table 13 are generally similar, with both L_{Aeq8hr} indices showing a good fit and both of the AHAAH indices showing a poor fit. Only the results for MIL-STD-1474D are problematic, being non-significant in Table 12 (i.e., showing a good fit) and statistically significant in Table 13.

In general, Figures 2 through 11 show the results of the GEE regression for the failure rates within the different exposure cells computed with the audiometric failure (or the audiometric and conditional failures) as a function of the level of the associated metric. The failure rates for the three exposures are depicted as open symbols (1-meter: circles, 3-meter: squares, 5-meter: diamonds). Horizontal and vertical reference lines are depicted for the 5% failure threshold and the current/suggested exposure limit for the metric being analyzed. At the intersection of the 5% failure rate with the regression curve and the upper confidence limit, two filled dots depict the $L_{(95,95)}$ and L_{95} points as derived from the analysis. The legend reports the fitted values for $L_{(95,95)}$ and L_{95}. The $L_{(95,95)}$ value corresponds to the point where the upper confidence limit crosses the 5% failure line. More sophisticated analysis (i.e. bootstrap or Monte Carlo methods) could be performed to estimate the error as a function of the associated metric which would yield a horizontal error bar.

Figure 2 illustrates GEE curve-fitting for only the audiometric failures with the MIL-STD-1474D metric. The 5% failure thresholds are $L_{(95,95)}$ = 193.5 and L_{95} = 196.3 dB. The current MIL-STD-1474D has effectively set 177 dB as the hazardous limit for single level hearing protection. When the effect of multiple impulses are added to the single shot exposures, only four exposures fell below the 177-dB limit. An important feature of the data is the organization of increasing failure rate with increased level. In one instance of the 1-meter exposure, the failure rate was substantially above the general trend and occurred at a level lower than the $L_{(95,95)}$ threshold.

Figure 3 illustrates GEE curve-fitting for both the audiometric and conditional failures with the MIL-STD-1474D metric. The 5% failure thresholds are $L_{(95,95)}$ = 185.1 and L_{95} = 188.7 dB. Only four exposures fell below the 177-dB limit using the MIL-STD-1474D. The organization of increasing failure with increased level is more evident with the addition of the conditional failures. Inclusion of the conditional failures increased the overall failure rates and lowered the $L_{(95,95)}$ threshold derived from the data. However, one of the 3-meter exposure cells was evaluated at about 184 dB and had a failure rate above 5%.

Figure 4 illustrates GEE curve-fitting for only the audiometric failure rates with the free-field L_{Aeq8hr} metric. The 5% failure thresholds are $L_{(95,95)}$ = 123.6 and L_{95} = 126.0 dBA. The data tended to exhibit a similar organization as seen in the MIL-STD-1474D evaluations. Two of the more energetic exposures at 1 meter and 3 meters were evaluated at levels above the L_{95} derived from the analysis.

Figure 5 illustrates GEE curve-fitting for both the audiometric and conditional failure rates with the free-field L_{Aeq8hr} metric. The 5% failure thresholds are $L_{(95,95)}$ = 115.4 and L_{95} = 118.8 dBA. Also for L_{Aeq8hr}, the threshold limit is 100 dBA at the ear and assumes a 15-dB effectiveness of the modified earmuff.

Figure 6 illustrates GEE curve-fitting for only the audiometric failure rates with the protected L_{Aeq8hr} metric. The 5% failure thresholds are $L_{(95,95)}$ = 109.5 and L_{95} = 113.1 dBA. Accounting for multiple impulses, none of the exposures were below the 85-dBA threshold limit for the L_{Aeq8hr} metric at the ear.

Figure 7 illustrates GEE curve-fitting for both the audiometric and conditional failure rates with the protected L_{Aeq8hr} metric. The 5% failure thresholds are $L_{(95,95)}$ = 101.4 and L_{95} = 104.6 and dBA. For L_{Aeq8hr}, the threshold limit is 85 dBA at the ear. Assuming that the hearing protector can provide about 15 dB of protection, the threshold limit would become 100 dBA (Chan et al. 2001). The $L_{(95,95)}$ value of 101.4 is in close agreement.

Figure 8 illustrates GEE curve-fitting for the audiometric failure rates with the unwarned AHAAH metric. The 5% failure thresholds are $L_{(95,95)}$ = 10,108.7 and L_{95} = 13,064.3 ARU. For AHAAH, the suggested damage-risk criterion is 500 ARU. Every exposure evaluated with the unwarned AHAAH model would have been judged to be hazardous.

Figure 9 illustrates GEE curve-fitting for both the audiometric and conditional failure rates with the unwarned AHAAH metric. The 5% failure thresholds are $L_{(95,95)}$ = 3,053.4 and L_{95} = 5,710.5 ARU. Adding the conditional failures clearly increases the failure rates as has been observed in the previous metrics. The unwarned AHAAH metric exhibits organization of increasing failure rate with increased hazard.

Figure 10 illustrates GEE curve-fitting for the audiometric failure rates with the warned AHAAH metric. The 5% failure thresholds are $L_{(95,95)}$ = 2,479.8 and L_{95} = 3,247.2 ARU. Assuming the warned condition, the majority of exposures were below the suggested 500 ARU.

Finally, Figure 11 illustrates GEE curve-fitting for both the audiometric and conditional failure rates with the warned AHAAH metric. The 5% failure thresholds are $L_{(95,95)}$ = 718.7 and L_{95} = 1,377.6 ARU. Above 500 ARU, the exposures appear to exhibit the organization that would be expected.

3. Discrimination

For just audiometric failures the areas under the ROC curves were calculated for the pre-specified [cells for which shots = 6 and/or level =6 or (level=5 and shots=100)] and *ad hoc* [those taken after audiometric or conditional failures] observations separately. The areas under the ROC curves using the pre-specified cells were calculated for the different noise exposure criteria based on 1,783 observations. The AUCs (Harrell's C-index) for each noise exposure criterion are shown in Table 15. The differences in the AUCs are given in Table 16. Thus, there is evidence that the AUC is significantly greater for the unprotected L_{Aeq8hr} criterion than the protected L_{Aeq8hr}, unwarned AHAAH, and warned AHAAH criteria. The AUC for the unwarned AHAAH is greater than that for the warned AHAAH. Shown in Figure 12 are the two ROC curves for unprotected L_{Aeq8hr} and unwarned AHAAH. For just audiometric failures the areas under the ROC curves were calculated using 77 observations from the *ad hoc* cells, as shown in Table 17. The results shown in Table 17 provide insufficient reason to reject the hypothesis that the AUC differs significantly from 0.5 for the L_{Aeq8hr} unprotected, L_{Aeq8hr} protected, MIL-STD-1474D, and unwarned AHAAH, given that the 95% confidence interval for these indices includes 0.5. In the same vein, the confidence interval for the warned AHAAH barely excludes 0.5, thus providing little reason to compare the levels.

The areas under the ROC curves were calculated for audiometric and conditional failures combined, again treating the pre-specified and *ad hoc* observations separately. The areas under the ROC curves for the different noise exposure criteria for pre-specified cells were, again based on 1,783 observations, are shown in Table 18. In Table 19, the differences demonstrate that the unprotected L_{Aeq8hr} yields a significantly larger AUC than any of the other indices, except the MIL-STD-1474D. In addition, the AUC for the MIL-STD-1474D is significantly larger than that for protected L_{Aeq8hr} and warned AHAAH. The ROC curves for audiometric and conditional failures combined for the unprotected L_{Aeq8hr} and unwarned AHAAH noise exposure criteria are shown in Figure 13. Based on 77 observations, the areas under the ROC curves for audiometric and conditional failures combined using the *ad hoc* cells are given in Table 20. As in the case of just the audiometric failures, there was no reason to reject the hypothesis that the AUC differs significantly from 0.5 for the L_{Aeq8hr} unprotected, L_{Aeq8hr} protected, MIL-STD-1474D, and unwarned AHAAH for the audiometric and conditional failures combined. Along the same lines, the lower confidence limit of the confidence interval for the warned AHAAH is only a little above 0.5. Thus, no compelling reason to compare the levels exists.

IV. Discussion

A. The Effect of Hearing Protection

Both the MIL-STD-1474D and the free-field L_{Aeq8hr} exposure criteria must make some accommodation for the effect of hearing protection on the impulse waveform. Using MIL-STD-1474, a single hearing protector allows an additional 29 dB of peak sound pressure level energy for an impulse. This allowance was derived from the difference in peak pressure level for the Light Anti-tank Weapon (LAW) rocket that produced the same TTS in human participants at different exposure differences [Price, personal communication]. The V-51R flanged earplug was formerly the standard issue hearing protection device for the U.S. Army and had a Noise Reduction Rating of 24 dB. While that estimate may be accurate for a well-fit product, the naïve subject-fit attenuation data for the V-51R suggest that about 50% of users achieve less than 10 dB of attenuation for frequencies below 500 Hz (Franks et al., 2000; Murphy et al., 2004) whereas the manufacturer's data would suggest that they achieve 20 to 24 dB of attenuation. Similarly, the E•A•R Classic earplug has a noise reduction rating of 29 dB and has been shown to be highly effective against impulse noise when properly worn (Johnson, 1994). When naïve subjects fit the E•A•R Classic, the A-weighted attenuations are typically about 15 to 20 dB. For experienced users, attenuations can range from 20 to 40 dB (Franks et al., 2000; Murphy et al., 2004; Murphy et al., 2009).

For L_{Aeq8hr}, the threshold limit is 85 dBA at the ear. Chan et al. (2001) assumed that the modified earmuff provided 15 dB of effective attenuation and set a threshold limit value of 100 dBA. For both the analyses with audiometric failure only and combined audiometric and conditional failures in Table 9, the differences between the free-field and protected $L_{(95,95)}$ thresholds is about 14 dBA, which supports Chan's assumption. The variability of the failure data affects the $L_{(95,95)}$ value. More failures or more variability across exposure levels would lower the $L_{(95,95)}$ for the protected condition, thus decreasing the effective attenuation. On the contrary, fewer failures or less variability would increase $L_{(95,95)}$ for the protected L_{Aeq8hr} and would increase the effective attenuation. This agreement between the 14-dB difference in our $L_{(95,95)}$ thresholds and Chan's 15-dB assumption is serendipitous.

Using the L_{Aeq8hr} data for the different exposures of the modified earmuff in the free-field and protected conditions, the differences are plotted in Figure 14. A quadratic function was fit to the data to illustrate the trend of increasing attenuation with free-field L_{Aeq8hr} level. In general, the differences start at about 9 dBA and rise to almost 20 dBA. The data for the unmodified muff are not plotted in this figure; however, the differences can be estimated to be between 18 and 20 dBA from the data in Table 8.

Hearing protectors cannot be assumed to provide a constant level of protection over a range of peak impulse levels when developing damage-risk criteria. The performance of protection has been shown to vary substantially from 0 dB for some nonlinear orifice products to well over 30 dB as the peak impulse sound pressure level increases (Allen and Berger, 1990; Berger and Hamery, 2008; Murphy et al., 2007; Zera and Mlynski, 2007). While none of the candidate noise exposure criteria evaluated in this paper explicitly dealt with the variable attenuation, the AHAAH model has a module where hearing protection is assumed to function as a minimum phase filter. Patterson and Ahroon (2005) evaluated the minimum phase filter module of the AHAAH model and demonstrated poor agreement between predicted and measured performance. A good hearing protection models must accurately predict the protector response and the transmission of sound through the protector to the ear.

B. Comparisons with Chan et al. (2001)

Chan et al. (2001) used a statistical model similar to the model used in this paper. They analyzed the results with propagation of failure (Patterson and Johnson, 1994) and without any presumed failures. When a participant suffered an audiometric failure at a lower level or number of impulses, the failure was propagated (counted) as a failure for all higher-level exposures and number of impulses. Propagation of failure was considered in the development of the current analysis. Price (2007b) noted that propagation of failure leads to the creation of results where persons were exposed in small numbers or were not exposed at all (See Tables 1 through 4). Patterson and Johnson assumed that an audiometric failure would propagate to higher energy levels and to greater numbers of impulses. For a conditional failure, the participant was presumed to have failed at the next more energetic level and at increasing numbers of impulses. For at least one participant, this presumption was invalid. The participant failed at Level 2-6 shots, passed Level 1-12 shots and proceeded to pass Level 2-12 shots and completed the series of exposures. Not propagating failure to other exposure cells will result in higher estimates of the L_{95} and $L_{(95,95)}$ (Chan et al. 2001). The present analysis also has included the conditional failures to capture those persons that might have exhibited audiometric failure at levels and numbers of impulses above that where the conditional failure was observed. Interestingly, the MIL-STD-1474D values Chan et al. (2001) obtained assuming failures for free-field data (Table IV in their paper) were rather close to those obtained assuming that failures consisted of audiometric and conditional failures. Chan et al. (2001) obtained an $L_{(95,95)}$ of 185.3 for MIL-STD-1474D, whereas we obtained 185.1 using audiometric and conditional failures. Similarly, Chan et al. (2001) obtained an $L_{(95,95)}$ for free-field L_{Aeq8hr} of 112.7 compared to the value of 115.4 that the present analysis found.

As can be seen from Tables 5 through 8, the next exposure level always resulted in a larger value for MIL-STD-1474D and free-field L_{Aeq8hr}. However for the AHAAH model, Levels 3 and 4 exhibited the greatest hazard and would yield different results. In all of the metrics, the estimated risk increased with the number of impulses: MIL-STD-1474D adds 5 log(N); L_{Aeq8hr} adds 10 log(N); AHAAH multiplies the ARUs for a single impulse by N. By not propagating failure across the table, the rank ordering of the risks will be different from those observed by others. For instance at 1-meter, exposure Level 6-50 has a lower failure rate than Level 6-12 and Level 6-25 when just audiometric failures are counted. However, to assume that a participant that failed at a lower number of shots would fail at a higher level is not assured and therefore biases the results. To verify such an assumption would not be an ethical use of human participants as the participant may or may not incur a permanent threshold shift (although it is assumed that they will incur a larger temporary threshold shift).

Where Patterson and Johnson (1994), Chan et al. (2001) and Patterson and Ahroon (2005) propagated failure from both lower to higher exposure levels and from a lesser to a greater number of impulses, Price (2007b) only propagated from a lesser to a greater number of impulses, since those exposure cells are the only ones guaranteed to have greater risk (ARUs) under the AHAAH model. For example, at Levels 5 and 6, the average warned ARUs are effectively equal for the 1-meter and 3-meter exposures (see Tables 5 and 6). For the 3-meter exposure, the average unwarned Level 6 ARUs were greater than Level 5 ARUs. For the 1-meter exposure, the average unwarned Level 6 ARUs were slightly less than Level 5 ARUs. Propagation only within a level assumes that a lower number of impulses are less hazardous than a greater number of impulses. For those levels with comparable hazard according to the AHAAH model, the propagation should be allowed to move diagonally (e.g. 1-meter Level 5-6 shots to Level 6-12 shots). Chan and Ho (2005) identified that this assumption changed the evaluation of several cells from hazardous to safe: 1 meter, Level 7-6 and 6-12; 3 meters Level 6-6, 6-12 and 7-6. These cells exhibited conditional and audiometric failures in varying degrees. These particular cells can no longer be classified as safe with 95% confidence when fewer than 59 subjects were exposed and did not experience TTS > 25 dB. Thus Price's analysis of the 1 and 3-meter exposures is not self-consistent and fails to propagate failure correctly.

Propagation of failure has been an item of contention among those evaluating the BOP studies. Initial efforts to conduct the analysis of the BOP data included different assumptions to propagate failures; however, the authors came to the realization that propagation effectively creates data where none was collected. The present analysis does not propagate failure and has reached conclusions regarding the quality of fit. The inclusion of the conditional failures provides further support to the rank ordering of the fits according to the AIC, BIC and QIC statistics. While these statistics may not settle the debate, they are useful and have provided evidence contrary to that claimed by proponents of the AHAAH model.

C. Effect of the warned/unwarned state of the ear

In this analysis, both the warned and unwarned conditions of the AHAAH model have been considered. Wholesale assignment of the status of the participants in the BOP study to the warned condition is questionable. Price (2007b) has maintained that the participants were exposed in a warned state because countdown preceded each of the impulses in the 1, 3 and 5-meter exposures. The countdowns presumably allowed the participants to "steel" themselves, tense their bodies and clench their teeth in anticipation of the pending blast wave. These actions are presumed to activate the participants' middle ear reflex.

Further evidence identified by Price was the limited testing conducted without a countdown (for a discussion of the results see Johnson, 1994). Following the completion of the 3-meter exposures, only those soldiers that completed the exposure with no TTS were given an opportunity to participate in the no-countdown study. Twenty soldiers volunteered to participate in the study; four soldiers completed the no-countdown exposure sequence. None of the participants were failed for audiometric reasons. One soldier had a potential conditional failure. The most common reason given for leaving the study prematurely was the stress produced by not knowing when the blast was going to occur. The participants may have been predisposed to include robust ears and possibly only those with functioning acoustic reflexes were chosen, considering they had completed the 3-meter exposures. The statistical power associated with testing 20 persons is insufficient to identify a safe exposure at the 95% confidence level using the binomial power calculation described in the original experimental protocol (Johnson, 1994; Patterson and Ahroon, 2005). If three audiometric failures had been observed, 20 subjects would be sufficient to identify an exposure as hazardous. Because the status of the reflex was never measured, the no-countdown study does not confirm the validity of the AHAAH model warned-ear hypothesis; instead it demonstrates the intolerability of the exposure conditions.

Price (2007b) has cited several studies to support the claim that middle ear reflexes can be conditioned and would have been the appropriate assumption for the BOP exposures because a countdown was provided. This assumption must be considered more carefully. Firstly, the degree to which the middle ear reflex is activated by a conditioned response is unknown. Even though an active middle ear reflex is normal, none of the BOP participants were given a test of the acoustic reflex during the pretest audiometric examination. Thus no evidence exists that ***all*** of the participants had functioning middle ear reflexes and no evidence of the timing of the reflex activation exists for ***any*** of the participants. In Marshall et al. (1975), thirteen participants were tested for a conditioned reflex response to handling or being shown a toy cap pistol. Two of the participants did not exhibit middle ear reflex. Bates et al. (1970) attempted to condition a middle ear reflex in twelve participants. Six of the twelve participants showed little or no evidence of conditioning. Fletcher and Riopelle (1960) conducted an investigation of providing a preimpulse stimulus to activate the middle ear reflex. In their study they found that seven of twenty-four participants did not exhibit an effective difference in TTS with and without the preimpulse stimulus. Furthermore, they found significant effects in the analysis of variance for the frequency of the stimulus, for the individual and for the effectiveness of the protection with individual. The middle ear muscles may have already been activated during the exposures conducted by Fletcher and Riopelle. If this assumption were true, then the protective effect of the middle ear reflex might not have been observed. In Price's study of cats exposed to high-level impulses, the effect of anesthesia (Atropine and Sodium Pentobarbital) eliminated any protection from the middle ear reflex and increased threshold shifts were observed (Price and Wansack 1989; Price, 2007a).

The important point to note in this discussion is not whether the middle ear reflex provides any prophylaxis against impulse noise; the middle ear reflex is commonly believed to provide 10 to 15 dB protection. Rather the pertinent questions are as follows:

- Is the reflex guaranteed to be present in all persons?

- Is the ability for the middle ear reflex to be conditioned a realistic assumption for all situations?
- Is the effectiveness of the reflex is the same across all persons?

The answer to those questions is no. Impulse noise cannot always be anticipated and the weapons are not always fired by a single person in isolation. When firing weapons on a range or in battle, an individual will almost certainly be exposed to impulses produced by other shooters in the area. Consequently, the warned AHAAH model will underestimate the actual exposure to unanticipated impulses. Furthermore, the location of the exposed ear relative to the impulse source presents a greater risk than knowing the condition of the middle ear reflex. Recent analysis of the levels around an M-16 suggests that proximity and angle relative to the muzzle can produce as much as 34-dB difference in the peak sound pressure levels (Rasmussen et al., 2009). The damage-risk criteria apply to all persons, not just to those who exhibit predefined protective effects. Thus the assumption of a warned condition is unsubstantiated.

D. Performance of different noise exposure criteria

One of the purposes of this analysis was to identify which noise exposure metric provides a best fit to the TTS data for the BOP data set. Of the five exposure models evaluated in Table 10, the free-field L_{Aeq8hr} had the lowest AIC, BIC and QIC values indicating it was a better fit than the other models. MIL-STD-1474D followed by the protected L_{Aeq8hr} metric had the next best scores in that order. The warned AHAAH model had the highest AIC, BIC and QIC scores indicating the poorest fit. Both audiometric failure and audiometric or conditional failure demonstrate a similar pattern for all five noise exposure criteria as shown in Table 1.

According to Fitzmaurice et al. (2004, p.385), when data are not missing completely at random (MCAR) the GEE method will yield biased estimates of the coefficients. In the blast overpressure study, when an audiometric or conditional failure occurred, the progression through the exposure matrix was modified resulting in generation of data that was not MCAR. Since participants likely to fail were not advanced to higher exposure levels and possibly removed from the study, the GEE method may have resulted in underestimating exposure failure rates. The effect of propagating failure would result in a higher failure rate for the more intense exposures thus moving the fitted curve to the left and lowering the failure threshold.

According to Raftery (1995, pp. 139-141) if BIC differences of greater than 10 exist between any two models, then there is strong evidence of differences in goodness-of-fit between the two models (See Table 14). For the audiometric failure data, the difference in the

BIC goodness-of-fit criteria (36.99) between the L_{Aeq8hr} free-field (275.2) and AHAAH-warned (312.2) was strong evidence in favor of the free-field L_{Aeq8hr} model. As one progresses from the lowest to the highest BIC score, one observes several other BIC differences of more than 10. The small difference in the BIC score for the protected L_{Aeq8hr} (311.1) and warned AHAAH model (312.2) demonstrates little or no difference in goodness-of-fit for these two noise exposure criteria. For the fits performed with both the audiometric and conditional failures, the patterns were essentially the same for the BIC. The free-field L_{Aeq8hr} had the lowest BIC score (550.7) and the warned AHAAH (590.8) the highest. Seven of the possible pairs of noise exposure criteria for conditional and audiometric failure in Table 10 displayed BIC differences greater than 10, including those of MIL-STD-1474D. The AHAAH-warned had the highest absolute BIC (590.81) which was very close to the AHAAH-unwarned BIC, suggesting that the goodness-of-fit for these two exposure criteria to the BOP injury data was virtually identical. Whether the conditional failures are included or excluded, the free-field L_{Aeq8hr} model provides the best fit to the TTS data.

The unprotected L_{Aeq8hr} exposure criteria showed discrimination superior to all the other criteria except MIL-STD-1474D through having significantly larger AUCs than the other criteria. The fact that the AUCs are significantly larger than those for the unwarned AHAAH and warned AHAAH criteria is particularly interesting in light of Price’s assertion that the AHAAH criteria show better discrimination in his contingency table (Price, 2007b). Recall that the ROC curve essentially summarizes the results of a whole series of two-by-two contingency tables. Price may have been correct in his assertion regarding the one table he chose to present. However, a more thorough examination of the problem, as provided with the ROC curves and the AUCs, indicates that the free-field L_{Aeq8hr} exposure criteria offers better discrimination than the AHAAH exposure criteria.

The main purpose of this paper has been to identify which, if any, noise exposure criterion best describes the BOP data. The unprotected L_{Aeq8hr} exposure criterion performs the best – both in terms of calibration (goodness-of-fit) and discrimination. The next step, establishing a risk threshold, requires careful thought. In doing so, the modeling approach (GEE or GLMM) and utilization of the model must be carefully considered. Furthermore, the accumulation of risk due to multiple impulses must be addressed and finally the damage-risk criteria must be applied to other exposures.

The two main modeling approaches presented in this paper are GEE and GLMM. Let us assume that the first issue is to identify a level of sound (for example, as measured by the unprotected L_{Aeq8hr} criterion) which represents a level at which 5% of the population would be at risk for injury. The GEE approach allows us to identify a level at which 5% (on average) of the population would suffer injury. The GEE approach makes no distinctions between individuals. Conversely, the GLMM approach takes into account variability between individuals. The fact that some individuals are probably more susceptible to hearing loss than others thus suggests that the GLMM approach may generate a more biologically realistic model. Practically speaking, the GLMM method identifies the level of sound at which the 5% most susceptible individuals suffer injury and should be used to establish protection limits.

The second important issue is how to define a 95% confidence interval around the level at which the failure rate is 5%. The $L_{(95,95)}$ does not actually do this. The $L_{(95,95)}$ is determined from the intersection of the upper limit of the GEE model 95% confidence interval with the 5% failure rate. By definition, this exposure level will be lower than L_{95}. A better estimate of the $L_{(95,95)}$ could be found by using computational statistics to resample the data for the variance of the L_{95}. One could perhaps utilize the methodology used in determining benchmark doses. In any event, the establishment of a better method to estimate the risk threshold is an important and rather urgent area for future research.

In Chan et al., (2001), a best-fit model was developed by fitting the MIL-STD-1474D coefficients for B-duration and accumulation of risk due to multiple impulses. The best-fit B-duration coefficient was -16.67 and the accumulation risk coefficient was 3.08. Kim (2008) has suggested that the differential response of a single degree of freedom oscillator to impulses with different A-duration (e.g. 1, 3 and 5-meter exposures) could explain the negative B-duration coefficient. The accumulation of risk is the more important issue in the development of any damage-risk criteria. The BOP data are flawed for this purpose. While the data provide increasing numbers of impulses, the exposure groups are flawed by the removal of the potentially more susceptible participants.

The AHAAH model and L_{Aeq8hr} seemingly have different rates for accumulating risk due to multiple impulses (N). AHAAH multiplies the ARUs by a factor of N, while L_{Aeq8hr} adds 10 log(N). At face value these seem to be quite different. The AHAAH model can be converted to a logarithmic quantity by dividing the ARUs by an appropriate factor. Two potential candidate factors would be the proposed risk threshold (500 ARU) or a minimum threshold output by the model (10^{-5} ARU). This approach allows ARUs to be treated as dB values,

$$Dose = 10 \log\left(\frac{\text{ARU}}{10^{-5}}\right) + 10 \log(N) \qquad (16)$$

Whereas L_{Aeq8hr} sums the energy in a waveform to which the ear is exposed, AHAAH essentially sums the squared displacement at individual locations on the basilar membrane. The AHAAH model uses the maximum sum to characterize the risk of an exposure and that maximum comes from one location on the basilar membrane. The proposed use of the ARU is to simply sum different events. However, the combination of risks from different places on the basilar membrane complicates the evaluation of blast noise exposures with varied spectra. The longer A-durations from large caliber weapons yield more low frequency (~100-400 Hz) content than small-caliber weapons (~500 to 1000 Hz). L_{Aeq8hr} and MIL-STD-1474D don't make a specific distinction with regards to spectral content. L_{Aeq8hr} assumes that the waveform is filtered by the transfer function of the outer and middle ear nominally described by A-weighting. MIL-STD-1474D decreases the damage-risk criteria for the X, Y and Z curves as the B-duration increases. B-duration in a free-field environment may be representative of the A-duration that governs the dominant frequency of the Friedlander-like impulse. However in reverberant environments, B-duration yields the decay envelope of the room where the impulse occurs. MIL-STD-1474D does not distinguish at all between different spectra for reverberant exposures.

Price has identified that the L_{Aeq8hr} metric doesn't address the differences in hearing loss observed in exposures to small caliber weapons (rifles) versus large caliber weapons (howitzers). The BOP exposures were one series of studies designed to vary three critical parameters of blast overpressure: the peak amplitude, the A-duration and the numbers of impulses. This analysis has not attempted to draw conclusions outside of the BOP data set. The meta-analysis conducted by Price (2007b) is thought provoking, but flawed. The AHAAH model requires waveform data to analyze the risk of a specific impulse. Many of the exposures in Price's analysis did not have the actual waveforms from the exposures reported. Price analyzed waveforms that were collected in similar conditions and from similar weapons in several cases. If the details of the waveform are critical to an accurate evaluation, then conjectures of the discrimination and superior performance of any model are irrelevant.

A similar criticism may be levied against the BOP data and the initial analyses. While the investigators carefully documented waveforms, they do not represent any individual's actual exposure. Johnson (1994) and Patterson and Johnson (1994), Chan et al., (2001) and Price (2007b) and the present analysis assumed that the data from the several impulses at a particular level were representative of the exposures for all impulses received by all participants for that exposure distance. As noted above, the exposures of the investigators (not the actual participants) to the blast impulses were used to fill in the missing data for some exposure conditions. Because the waveforms were collected using the same level of charge and in the same positions as the actual participants the substitution is reasonable.

Without question, other human exposures have bearing on the interpretation of these results. However, the lack of detailed information regarding the exposures, the exposure waveforms and the outcomes precludes combining the results for a meta-analysis. Additional exposures that may replicate past studies should be conducted. Exposure conditions should be designed that differentiate predictions of the AHAAH model, L_{Aeq8hr}, and MIL-STD-1474D. The BOP study has been debated for more than a decade to prove/disprove theories of the effects of impulse noise on the human auditory system. Unless critical experiments are designed specifically to exploit features or weaknesses, the debate will continue unresolved.

V. Conclusions

The development of damage-risk criteria for impulse noise is a critical need for the military and an important occupational issue. The exposure limits that undergird MIL-STD-1474D also form the basis of maximum exposure limits to impulsive or continuous noise recommended by OSHA, MSHA and NIOSH (Ward, 1968; OSHA 1983; NIOSH 1998; MSHA 1999).

The current MIL-STD-1474D makes allowance for single and double hearing protection. As discussed above, the use of a 29-dB attenuation factor for a single level of hearing protection is overly optimistic. The L_{Aeq8hr} free-field metric does not consider any hearing protector. If peak reduction data were available for a specific protector, then free-field measurements could be transformed to a protected level. A minimum-phase filter model for the transmission of impulsive sound through a hearing protector has been a part of the AHAAH model almost since its inception. However, the minimum-phase model performed poorly in the analysis by Patterson and Ahroon (2005). Subsequent models of hearing protection have been developed using acoustic elements to model the response. The models may have merit and should be evaluated with data collected with actual subjects or with acoustic test fixtures.

Regardless of the shortcomings of the damage-risk criteria considered in this report, hearing protection is often the primary means of controlling the exposures of military personnel. Better models of protector performance must be developed and incorporated into the analysis of impulsive noise exposures.

The performance of models of various phenomena may be evaluated in terms of how well they fit the data (calibration) and how well they predict future events (discrimination). Noise exposure criteria are models of how noise affects hearing. These models have been evaluated and the unprotected L_{Aeq8hr} noise exposure criterion performed better than the others both in terms of goodness-of-fit and discrimination, except, as noted for the MIL-STD-1474D criteria with regard to discrimination (no difference).

While the actual determination of a risk threshold has not been addressed in this paper, the groundwork for such a determination has been laid and the associated issues have been presented.

VI. Acknowledgements

The authors wish to thank Drs. Edward Krieg and Rickie Davis (NIOSH), Dr. Philemon Chan (L3Com), Dr. William Ahroon (USAARL) and Dr. G. Richard Price (AHAAH Analysis) for their careful review of the draft manuscript. Their contributions significantly improved the manuscript. We also wish to thank the U.S. Army Medical Research and Materiel Command, U.S. Army Aeromedical Research Laboratories and the U.S. Army Research Laboratory for their direct and indirect support of this effort.

VII. References

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

C.H. Allen and E.H. Berger. Development of a unique passive hearing protector with level-dependent and flat attenuation characteristics. Noise Cont. Eng. J., 34:97-105, 1990.

M.A. Bates, M. Loeb, R.P. Smith, and J.L. Fletcher. Attempts to condition the acoustic reflex. J. Aud. Res., 10:132-135, 1970.

E.H. Berger and P. Hamery. Empirical evaluation using impulse noise of level-dependency for passive earplug designs. In Acoustics `08, Paris, July 2008. Acoustical Society of America.

P.C. Chan and K.C. Ho. Issues of the AHAAH model for impulse noise hazard assessment. Technical Report J0189-05-275, Contract No. F41624-01-C-7002, L3Com, 2005.

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

J. Cui. QIC program and model selection in GEE analyses. The Stata Journal, 7:209-220, 2007.

Direction Technique de Armements Terrestres. Recommendations on evaluating the possible harmful effects of noise on hearing. Technical Report AT-83/27/28, Establissement Technique de Bourges, 1983.

J.J. Earshen. Sound Measurement: Instrumentation and Noise Descriptors, pages 52{55. AIHA Press, Fairfax, VA, 2003.

G.M. Fitzmaurice, N.M. Laird, and J.H. Ware. Applied Longitudinal Analysis, page 385. Wiley, New York, 2004.

J.L. Fletcher and A.J. Riopelle. Protective effect of the acoustic reflex for impulsive noises. J. Acoust. Soc. Am., 32:401-404, 1960.

J.R. Franks, W.J. Murphy, J.L. Johnson, and D.A. Harris. Four earplugs in search of a rating system. Ear Hear., 21:218-226, 2000.

J.W. Hardin and J.M. Hilbe. Generalized Estimating Equations, page 168. Chapman and Hall, Boca Raton, 2003.

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.

N.J. Horton, J. D. Bebchuk, C.L. Jones, S.R. Lipsitz, P.J. Catalano, G.E.P. Zahner, and G.M. Fitzmaurice. Goodness-of-fit for GEE: An example with mental health service utilization. Statistics in Medicine, 18:213-222, 1999.

D.W. Hosmer, T. Hosmer, S. Le Cessie, and S. Lemeshow. A comparison of goodness-of-fit tests for the logistic regression model. Statistics in Medicine, 16:965-980, 1997.

D.W. Hosmer and S. Lemeshow. Applied Logistic Regression, page 162. Wiley, New York, 2000.

D.L. Johnson. Blast overpressure studies with animals and men: A walk-up study. Technical Report USAARL 94-2, U.S. Army Aeromedical Research Laboratories, 1994.

D.L. Johnson. Blast overpressure studies: Final task report for U.S. Army Medical Research and Materiel Command. Technical Report USAARL CR DAMD17-93-C-3101, U.S. Army Aeromedical Research Laboratories, 1997.

J. Kim. Effect of A-duration on system response. Personal communication, April 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.

K.Y. Liang and S.L. Zeger. Longitudinal data analysis using generalized linear models. Biometrika, 73:13-22, 1986.

L. Marshall, J.F. Brandt, and L.E. Marston. Anticipatory middle-ear reflex activity from noisy toys. J. Speech Hear. Dis., 40:320-326, 1975.

MIL-STD-1474D. Department of Defense Design Criteria Standard, Noise Limits. Department of Defense, Washington D.C., 1997.

MSHA. Health standards for occupational noise exposures: Final rule. Mine Safety and Health Administration. 20 CFR part 62, 64 Fed Reg: 49548-634, 49636-7, 1999.

W.J. Murphy, D.C. Byrne, and J.R. Franks. Firearms and hearing protection. Hear. Rev., pages 36-38, 2007.

W.J. Murphy, J.R. Franks, E.H. Berger, A. Behar, J.G. Casali, C. Dixon Ernst, E.F. Krieg, B.T. Mozo, D.W. Ohlin, J.D. Royster, L.H. Royster, S.D. Simon, and C. Stephenson. Development of a new standard laboratory protocol for estimation of the ¯eld attenuation of hearing protection devices: Sample size necessary to provide acceptable reproducibility. J. Acoust. Soc. Am., 115:311-323, 2004.

W.J. Murphy and R. L. Tubbs. Assessment of noise exposure for indoor and outdoor ¯ring ranges. J. Occ. Env. Hyg., 4:688-697, 2007.

W.J. Murphy, D.C. Byrne, D. Gauger, W.A. Ahroon, E.H. Berger, S>N.Y Gerges, R. McKinley, B. Witt, E.F. Krieg, Results of the National Institute for Occupational Safety and Health – U.S. Environmental Protection Agency Interlaboratory Comparison of American National Standards Institute S12.6-1997 Methods A and B. J. Acoust. Soc. Am. 125:3262-3277, 2009.

R. Newson. Parameters behind non-parametric statistics: Kendalls τ_{α}, somer's d , and median differences. The Stata Journal, 1:1-20, 2001.

NIOSH. Criteria for a recommended standard: Occupational noise exposure; Revised Criteria. Report no. (NIOSH) 98-126. DHHS, Centers for Disease Control, National Institute for Occupational Safety and Health, 1998.

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

OSHA. CPL 2-2.35A-29 CFR 1910.95(b)(1) Guidelines for noise enforcement: Appendix A, U.S. Department of Labor, Occupational Safety and Health Administration, December 19, 1983.

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

J.H. Patterson and W.A. Ahroon. Evaluation of an auditory hazard model using data from human volunteer studies. Technical Report USAARL 2005-01, U.S. Army Aeromedical Research Laboratories, 2005.

J.H. Patterson and D.L. Johnson. Temporary threshold shifts produced by high intensity free-field impulse noise in human wearing hearing protection. Technical Report USAARL 94-46, U.S. Army Aeromedical Research Laboratories, 1994.

M.S. Pepe, G. Longton, and H. Janes. Estimation and comparison of receiver operating characteristic curves. The Stata Journal, 9:1-16, 2009.

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. 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. Stata Press, College Station TX, 2005.

A.E. Raftery. Bayesian model selection in social research, volume 26, pages 111-163. Oxford, Blackwell, 1996.

P. Rasmussen, G. Flamme, M. Stewart, D. Meinke, and J Langford. Measuring recreational firearm noise. Sound and Vib., pages 2-6, August 2009.

G.F. Smoorenburg. Damage risk criteria for impulse noise. In R.P. Hamernik, D. Henderson, and R. Salvi, editors, New Perspectives on Noise-Induced Hearing Loss, pages 471-490, New York, 1982. Raven Press.

W.D. Ward ed. Proposed damage-risk criteria for impulse noise (gunfire) (U). Washington, DC: National Academy of Sciences, National Research Council Committee on Hearing, Bioacosutics, and Biomechanics, (1968).

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

J. Zera and R. Mlynski. Attenuation of high-level impulses by earmuffs. J. Acoust. Soc. Am., 122:2082-2096, 2007.

Definitions of Terms

AHAAH	Auditory Hazard Assessment Algorithm
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
β_0, β_1	parameters in the linear predictor for GEE and GLMM
$\hat{\beta}_0$, $\hat{\beta}_1$	Estimates of parameters β_0, and β_1
GEE	Generalized estimating equation (these are population-average or marginal models)
GLMM	Generalized linear mixed model (these are subject-specific models)
L_{95}	Value of a hazard index for which the mean failure rate is equal to 0.05
$L_{(95,95)}$	Level of the hazard index for which the upper confidence limit of predicted failure rate is equal to 0.05
L_{Aeq8hr}	Eight-hour equivalent of A-weighted sound exposure level
L_m	Effective exposure level for MIL-STD-1474D
L_p	Peak sound-pressure level
$P_{\max}$	Peak pressure of the free-field incident wave
P_{ref}	Reference value of 20 μPa
QIC_u	Quasi-likelihood under the independence model criterion
T_b	*B* duration in milliseconds
TTS	Temporary threshold shift

Table 1: 1-Meter BOP Study Injury Data using Modified Earmuff

Level	6 Blasts	12 Blasts	25 Blasts	50 Blasts	100 Blasts
7	54,0,2,0				
6	59,1,0,0	53,2,2,0	49,4,2,0	42,3,0,1	24,2,1,0
5	60,0,1,0	1,2,0,0	4,0,0,0	3,0,3,0	11,2,4,0
4	62,0,0,0	1,0,1,0	2,1,0,0	1,0,0,0	2,1,0,0
3	62,0,0,0	1,0,0,0	0,0,0,0	0,1,0,0	1,0,0,0
2	63,1,0,0	0,0,1,0	0,1,0,0	1,0,0,0	0,1,0,0
1	65,0,0,0	1,0,0,0	0,0,0,0	1,0,0,0	0,0,0,0

First number = Pass
Second Number = Conditional Failure
Third number= Audiometric failure
Fourth number= Incomplete testing

Table 2: 3-Meter BOP Study Injury Data using Modified Earmuff

Level	6 Blasts	12 Blasts	25 Blasts	50 Blasts	100 Blasts
7	40,2,0,4				
6	59,0,2,0	57,1,1,0	52,2,2,1	48,0,2,0	25,0,2,1
5	61,3,1,0	6,1,0,0	6,0,0,0	4,1,1,0	19,0,0,0
4	65,3,0,0	8,1,0,0	2,0,0,0	1,0,0,0	0,1,0,0
3	67,1,0,0	4,0,0,0	0,1,0,0	0,0,0,0	0,0,0,0
2	68,0,0,0	1,0,0,0	0,0,0,0	1,0,0,0	0,0,0,0
1	68,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0

First number = Pass
Second Number = Conditional Failure
Third number= Audiometric failure
Fourth number= Incomplete testing

Table 3: 5-meter BOP Study Injury Data using Modified Earmuff

Level	6 Blasts	12 Blasts	25 Blasts	50 Blasts	100 Blasts
7	54,0,0,0				
6	56,1,0,0	55,1,0,0	56,0,0,0	55,1,0,0	52,0,2,2
5	56,1,0,0	1,0,0,0	1,0,0,0	0,0,0,0	1,0,0,0
4	58,0,0,0	1,0,0,0	0,1,0,0	0,0,0,0	0,0,0,0
3	59,0,0,0	0,0,0,0	0,0,0,0	0,1,0,0	0,0,0,0
2	59,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0	0,1,0,0
1	59,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0

First number = Pass
Second Number = Conditional Failure
Third number= Audiometric failure
Fourth number= Incomplete testing

Table 4: 5-meter BOP Study Injury Data using Unmodified Earmuff

Level	6 Blasts	12 Blasts	25 Blasts	50 Blasts	100 Blasts
7	50,0,0,0				
6	58,1,0,0	56,0,0,0	53,0,0,0	45,0,0,0	39,0,0,0
5	59,0,0,0	1,0,0,0	1,0,0,0	1,0,0,0	0,0,0,0
4	62,0,0,0	2,0,0,0	2,0,0,0	2,0,0,0	2,0,0,0
3	61,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0
2	62,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0
1	62,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0	0,0,0,0

First number = Pass
Second Number = Conditional Failure
Third number= Audiometric failure
Fourth number= Incomplete testing

Table 5: Analyzed waveform data for single blast of 1-meter exposures with modified earmuff

	Impulse measured in free-field			Impulse measured under hearing protector			
Exposure Level	**N**	**MIL-STD 1474 (dB)**	**L_{Aeq8hr} (dBA)**	**N**	**L_{Aeq8hr} (dBA)**	**AHAAH Warned (ARU)**	**AHAAH Unwarned (ARU)**
7	NA	NA	NA	8	94.0	17.0	97.0
6	4	188.9	110.1	14	93.1	36.9	142.6
5	10	182.7	108.3	12	91.7	36.3	158.3
4	4	179.1	104.8	4	91.9	70.3	293.8
3	8	175.8	101.6	11	89.9	55.7	241.9
2	12	174.3	99.2	17	89.0	46.4	175.7
1	6	171.8	96.2	9	86.0	34.7	165.1

Table 6: Analyzed waveform data for single blast of 3-meter exposures with modified earmuff

	Impulse measured in free-field			Impulse measured under hearing protector			
Exposure Level	**N**	**MIL-STD 1474 (dB)**	**L_{Aeq8hr} (dBA)**	**N**	**L_{Aeq8hr} (dBA)**	**AHAAH Warned (ARU)**	**AHAAH Unwarned (ARU)**
7	6	188.4	110.6	10	92.7	33.9	130.2
6	23	184.0	107.7	32	91.2	33.5	171.2
5	25	180.6	104.6	16	91.2	24.1	144.0
4	19	178.8	103.1	23	88.9	33.6	179.6
3	11	174.7	99.6	18	86.7	37.1	185.9
2	12	172.2	96.4	15	83.7	29.5	156.3
1	NA	NA	NA	NA	NA	NA	NA

Table 7: Analyzed waveform data for single blast of 5-meter exposures with modified earmuff

	Impulse measured in free-field			Impulse measured under hearing protector			
Exposure Level	**N**	**MIL-STD 1474 (dB)**	**L_{Aeq8hr} (dBA)**	**N**	**L_{Aeq8hr} (dBA)**	**AHAAH Warned (ARU)**	**AHAAH Unwarned (ARU)**
7	2	183.5	107.5	3	96.8	37.7	118.4
6	2	180.7	105.4	3	94.6	36.7	111.8
5	4	177.8	102.8	6	91.0	28.0	92.4
4	3	173.8	99.1	6	90.0	29.2	88.7
3	4	171.4	96.7	9	87.0	24.0	97.5
2	4	168.7	94.5	9	84.8	18.9	80.2
1	6	165.5	91.2	8	83.3	18.6	111.2

Table 8: Analyzed waveform data for single blast of 5-meter exposures with unmodified earmuff

	Impulse measured in free-field			Impulse measured under hearing protector			
Exposure Level	N	MIL-STD 1474 (dB)	L_{Aeq8hr} (dBA)	N	L_{Aeq8hr} (dBA)	AHAAH Warned (ARU)	AHAAH Unwarned (ARU)
7	8	183.5	107.5	12	92.3	27.5	109.7
6	1	180.7	106.8	12	88.3	13.2	75.1
5	10	177.8	102.6	18	85.2	9.4	58.1
4	12	173.8	99.9	18	80.7	4.7	32.8
3	12	171.4	97.8	18	77.2	3.1	23.4
2	12	168.7	94.7	18	74.2	2.0	17.3
1	12	165.5	92.1	18	72.6	1.3	13.7

Table 9: L_{95} and $L_{(95,95)}$ for different indices and outcome variables

	Audiometric Failure		Audiometric & Conditional Failure	
Independent Variable	L_{95}	$L_{(95,95)}$	L_{95}	$L_{(95,95)}$
MIL-STD-1474D (dB)	196.3	193.5	188.7	185.1
L_{Aeq8hr} Free-field (dBA)	126.0	123.6	118.8	115.4
L_{Aeq8hr} Protected (dBA)	113.1	109.5	104.6	101.4
Unwarned AHAAH (ARU)	13064.3	10108.7	5710.5	3053.4
Warned AHAAH (ARU)	3247.2	2479.8	1377.6	718.7

Table 10: Information criteria for assessing quality of fit of different metrics

	Audiometric Failure			Audiometric & Conditional Failure		
Independent Variable	AIC for GLMM	BIC for GLMM	QIC (ind) for GEE	AIC for GLMM	BIC for GLMM	QIC (ind) for GEE
MIL-STD-1474D (dB)	271.45	288.03	282.01	549.24	565.83	636.25
L_{Aeq8hr} Free-field (dBA)	258.63	275.21	273.73	534.06	550.65	630.26
L_{Aeq8hr} Protected (dBA)	294.44	311.11	305.80	562.85	579.52	650.05
Unwarned AHAAH (ARU)	292.83	309.51	304.51	572.76	589.43	650.85
Warned AHAAH (ARU)	295.53	312.20	307.13	574.14	590.81	653.99

Table 11: Results of Goodness-of-fit test for the General Estimating Equation using three categories of risk with failure defined as an audiometric failure. p > 0.05 indicates a good fit.

Hazard Index	p-value
LAeq_8hr Free-field	0.9182
LAeq_8hr protected	0.5496
MIL-STD-1474D	0.2818
Unwarned AHAAH	0.0340
Warned AHAAH	0.0103

Table 12: Results of Goodness-of-fit test for the General Estimating Equation using three categories of risk with failure defined as either an audiometric or conditional failure. p > 0.05 indicates a good fit.

Hazard Index	p-value
MIL-STD-1474D	0.5253
LAeq_8hr protected	0.3224
LAeq_8hr unprotected	0.3220
Warned AHAAH	0.0011
Unwarned AHAAH	0.0000

Table 13: Results of GEE Goodness-of-fit tests – using four categories of risk with failure defined as either an audiometric or conditional failure p > 0.05 indicates a good fit.

Hazard Index	p-value
LAeq_8hr protected	0.8060
LAeq_8hr unprotected	0.1727
MIL-STD-1474D	0.0431
Warned AHAAH	0.0317
Unwarned AHAAH	0.0010

Table 14: Raftery's classification of differences in goodness-of-fit using the Bayesian Information Criteria.

Absolute difference in BIC	Evidence
0-2	Weak
2-6	Positive
6-10	Strong
> 10	Very Strong

Table 15: Area under ROC curves (AUC or Harrell's C) for different damage-risk criteria for just audiometric failures for the pre-specified cells. Note that larger AUC implies better discrimination.

Hazard Index	Harrell's C (AUC)	Jackknife Std. Error	95% Conf. Interval
LAeq unprotected	.8213	.0375	(.7478, .8949)
LAeq protected	.7685	.0425	(.6853, .8517)
MIL-STD-1474D	.7974	.0303	(.7380, .8568)
Unwarned AHAAH	.7923	.0474	(.6993, .8852)
Warned AHAAH	.7546	.0568	(.6433, .8659)

Table 16: Differences in AUC for different damage-risk criteria for just audiometric failures.

Difference in Hazard Indices	Difference in Harrell's C	Std. Err.	z	p-value (P>|z|)	95% Conf. Interval
LAeq unprotected - LAeq protected	.0528	.0113	4.68	0.000	(.0307, .0749)
LAeq unprotected - MIL-STD-1474D	.0239	.0229	1.04	0.297	(-.0210, .0688)
LAeq unprotected – Unwarned AHAAH	.0291	.0134	2.17	0.030	(.0028, .0553)
LAeq unprotected – Warned AHAAH	.0667	.0211	3.17	0.002	(.0254, .1080)
LAeq protected - MIL-STD-1474D	-.0289	.0318	-0.91	0.364	(-.0912, .0334)
LAeq protected – Unwarned AHAAH	-.0238	.0145	-1.63	0.102	(-.0523, .0047)
LAeq protected – Warned AHAAH	.0139	.0193	0.72	0.471	(-.0239, .0518)
MIL-STD-1474D –Unwarned AHAAH	.0051	.0328	0.16	0.876	(-.0592, .0695)
MIL-STD-1474D – Warned AHAAH	.0428	.0381	1.12	0.262	(-.0319, .1176)
Unwarned AHAAH – Warned AHAAH	.0377	.0164	2.30	0.021	(.0056, .0698)

Table 17**:** Area under ROC curves (AUC or Harrell's C) for different damage-risk criteria for just audiometric failures for the *ad hoc* observations.

Hazard Index	**Harrell's C (AUC)**	**Jackknife Std. Error**	**95% Conf. Interval**
LAeq unprotected	.7000	.1490	(.4080, .9920)
LAeq protected	.6759	.1138	(.4528, .8990)
MIL-STD-1474D	.6976	.1511	(.4014, .9938)
Unwarned AHAAH	.6157	.1049	(.4101, .8213)
Warned AHAAH	.6807	.0840	(.5162, .8453)

Table 18**:** Area under ROC curves (AUC or Harrell's C) for different damage-risk criteria for audiometric and conditional failures combined for the pre-specified cells. Note that larger AUC implies better discrimination.

Hazard Index	**Harrell's C (AUC)**	**Jackknife Std. Error**	**95% Conf. Interval**
LAeq unprotected	.7415	.0351	(.6727, .8104)
LAeq protected	.6944	.0350	(.6258, .7630)
MIL-STD-1474D	.7473	.0327	(.6832, .8113)
Unwarned AHAAH	.6968	.0419	(.6146, .7790)
Warned AHAAH	.6751	.0432	(.5903, .7598)

Table 19: Differences in AUC for different damage-risk criteria for audiometric and conditional failures combined.

Difference in Hazard Indices	Difference in Harrell's C	Std. Err.	z	p-value (P>|z|)	95% Conf. Interval
LAeq unprotected - LAeq protected	.0471	.0107	4.41	0.000	(.0262, .0681)
LAeq unprotected - MIL-STD-1474D	-.0057	.0138	-0.42	0.677	(-.0327, .0212)
LAeq unprotected – Unwarned AHAAH	.0448	.0201	2.22	0.026	(.0053, .0842)
LAeq unprotected – Warned AHAAH	.0665	.0179	3.71	0.000	(.0314, .1016)
LAeq protected - MIL-STD-1474D	-.0529	.0207	-2.55	0.011	(-.0935, -.0122)
LAeq protected – Unwarned AHAAH	-.0024	.0244	-0.10	0.923	(-.0502, .0455)
LAeq protected – Warned AHAAH	.0193	.0193	1.00	0.316	(-.0185, .0572)
MIL-STD-1474D – Unwarned AHAAH	.0505	.0288	1.75	0.079	(-.0059, .1069)
MIL-STD-1474D – Warned AHAAH	.0722	.0266	2.71	0.007	(.0200, .1244)
Unwarned AHAAH – Warned AHAAH	.0217	.0150	1.45	0.148	(-.0077, .0511)

Table 20**:** Area under ROC curves (AUC or Harrell's C) for different damage-risk criteria for audiometric and conditional failures combined for the *ad hoc* cells.

Hazard Index	Harrell's C (AUC)	Jackknife Std. Error	95% Conf. Interval
LAeq unprotected	.5508	.0821	(.3898, .7118)
LAeq protected	.6222	.0729	(.4792, .7651)
MIL-STD-1474D	.5035	.0834	(.3400, .6670)
Unwarned AHAAH	.6208	.0723	(.4791, .7626)
Warned AHAAH	.6708	.0657	(.5419, .7996)

VIII. Figures

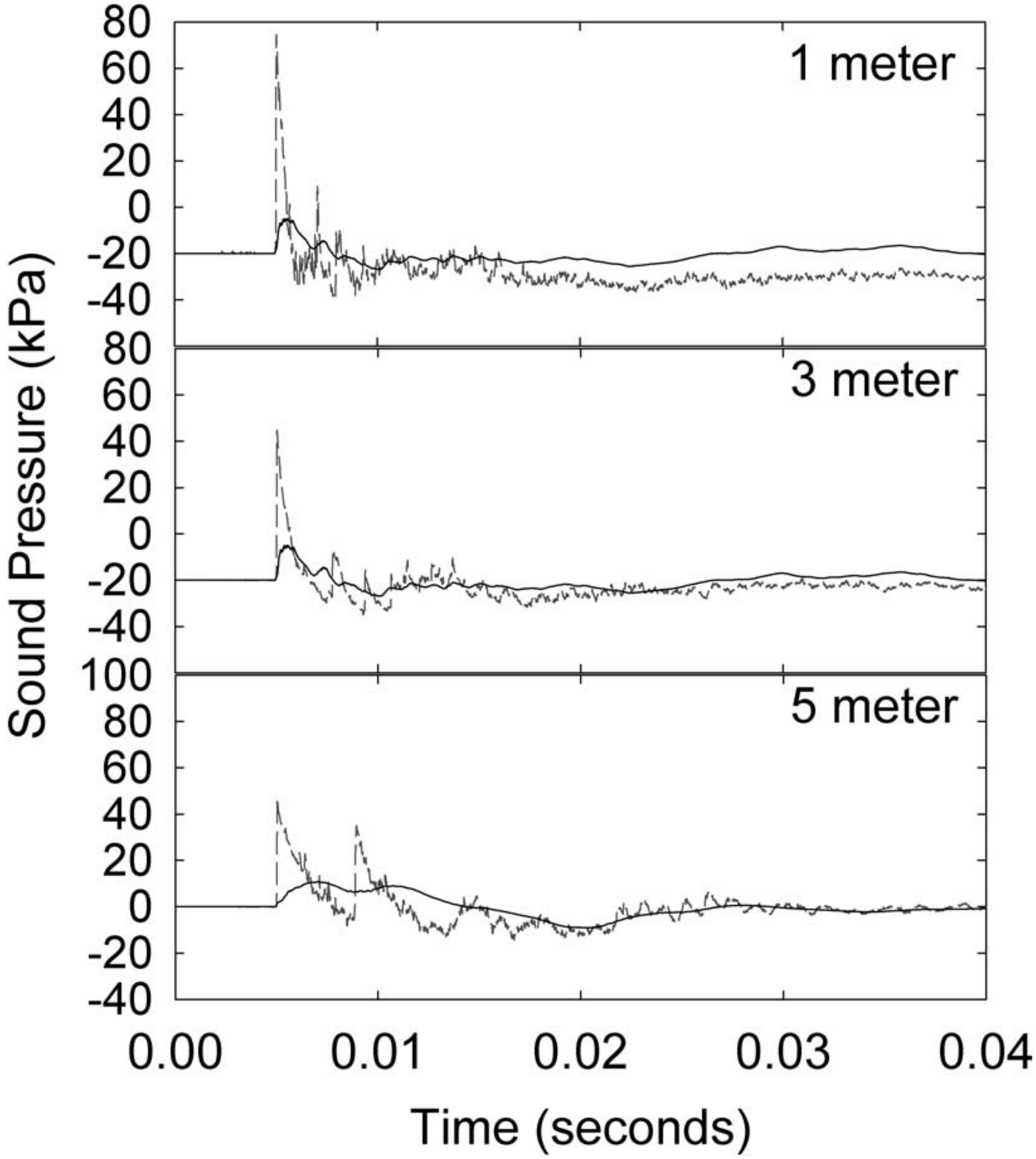

Figure 1: Example waveforms from the 1, 3 and 5-meter exposures for the free-field (dashed line) and protected (solid line) conditions.

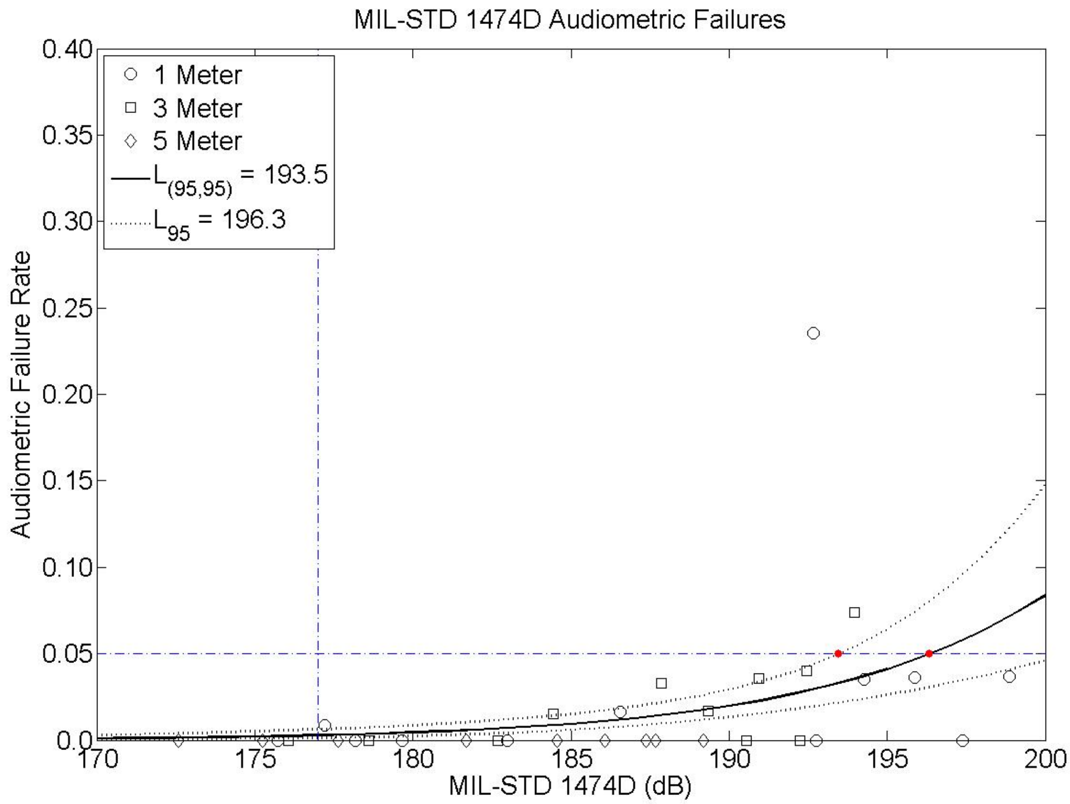

Figure 2: Evaluations of 1, 3 and 5-meter exposures using the MIL-STD-1474D metric using the audiometric and conditional failures. The logistic regression curve and 95% confidence intervals are shown as the solid and dotted lines. The $L_{(95,95)}$ and L_{95} points are plotted with a solid dot symbol. The 5% failure criterion and the 177 dB design limit are shown as horizontal and vertical dashed-dotted lines.

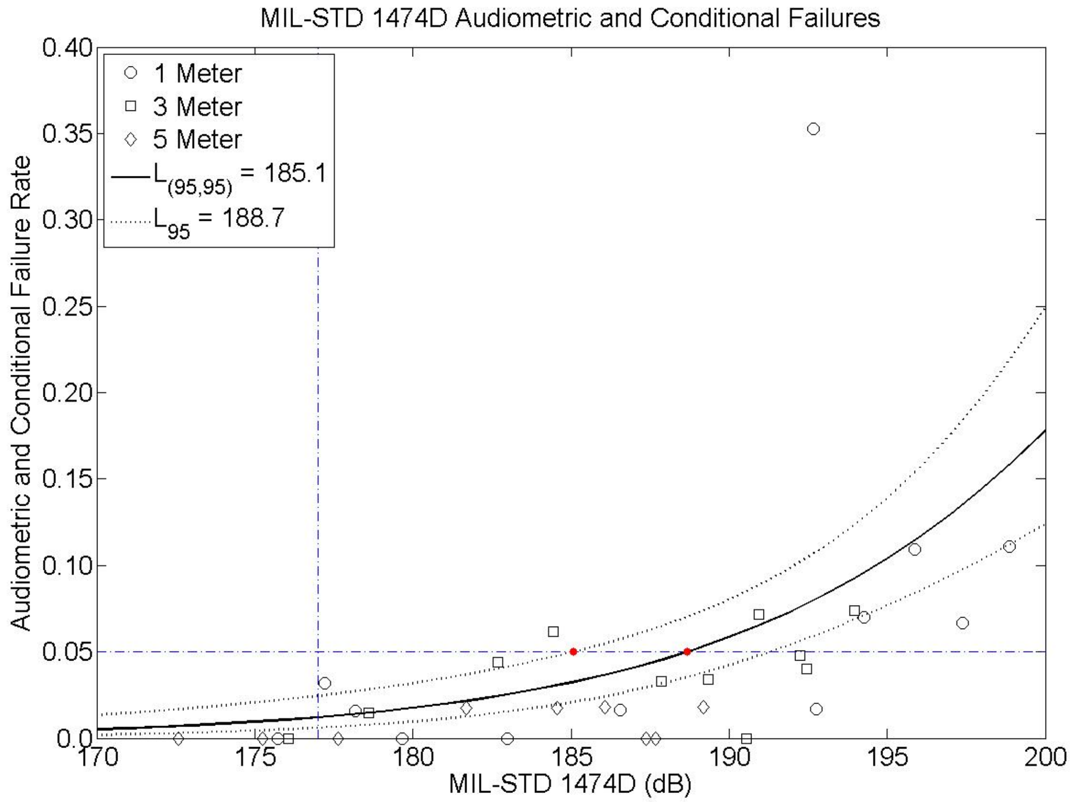

Figure 3: Evaluations of 1, 3 and 5-meter exposures using the MIL-STD-1474D metric using the audiometric and conditional failures. The logistic regression curve and 95% confidence intervals are shown as the solid and dotted lines. The $L_{(95,95)}$ and L_{95} points are plotted with a solid dot symbol. The 5% failure criterion and the 177 dB design limit are shown as horizontal and vertical dashed-dotted lines.

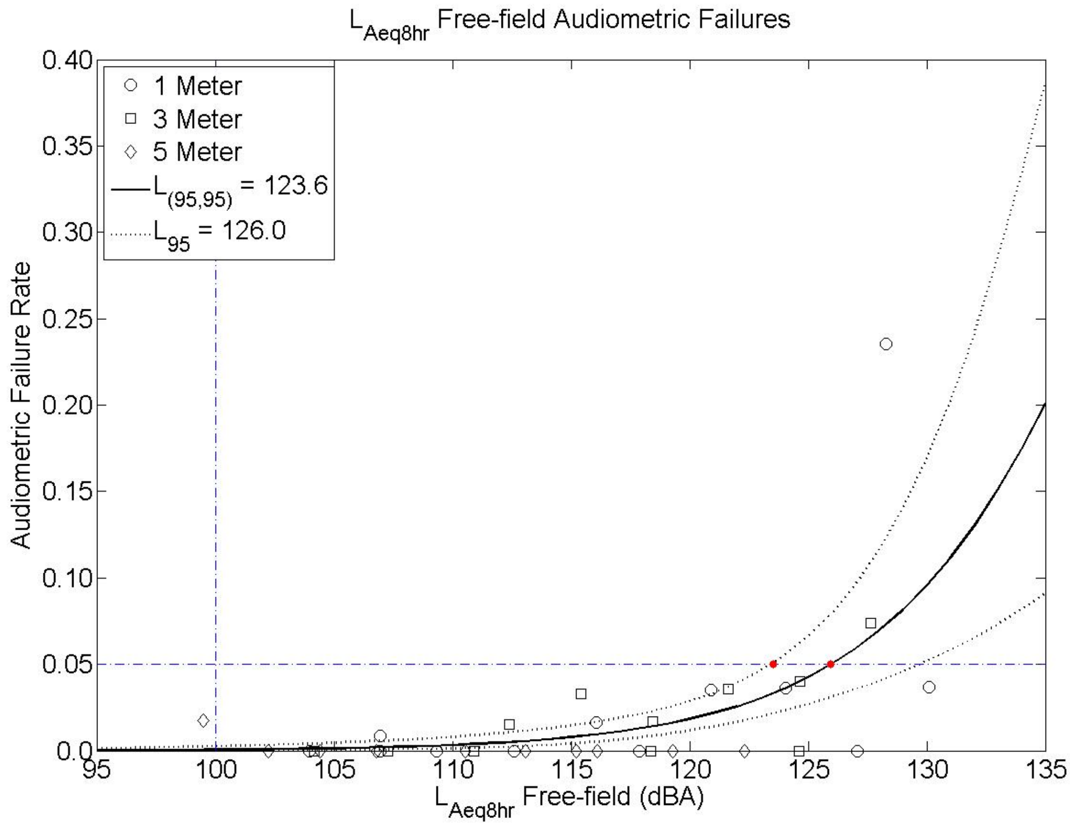

Figure 4: Evaluations of 1, 3 and 5-meter exposures using the protected 8-hour equivalent A-weighted Level, L_{Aeq8hr}, metric using the audiometric failures. The logistic regression curve and 95% confidence intervals are shown as the solid and dotted lines. The $L_{(95,95)}$ and L_{95} points are plotted with a solid dot symbol. The 5% failure criterion and the 100 dBA limit (85 dBA protected DRC+ 15 dBA hearing protection) are shown as horizontal and vertical dashed-dotted lines.

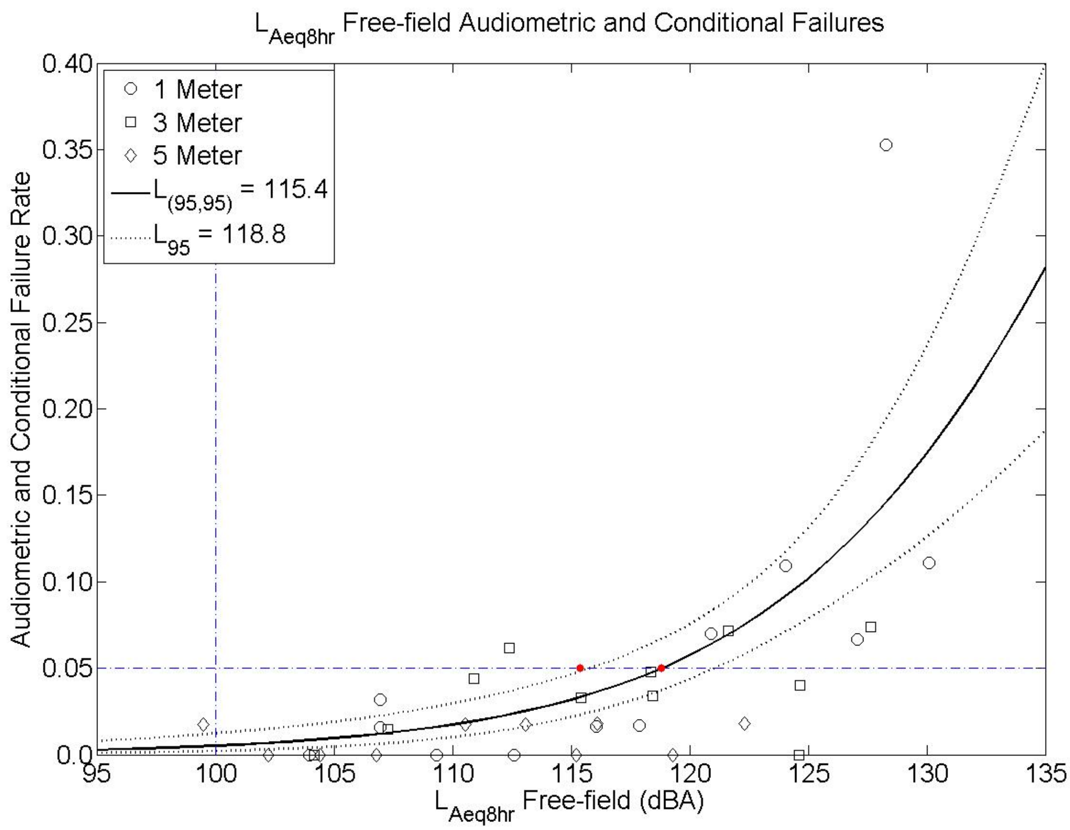

Figure 5: Evaluations of 1, 3 and 5-meter exposures using the protected 8-hour equivalent A-weighted Level, L_{Aeq8hr}, metric using the audiometric and conditional failures. The logistic regression curve and 95% confidence intervals are shown as the solid and dotted lines. The $L_{(95,95)}$ and L_{95} points are plotted with a solid dot symbol. The 5% failure criterion and the 100 dBA limit (85 dBA protected DRC+ 15 dBA hearing protection) are shown as horizontal and vertical dashed-dotted lines.

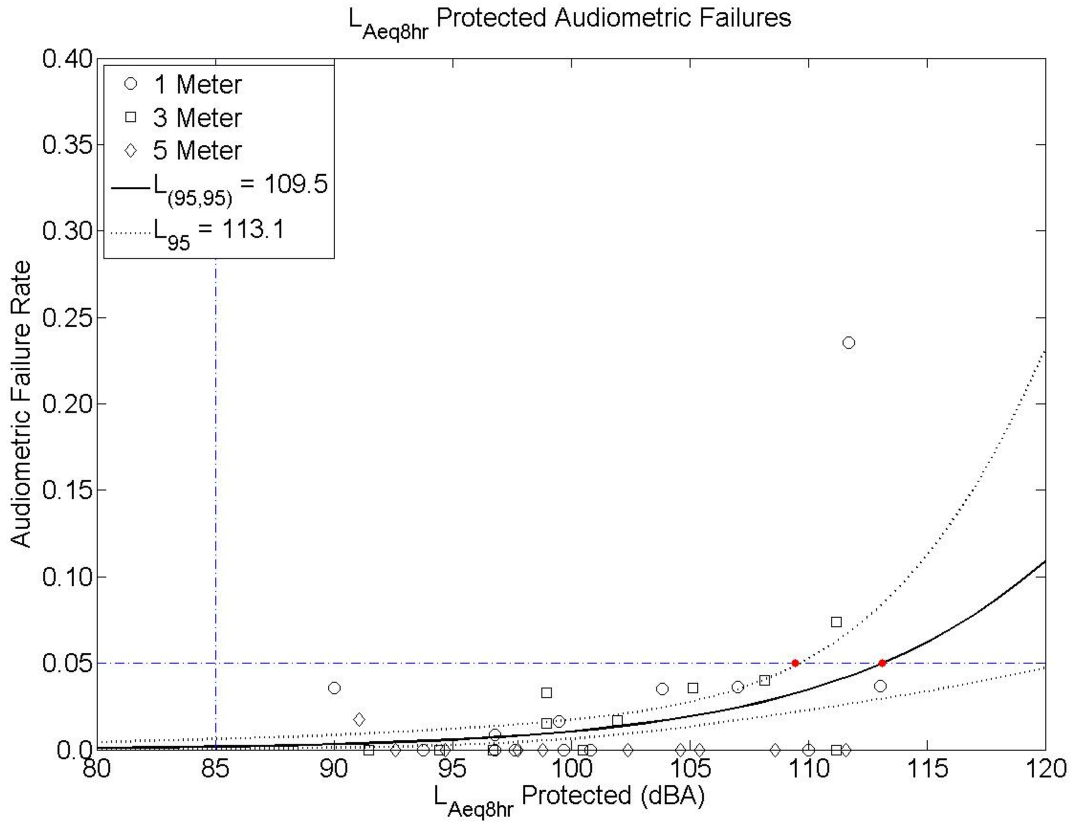

Figure 6: Evaluations of 1, 3 and 5-meter exposures using the protected 8-hour equivalent A-weighted Level, L_{Aeq8hr}, metric using the audiometric failures. The logistic regression curve and 95% confidence intervals are shown as the solid and dotted lines. The $L_{(95,95)}$ and L_{95} points are plotted with a solid dot symbol. The 5% failure criterion and the 85 dBA suggested damage-risk criterion are shown as horizontal and vertical dashed-dotted lines.

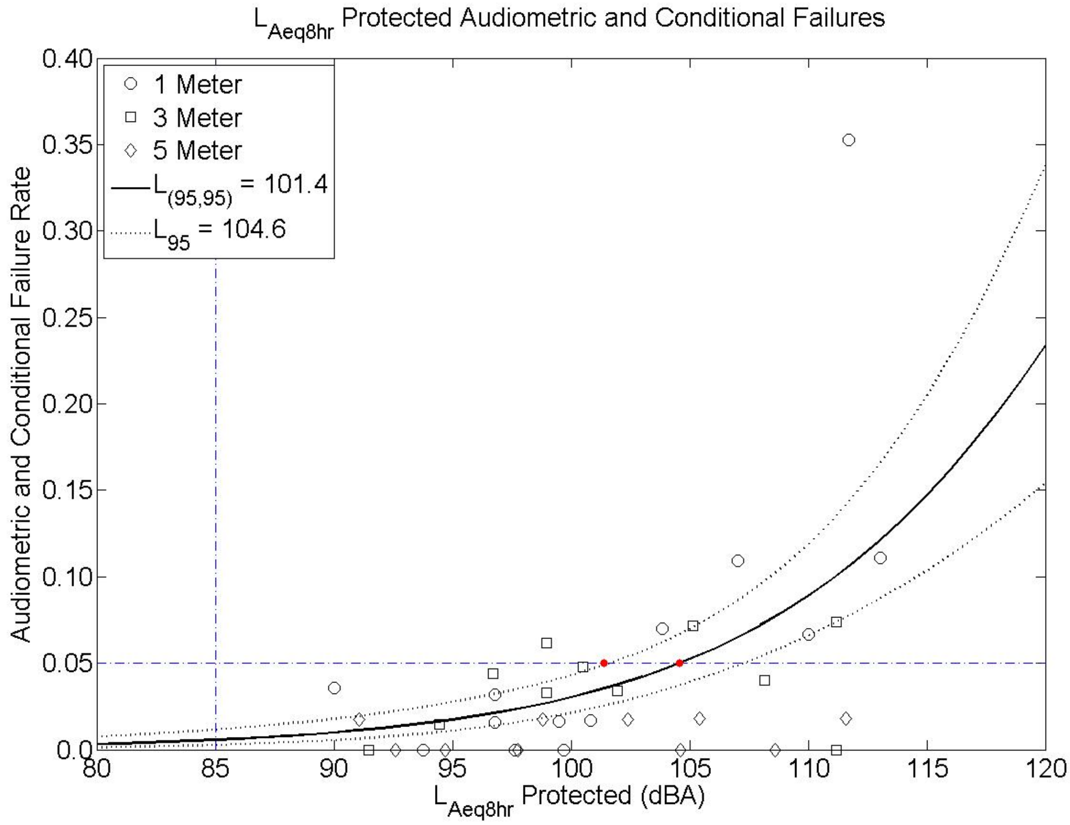

Figure 7: Evaluations of 1, 3 and 5-meter exposures using the protected 8-hour equivalent A-weighted Level, L_{Aeq8hr}, metric using the audiometric and conditional failures. The logistic regression curve and 95% confidence intervals are shown as the solid and dotted lines. The $L_{(95,95)}$ and L_{95} points are plotted with a solid dot symbol. The 5% failure criterion and the 85 dBA suggested damage-risk criterion are shown as horizontal and vertical dashed-dotted lines.

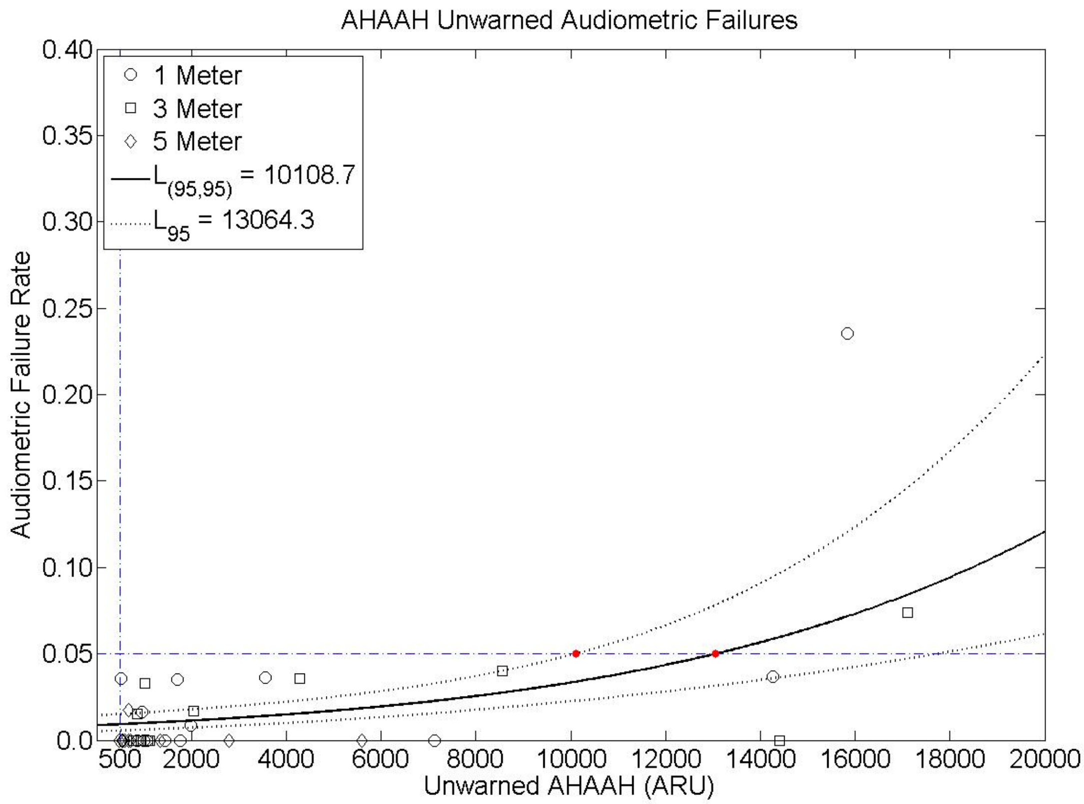

Figure 8: Evaluations of 1, 3 and 5-meter exposures using the unwarned AHAAH model using the audiometric failures. The logistic regression curve and 95% confidence intervals are shown as the solid and dotted lines. The $L_{(95,95)}$ and L_{95} points are plotted with a solid dot symbol. The 5% failure criterion and the 500 ARU suggested damage-risk criterion are shown as horizontal and vertical dashed-dotted lines.

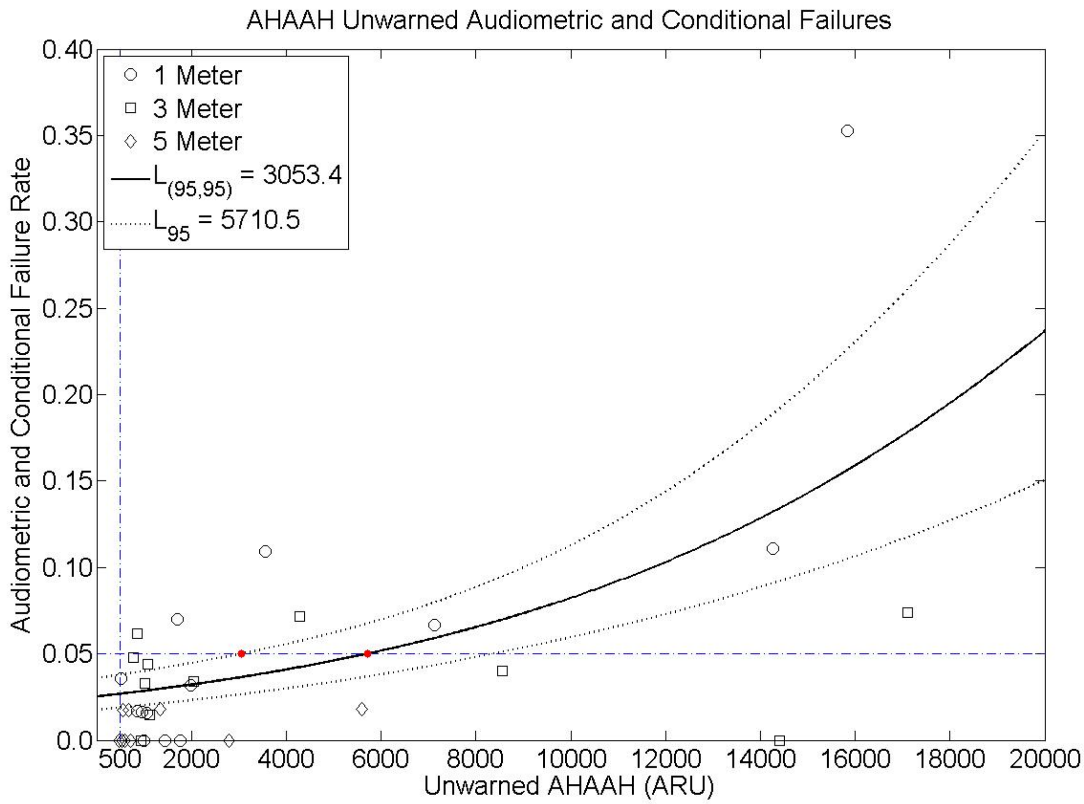

Figure 9: Evaluations of 1, 3 and 5-meter exposures using the unwarned AHAAH model using the audiometric and conditional failures. The logistic regression curve and 95% confidence intervals are shown as the solid and dotted lines. The $L_{(95,95)}$ and L_{95} points are plotted with a solid dot symbol. The 5% failure criterion and the 500 ARU suggested damage-risk criterion are shown as horizontal and vertical dashed-dotted lines.

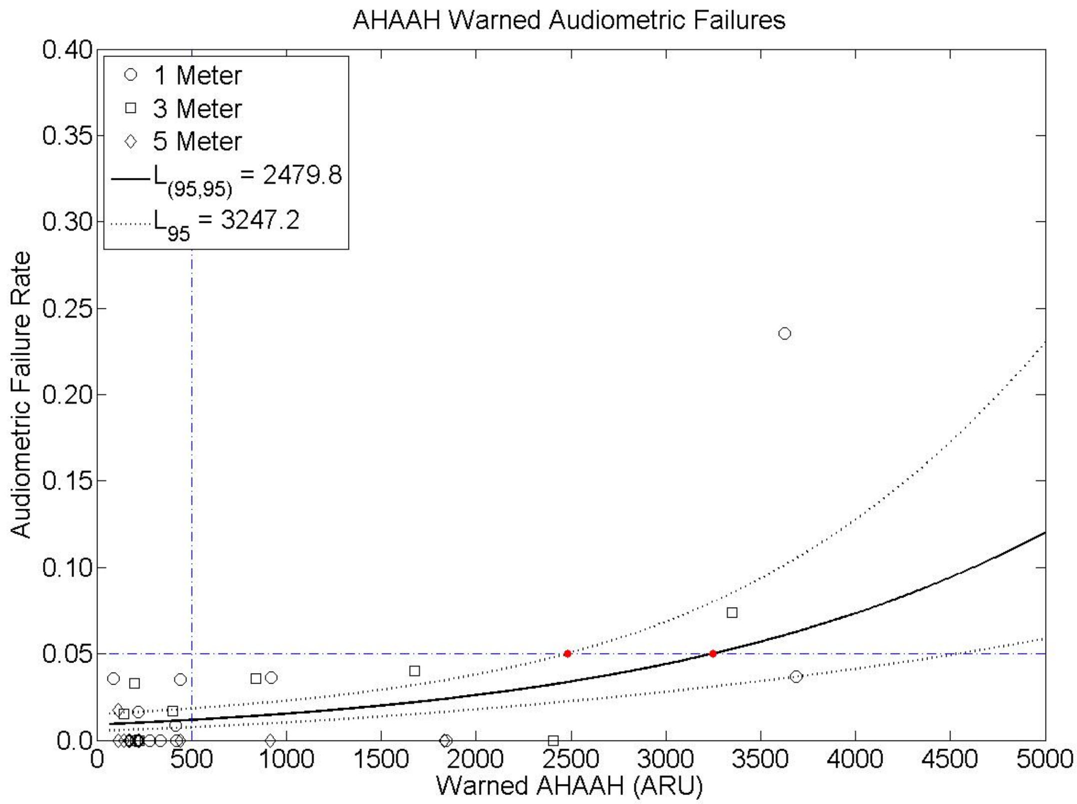

Figure 10: Evaluations of 1, 3 and 5-meter exposures using the warned AHAAH model using the audiometric failures. The logistic regression curve and 95% confidence intervals are shown as the solid and dotted lines. The $L_{(95,95)}$ and L_{95} points are plotted with a solid dot symbol. The 5% failure criterion and the 500 ARU suggested damage-risk criterion are shown as horizontal and vertical dashed-dotted lines.

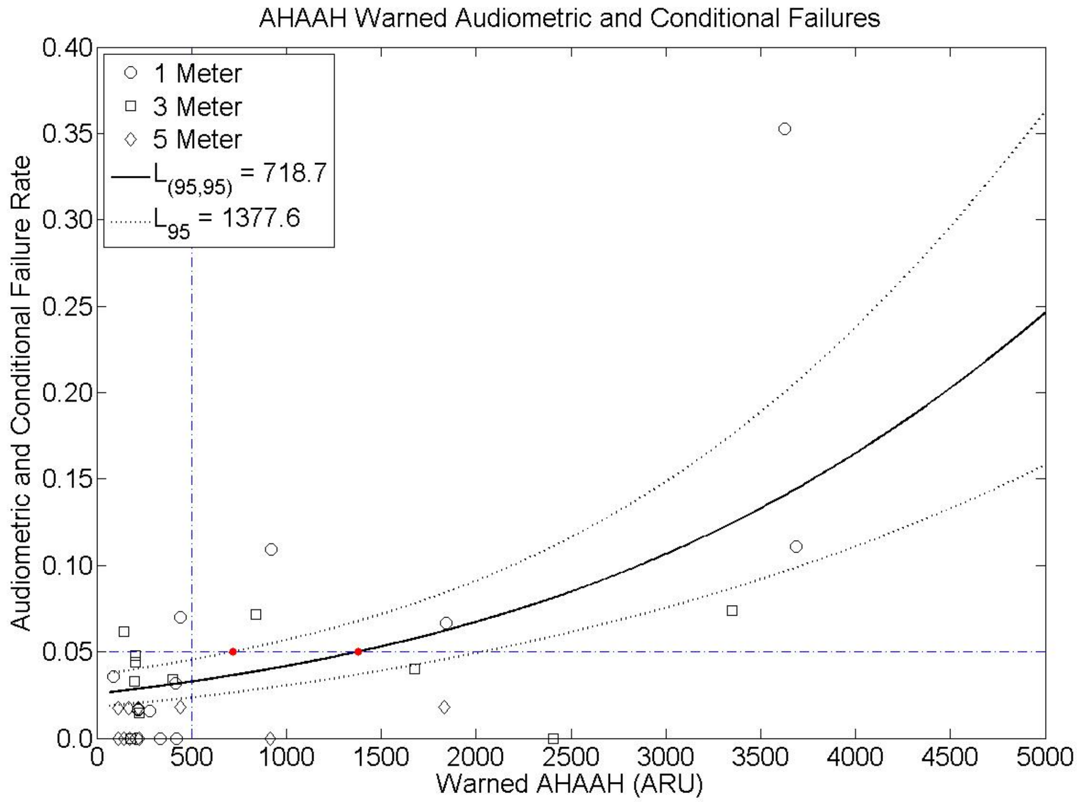

Figure 11: Evaluations of 1, 3 and 5-meter exposures using the warned AHAAH model using the audiometric and conditional failures. The logistic regression curve and 95% confidence intervals are shown as the solid and dotted lines. The $L_{(95,95)}$ and L_{95} points are plotted with a solid dot symbol. The 5% failure criterion and the 500 ARU suggested damage-risk criterion are shown as horizontal and vertical dashed-dotted lines.

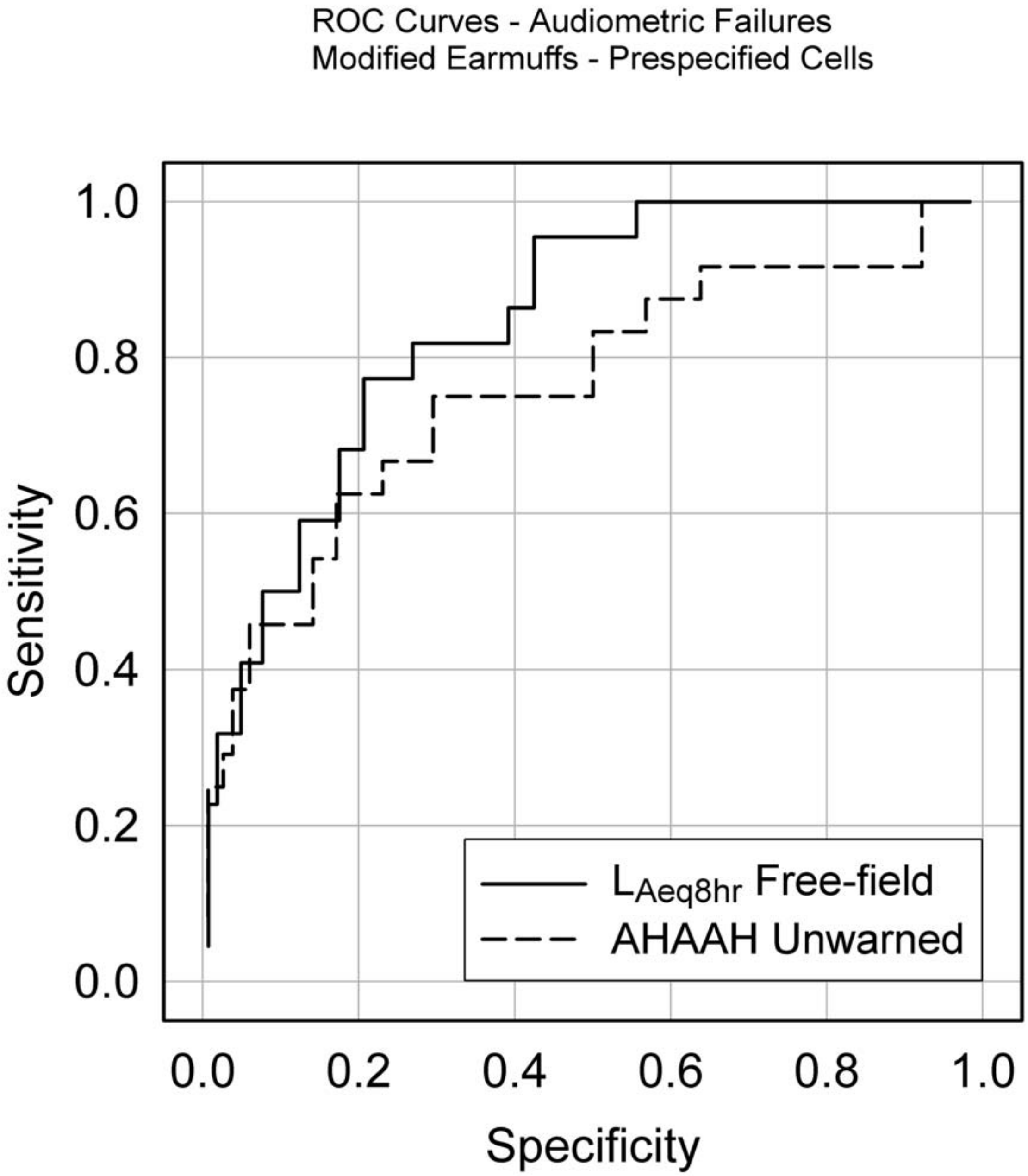

Figure 12: ROC curves for just audiometric failures for pre-specified cells for L_{Aeq8hr} and unwarned AHAAH damage-risk criteria.

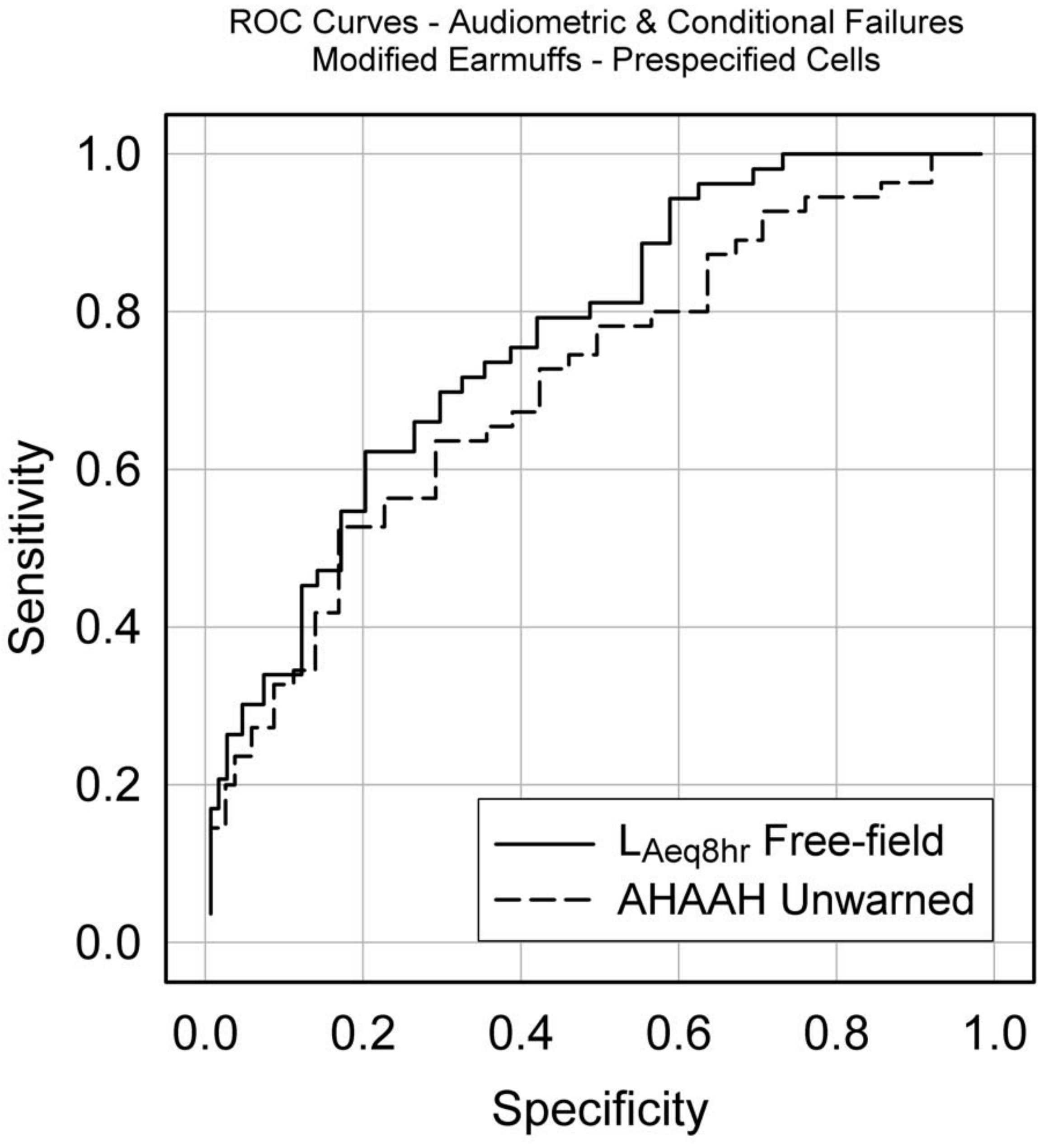

Figure 13: ROC curves for audiometric and conditional failures combined for pre-specified cells for L_{Aeq8hr} and unwarned AHAAH damage-risk criteria.

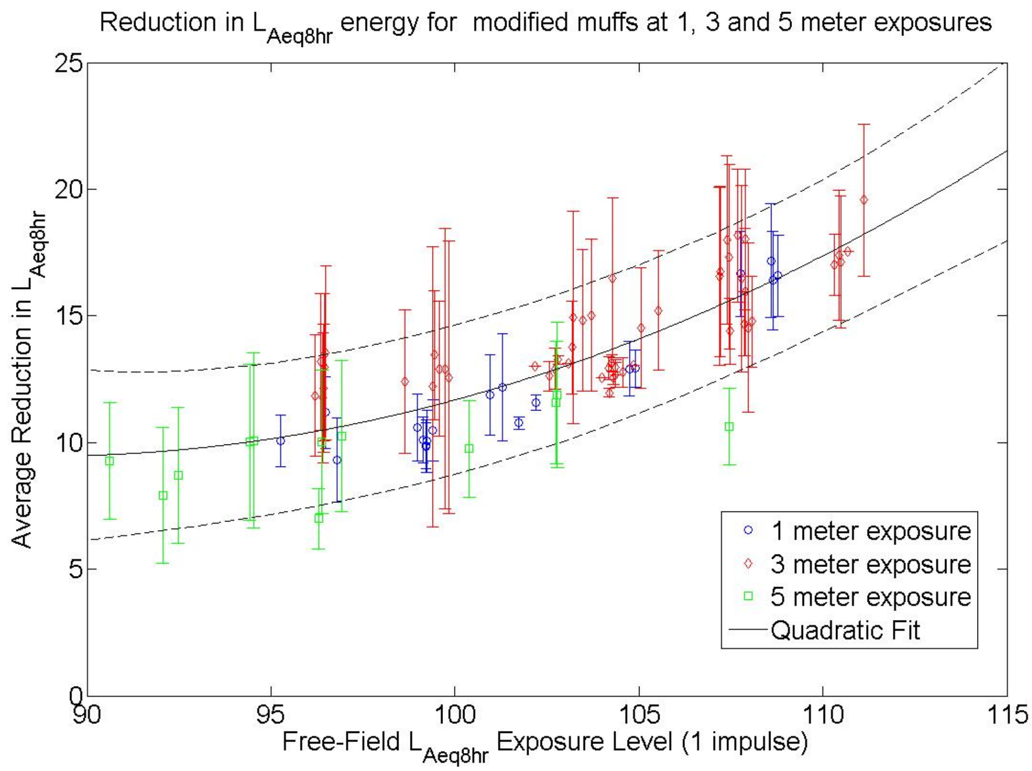

Figure 14 Comparison of the reduction in L_{Aeq8hr} from free-field to underneath the modified RACAL earmuff. The average and standard deviation of the peak impulse reduction measured between the free-field probe and the protected microphone were determined for each set of impulses. The different exposure distances are denoted by circles for 1-meter, diamonds for 3-meter and squares for 5-meter exposure distances. A quadratic fit is show to illustrate the trend of the peak reduction with increasing free-field peak level.