

Alternative Metrics for Noise Exposure Among Construction Workers

NOAH SEIXAS^{1*}, RICK NEITZEL¹, LIANNE SHEPPARD^{1,2} and BRYAN GOLDMAN^{1,2}

¹Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, WA 98195-7234, USA; ²Department of Biostatistics, University of Washington, Seattle, WA 98195, USA

Received 21 October 2004; in final form 16 February 2005; published online 29 March 2005

Although the exposure–response relationships for noise-induced hearing loss are relatively well established, there is not complete agreement on which metrics of noise exposure best represent risk of hearing damage. In particular, while L_{eq} , based on a 3 dB exchange rate (ER) is used by most agencies, US OSHA's standard is based on the L_{avg} , which uses a 5 dB ER. In addition, peak levels of exposure, which are commonly found in some industries, including construction, are believed to increase risk above that predicted by the L_{eq} . This paper presents an analysis of a large database of noise exposures among construction workers, comparing several noise metrics, and their application to a cohort of construction workers. Metrics examined were the L_{avg} , L_{eq} and L_{max} , expressing average levels of exposure across an exposure interval. Two novel metrics were derived from these monitored metrics, L_{eq}/L_{avg} and L_{max}/L_{eq} , as measures of exposure variability and 'peakiness', respectively. A total of 730 workshifts, including data on 361 492 min of exposure to workers in nine trades were examined. Correlations between average metrics (L_{eq} , L_{avg} and L_{max}) are generally very high, while the variability metrics are poorly correlated with either average levels, or with each other, indicating that they characterize different aspects of exposure. Alternative models for estimating exposure for the cohort were considered and the use of a task-within-trade specific mean level was adopted. The task-specific estimates of exposure using the various metrics will be applied to the cohort's work history to explore the importance of these alternative metrics in estimating risk of noise-induced damage.

Keywords: noise; exposure assessment; occupational noise exposure; exposure metrics; construction workers

INTRODUCTION

In order to study the health effects of environmental or occupational exposures using epidemiological techniques, an individual's exposures to the agent or agents of interest must be summarized into a coherent representation over time. The summary variable used for the exposure–response analysis, called the exposure metric, must capture the characteristics of exposure that are hypothesized or known to exert the damaging or toxic effect being studied—that is, the exposure metric must be consistent with the underlying model of the disease process (Seixas and Checkoway, 1995; Nieuwenhuijsen, 2003; Checkoway *et al.*, 2004). Selection of the 'correct'

exposure metric for a particular exposure-disease process is important because misspecification of the metric will introduce error into the analysis, thereby biasing any observed effect (Armstrong *et al.*, 1992). However, the degree to which the chosen metric is correlated with the true metric will determine how well it performs in the analysis.

Selection of an appropriate metric for noise exposure in relation to noise-induced hearing loss is particularly relevant. The standard metric for monitoring noise is the L_{eq} , a measure based on the equal energy hypothesis (EEH), which states that equal amounts of sound energy produce equal amounts of damage regardless of their distribution over time. The EEH implies a 3 dB exchange rate—the increase in dB allowed with a halving of exposure duration in order to keep the total energy equal. Thus the EEH contains the same assumptions as cumulative exposure about

*Author to whom correspondence should be addressed.
Tel: +1-206-685-7189; fax: +1-206-543-9616;
e-mail: nseixas@u.washington.edu

linearity of time and duration, and some authors argue that it is the most appropriate metric for integrating noise exposure of any type or duration (Atherley and Martin, 1971; Dancer *et al.*, 1998). Most national and international occupational and environmental health agencies use the L_{eq} (NIOSH, 1998). However, the US Occupational Safety and Health Administration (OSHA) standard specifies a 5 dB exchange rate and is termed the L_{avg} (OSHA, 1983). The L_{avg} is based on the premise that damage accrued in the ear during periods of high noise is partially repaired during intermittent low noise periods. This assumption results in lower estimates of damage from high noise periods than would be predicted using the EEH. Whether the L_{eq} or the L_{avg} is the correct metric for predicting noise induced hearing loss continues to be a subject of debate (NIOSH, 1998).

Furthermore, there is substantial evidence that high peak noise levels produce more damage than would be expected based on the EEH (Henderson *et al.*, 1994; Levine *et al.*, 1998). In animal studies, peaks over ~120–135 dBA induce mechanical damage to the hearing mechanism, while lower-level chronic noises produce toxic effects through metabolic alterations (Lataye and Campo, 1996). Typical continuous-type industrial noise levels are well under 120 dBA, but many processes including those in the construction industry do produce peaks in this range.

Although it may be important to include peak levels in an exposure metric, methods for summarizing peaks in a highly variable exposure environment over extended periods of time are not well developed. In the context of a longitudinal study of noise-induced hearing damage among construction workers, we have developed a series of exposure metrics that attempt to capture the important characteristics of noise exposure—exposure level, intermittency and ‘peakiness’—using available dosimetry measurements. This paper describes the development of these metrics and their relationship to each other.

METHODS

Dataset

A longitudinal study of noise-induced hearing loss among construction trade apprentices forms the basis of this study (Seixas *et al.*, 2004, 2005). Apprentices were recruited during their first year of apprenticeship from nine trade groups. After baseline questionnaires and hearing tests, the subjects were tested annually for 4 years. At each testing session, questionnaires were administered including a detailed task-based work history for the time period since last examination.

A database of full-shift noise dosimetry measurements on construction workers has been developed over the past 7 years. Some of these data have been

previously reported (Neitzel *et al.*, 1999; Seixas *et al.*, 2001; Reeb-Whitaker *et al.*, 2004). All noise surveys have used essentially the same techniques. Workers in nine trades were monitored on numerous commercial construction projects in the vicinity of Seattle, WA. Noise dosimeters (Quest, Q-300) were worn for a full shift while data-logging on 1 min intervals. During monitoring, each subject was asked to fill out a task-card designating the start and end time in which they conducted a number of pre-defined tasks (Neitzel *et al.*, 1999). Cards were designed for each trade so that unique trade-specific tasks could be identified. A total of ~120 possible tasks were listed on the 11 different trade-specific cards. This inventory of tasks was developed with the assistance of contractors, construction workers and construction apprenticeship coordinators. Workers also reported their environment (enclosed, partially enclosed, outdoors, unknown), presence of a nearby high noise source (yes, no) and the number of other workers nearby (>3, ≤3, unknown) on the cards. The completed task card information was merged with the dosimetry files to identify the task conducted during each minute of recorded exposure.

The dataset was extensively checked and errors and untenable data were adjusted or removed from the dataset as follows. First, short duration shifts (<300 min) were removed. Second, means and SDs of 1 min L_{eq} and L_{avg} within a shift were calculated, and shifts with means <10 dB (indicating an extremely biased full-shift average based almost completely on sub-threshold levels of 0 dB) or SDs >20 dB (again indicating a large portion of sub-threshold shift minutes) were removed. Third, all values of L_{eq} and L_{avg} were given a minimum value of 69.9 dB—the lowest level recorded accurately by the dosimeters and a level below which exposure on a working construction site is unlikely to drop. Fourth, relationships between each of the 1 min metrics were examined to insure that $L_{avg} \leq L_{eq} \leq L_{max} \leq L_{peak} \leq 146.6$ dB. If any of these relationships were violated, the measured L_{eq} was assumed correct and the other metrics were accordingly adjusted down (in the case of L_{avg}) or up (for all other metrics). Full shift measurements with an error >10% in minutes in these relationships were removed. Finally, extreme outliers in the ratio metrics defined below ($L_{eq}/L_{avg} > 50$, or $L_{max}/L_{eq} > 1000$) were also removed from the dataset. Most of these extreme values represented shifts with large portions of time spent below threshold, or where two channels reported such widely divergent values as to be unlikely or impossible. Although the adjusted and removed data points affected the calculated values of the metrics, it was felt that they generally represented frank errors in dosimetry measurement, rather than unusual but accurately recorded exposures. These data-cleaning criteria resulted in the removal of only a small fraction of the total data.

Metrics definition

Several exposure metrics were logged each minute during every monitored workshift: L_{eq} , using a 3 dB exchange rate, slow response time (1 s) and no threshold; L_{avg} , using a 5 dB exchange rate, slow response time and a 80 dB threshold; and L_{max} , with a fast (0.125 s) response time. L_{max} with a slow (1 s) response time and L_{peak} were also recorded, but because of their similarity to L_{max} (fast response) they were not analysed for this paper. All levels described in this paper were measured using A-weighting. The L_{eq} and L_{avg} represent average levels integrated over a 1 min period and L_{max} represents the highest maximum level measured during that same period.

The three metrics defined above (L_{eq} , L_{avg} and L_{max}) were summarized for a full shift, or within a specific task-event (e.g. the duration in which a specific task was conducted within a single workshift by a subject), using the following:

$$L_{ij} = q \log_{10} \left[\frac{1}{\sum_k w_{ijk}} \sum_{k=1}^{n_{ij}} w_{ijk} 10^{L_{ijk}/q} \right], \quad (1)$$

where L_{ij} is the level for individual i over shift (or task) j summarized over the $k = 1$ to n_{ij} time intervals. The n time intervals were nominally 1 min, but where more than one task was reported within a minute, the time-periods were fractions of the minute represented by w . The L_{eq} and L_{max} were calculated using $q = 10$, while L_{avg} was calculated using $q = 16.6$ (Earshen, 2000). Note that L_{max} across a sampling period represents the average of the L_{max} levels measured within that period, and as with L_{eq} and L_{avg} levels, averaging was conducted on the exponential scale to appropriately weight the higher-level maxima.

These three metrics represent different, widely accepted measures of average levels across the work period. Two novel metrics were also developed to represent the variability of exposure. It is well known that L_{avg} and L_{eq} are equivalent in steady-state noise, and that their difference becomes larger with increasingly variable noise levels within an exposure interval. Thus, the ratio of L_{eq} to L_{avg} represents the degree of variability, or intermittency independent of the average exposure level during the time interval. L_{eq}/L_{avg} was thus calculated:

$$L_{eq}/L_{avg} = \frac{1}{\sum_k w_{ijk}} \sum_{k=1}^{n_{ij}} w_{ijk} \frac{10^{L_{eq,ijk}/10}}{10^{L_{avg,ijk}/10}}. \quad (2)$$

Similarly, the degree of 'peakiness' of an exposure can be represented by the ratio of the L_{max} to the L_{eq} . The ratio of L_{max}/L_{eq} can be thought of as similar to an average crest factor (ratio of peak to average level) over the exposure interval and was calculated with equation (2), exchanging L_{max} for L_{eq} and L_{eq} for L_{avg} .

In summary, five metrics were compared. L_{eq} , L_{avg} and L_{max} indicate average levels of exposure over 1 min to full-shift sampling periods. Two ratio metrics were also developed: L_{eq}/L_{avg} expresses the average variability in levels and increases in increasingly intermittent exposures and L_{max}/L_{eq} expresses the degree of 'peakiness' across the exposure period. Each of these metrics was described by trade and compared using scatter plots and Pearson's correlation coefficients on a 1 min scale and for full shifts. These full-shift metrics were also described by using one-way analysis of variance models with trade as the predictor.

Estimation of metrics for cohort exposure assignment

In order to estimate exposure levels for the cohort, a task-based approach was adopted to exploit the substantial differences in exposure between tasks and the variable time study subjects reported doing specific tasks. These analyses used a 'trade/task event' as the basic unit of analysis. A trade/task event was defined as the period of time during a single workshift that an individual subject (belonging to a particular trade) reported a single task. Although many tasks were reported by several trades, they were kept independent to allow for different trade-specific exposure levels.

Three alternative estimation strategies were developed and compared: calculation of a simple mean exposure by trade, calculation of mean by trade/task and linear modelling using trade/task and other available covariates. To compare these different estimation approaches, model development and validation data subsets were created by selecting 1 min data based on a 50% random sample of trade/task events.

Exposure estimates were first made by calculating mean exposures by trade-task or by trade in the model development data subset. The means were estimated using equation (1) to aggregate minute-long exposure levels by trade/task event, and then the mean of those levels in dB were used to calculate the trade, or trade/task exposure level.

Linear models were also used to estimate exposure level by trade/task while controlling for available covariates. Model selection was conducted using the SAS procedure MIXED with maximum likelihood estimation, with a random effect for subject and fixed effects for trade/task, environment, near high noise source and number of other workers nearby. Variable selection was conducted for all five metrics by comparing nested models with a likelihood ratio test. Final model estimates were produced by running the same models using a restricted maximum likelihood method. Because linear models produce estimated means, and noise levels in decibels combine exponentially, the levels estimated by the

linear model results were adjusted with the residual variance (Seixas *et al.*, 2003).

Estimated levels (trade mean, trade-task mean and model-predicted) were calculated for the validation data subset using the three sets of predictions from the training data subset. Estimation methods were evaluated by comparing predicted and observed levels in the validation data subset in terms of bias and precision (Hornung, 1991) and correlation.

Based on these analyses, exposure estimates based on the trade/task mean level were selected to predict the exposures for the cohort. The annualized 2000 h equivalent exposure level was estimated for each subject for the interval between annual hearing tests. The hours per year each individual spent doing a particular trade/task ($H_{tt,i}$) was calculated from the duration of jobs and the time spent doing specific tasks within each job, as reported in each follow-up survey. Individual hours per year were then combined with the trade-task specific exposure level to obtain the individual 2000 h equivalent exposure level. By standardizing to 2000 h, these measures increase with longer durations between examinations or more hours worked per year. For L_{eq} , this was:

$$L_{eq\ 2000,i} = q \text{Log}_{10} \left[\frac{1}{2000} \sum_{TT} H_{tt,i} 10^{L_{tt,i}/q} \right]. \quad (3)$$

The same calculation was used for L_{avg} and L_{max} . As before, q was set to 10 for L_{eq} and L_{max} , and to 16.6 for L_{avg} .

For the ratio metrics, an average level, not standardized to 2000 h, was calculated to represent the average variability of exposure over the examination interval:

$$(L_{eq}/L_{avg})_i = \frac{1}{\sum H_i} \sum_{TT} H_{tt,i} (L_{eq}/L_{avg})_{tt}. \quad (4)$$

The same equation was used to obtain average L_{max}/L_{eq} for each subject and survey interval.

Subject specific exposure metrics were then described and the correlation between the subject-specific metrics calculated to demonstrate the potential importance of alternative exposure metrics for the epidemiologic analysis.

RESULTS

A total of 770 shift-long measurements were collected on numerous commercial construction sites from 1997 to 2004. Nine different trades were measured: carpenters, cement masons, electricians, insulators, ironworkers, labourers, masons, operating engineers and sheet metal workers. The 770 work shifts included almost 375 000 min of work time (Table 1). Since several tasks could be reported during the same minute, the data were restructured to include 455 447 exposure periods with associated weights, w , ranging from 1 (only 1 task reported) to 0.167 (6 tasks reported).

Twenty five shifts were removed because they represented only partial shifts (<300 min). After examining the distribution of the remaining data, an additional 15 shifts with very low mean exposure levels ($L_{avg} < 10$ dBA) or high within-shift SDs (L_{eq} SD > 20 dBA) were deleted. Additional data errors in which metrics displayed impossible relations to each other (e.g. $L_{avg} > L_{eq}$) were examined and adjustments made where indicated. In no case did these errors represent >10% of a work shift and so no additional shifts were removed. Finally, the distributions of the ratio metrics were examined and 381 additional minutes with extremely high values were removed. In total, 5.2% of work shifts and 3.5% of work time were edited out of the dataset, leaving 361 492 min over 730 work shifts for analysis (Table 1).

The full dataset of 1 min exposure metrics is described by trade in Table 2, and full-shift measurements are described by trade in Table 3. Note that the 1 min data in Table 2 are described arithmetically, and the means, therefore, under-represent integrated exposure levels that would be calculated using appropriate averaging equations [e.g. equation (1)]. Over all trades, the 1 min L_{avg} was ~3.5 dBA < L_{eq} , while L_{max} was ~14 dBA higher. L_{avg} and L_{eq} were very highly correlated ($r = 0.96$) and L_{max} was well correlated with L_{eq} ($r = 0.88$) and somewhat less well with L_{avg} ($r = 0.77$). The ratio metrics were poorly correlated with other measures and only slightly correlated with each other ($r = 0.35$). Differences in average metrics between trades were relatively small with the exception of the L_{max}/L_{eq} ratio, which fluctuates greatly both within and between

Table 1. Results of data cleaning on number of observations, minutes and shifts for analysis

Action	Observations (k) ^a		Minutes		Shifts	
	Number	Cumulative % removed	Number	Cumulative % removed	Number	Cumulative % removed
Initial dataset	455 447	0	374 760	0	770	0
Remove shifts <300 min	448 740	1.5	369 383	1.4	745	3.2
Remove shifts with outlier means, SDs	441 167	3.1	361 873	3.4	730	5.2
Remove minutes with outlier ratios	440 656	3.2	361 492	3.5	730	5.2

^aObservations (k) are minutes, or fractions of minutes when multiple tasks reported, as in equation (1).

Table 2. One-minute exposure metrics by trade [mean (SD)]

Trade	<i>n</i> (min)	L_{avg} (dBA)	L_{eq} (dBA)	L_{max} (dBA)	L_{eq}/L_{avg}	L_{max}/L_{eq}
All	361 492	77.8 (8.7)	81.2 (8.1)	95.3 (10.5)	2.6 (1.7)	45.9 (53.9)
Carpenters	66 231	77.8 (8.6)	81.3 (8.4)	96.1 (11.4)	2.6 (1.7)	53.7 (60.1)
Cement Masons	14 764	77.5 (8.7)	80.8 (7.8)	94.0 (9.5)	2.5 (1.5)	36.0 (42.7)
Electricians	114 827	76.8 (8.0)	80.5 (7.6)	94.9 (10.1)	2.8 (1.7)	46.6 (54.2)
Insulators	11 597	74.5 (6.7)	77.9 (6.2)	91.1 (8.3)	2.7 (1.7)	39.2 (47.4)
Ironworkers	35 752	79.5 (9.1)	82.9 (8.5)	97.9 (11.3)	2.6 (1.6)	54.4 (59.7)
Labourers	31 432	78.2 (9.1)	81.7 (8.6)	96.3 (10.7)	2.6 (2.0)	50.3 (57.4)
Masons	34 437	77.2 (9.4)	80.5 (8.9)	94.0 (10.9)	2.5 (1.5)	40.6 (48.2)
Operating engineers	31 296	81.3 (9.3)	83.8 (8.3)	95.4 (9.9)	2.1 (1.9)	29.9 (40.0)
Sheet metal	21 156	76.9 (7.1)	80.6 (6.4)	94.3 (9.1)	2.7 (1.7)	39.4 (44.3)

Table 3. Full shift exposure metrics description, and r^2 from ANOVA by trade [mean (SD)]

ANOVA	<i>n</i>	L_{eq}	L_{avg}	L_{max}	L_{eq}/L_{avg}	L_{max}/L_{eq}
Adj r^2		0.085	0.077	0.152	0.404	0.232
All	730	87.8 (5.8)	82.1 (5.7)	105.0 (5.45)	2.61 (0.44)	45.9 (14.6)
Carpenters	135	88.7 (5.5)	82.4 (5.2)	106.7 (5.6)	2.59 (0.33)	53.7 (14.3)
Cement masons	31	87.7 (5.6)	82.6 (5.6)	102.4 (3.5)	2.48 (0.37)	36.1 (8.5)
Electricians	230	86.7 (5.5)	80.8 (5.2)	104.6 (5.2)	2.77 (0.39)	46.6 (13.0)
Insulators	23	81.8 (3.8)	77.1 (4.8)	98.9 (2.6)	2.65 (0.52)	39.5 (12.6)
Ironworkers	71	91.1 (5.4)	84.8 (5.2)	109.1 (4.2)	2.56 (0.38)	55.4 (11.4)
Labourers	63	89.1 (6.6)	83.1 (6.5)	106.6 (5.5)	2.62 (0.31)	50.2 (13.5)
Masons	73	88.5 (6.7)	82.7 (6.7)	102.9 (5.3)	2.52 (0.42)	40.4 (13.0)
Operating engineers	61	88.4 (5.0)	84.5 (5.8)	103.5 (5.2)	2.14 (0.62)	30.4 (12.9)
Sheet metal workers	43	85.7 (4.2)	80.3 (3.7)	102.9 (4.2)	2.73 (0.42)	39.3 (9.7)

trade. Although operating engineers and ironworkers have relatively similar L_{eq} levels (84 and 83 dBA, respectively), the exposure variability ratio metrics were much higher for ironworkers, as might be expected given the nature of their work. For operating engineers and ironworkers the L_{eq}/L_{avg} was 2.1 and 2.6, and L_{max}/L_{eq} was 30 and 54 respectively. Electricians also had high variable exposures with an L_{avg} among the lowest of the trades (77 dBA) but among the highest ratio metrics, 2.7 and 47 for L_{eq}/L_{avg} and L_{max}/L_{eq} , respectively.

After aggregating by shift using equations (1) and (2), the same five metrics across 730 workshifts are shown in Fig. 1 and described in Table 3. Unlike the 1 min data in Table 2, the averages in Table 3 represent the full-shift levels experienced by construction workers. The mean L_{avg} was 82.1 ± 5.7 dBA and the average L_{eq} was 87.8 ± 5.8 dBA. There was relatively little difference in mean levels between trades: average L_{eq} levels ranged from 81.2 dBA for insulators to 91.1 dBA for ironworkers. Trade was a highly significant predictor for all five metrics in a one-way analysis of variance, although only 8.5% of the variability in L_{eq} was explained by trade. The range of trade-specific mean L_{eq}/L_{avg} ratios did not appear to vary greatly (2.1 ± 0.6 for operating

engineers to 2.8 ± 0.4 for electricians) even though trade explained a surprising 40% of the overall variability. Ironworkers had the highest L_{max}/L_{eq} value (55.4 ± 11.4), followed closely by carpenters (53.7 ± 14.3). In a pattern similar to the 1 min data, the shift averaged L_{avg} and L_{eq} were highly correlated ($r = 0.93$), L_{max} was well correlated with L_{eq} ($r = 0.85$) and somewhat less so with L_{avg} ($r = 0.71$) (Fig. 1). The ratio metrics were poorly correlated with the average levels.

Three alternative methods used to estimate task-based exposure levels (trade mean, trade/task mean, model predicted) for integration with the cohort's work history were developed and tested in the model development data subset. Of the 108 trade/tasks in the complete dataset, 44 occurred more than 10 times (1536 trade/task events). A 50% random sample of these trade/task events was selected and removed to a 'validation' data subset containing 768 trade/task events in 44 trade/tasks. The remaining model development data subset consisted of 212 394 min representing all 108 trade/tasks.

Using the model development data subset, equations (1) and (2) were used to aggregate by trade/task event and these levels were then averaged by trade, and by trade/task. In addition, linear models including

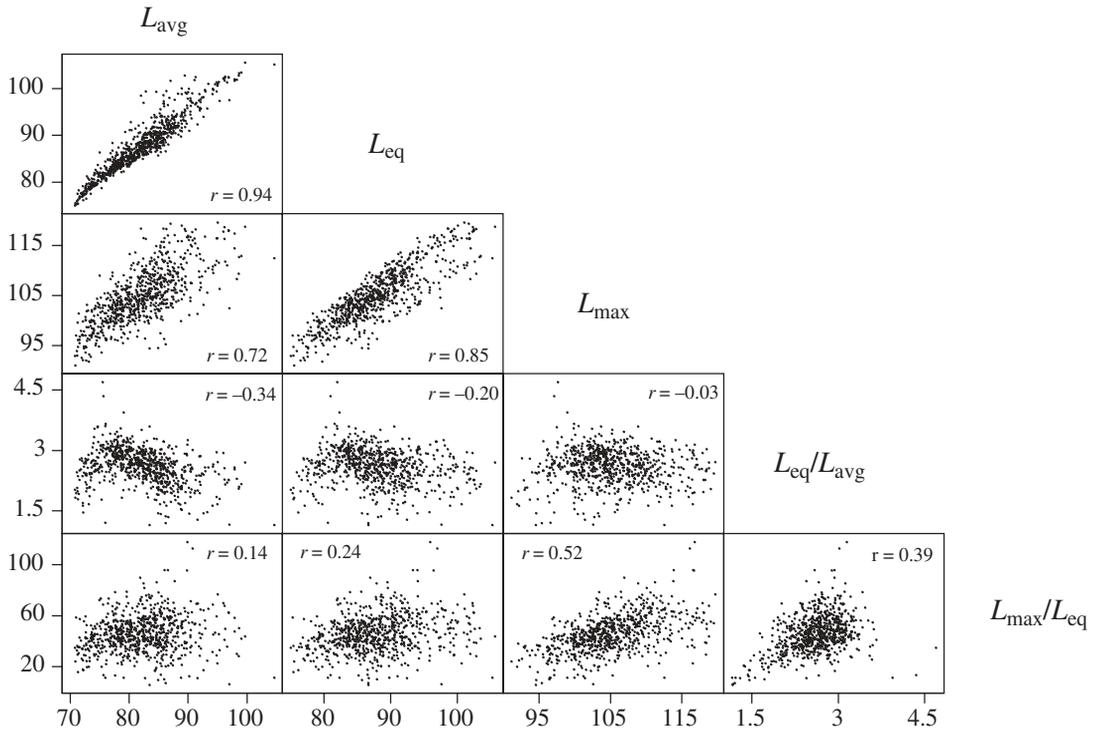


Fig. 1. Scatter plot matrix of full shift average alternative noise exposure metrics ($n = 730$).

trade/task, other sources of high noise in the area (Y/N), work environment (indoors, outdoors, partially enclosed) and number of workers in the vicinity were developed for each of the five metrics. The number of workers in the vicinity did not increase the model fit and was dropped from the analysis.

Using the results of the three estimation methods, predicted exposure metrics were calculated for comparison with the 768 observed levels in the validation data subset. Trade and trade-task mean levels were directly computed, and coefficients from the multivariate model were used to predict the third estimate. The model-predicted estimate was adjusted with the model's MSE to account for non-linear averaging (Seixas *et al.*, 2003). The comparisons between the three estimates of L_{eq} and the observed levels in the validation subset are displayed in Fig. 2. All three comparisons show that the estimation procedures severely restrict the variability in the observed data. The most dramatic restriction of variability is shown using the trade-specific mean levels, which take on only nine possible values and fall within a very small range. Although the model-predicted estimates take on more possible values (trade/task adjusted for the other covariates in the model), the pattern is similar to the trade/task specific means.

These observations are further characterized for all five metrics in terms of bias (average difference in observed and predicted values) and precision (SD of the differences) and the correlation of the observed

and predicted values, in Table 4. Trade/task means resulted in very little bias for any of the metrics, while the model-predicted and trade mean values for the noise metrics had a small negative bias. The precision of the trade means is larger (less precise) than for the model-predicted levels, or trade/task means. The correlation of the predicted and observed values is lower when using the trade means, and the trade/task and model-predicted mean correlations are higher and very similar. Based on these results, the trade/task mean exposure metric was adopted. Use of the trade-task mean preserves the task-based estimation approach, while considerably simplifying the computations needed for the multivariate modelling method.

Trade/task means were re-estimated based on the whole dataset, including both model development and validation data subsets, and the results merged with the cohort's task-based work history for each examination interval using equations (3) and (4). The 289 apprentices in the study had 700 follow-up interviews, yielding 2.4 intervals per subject. Within these examination intervals, subjects reported working 1357 jobs consisting of 5853 specific trade/tasks for an average of 1.9 jobs per interval and 4.3 tasks per job.

The correlation of the resulting individual interval-specific metrics (annualized L_{eq} , L_{avg} , L_{max} and average ratio metrics) is shown in Fig. 3. Similar to the 1 min and full shift metrics, L_{avg} and L_{eq} , and to a lesser degree L_{max} were highly correlated with each

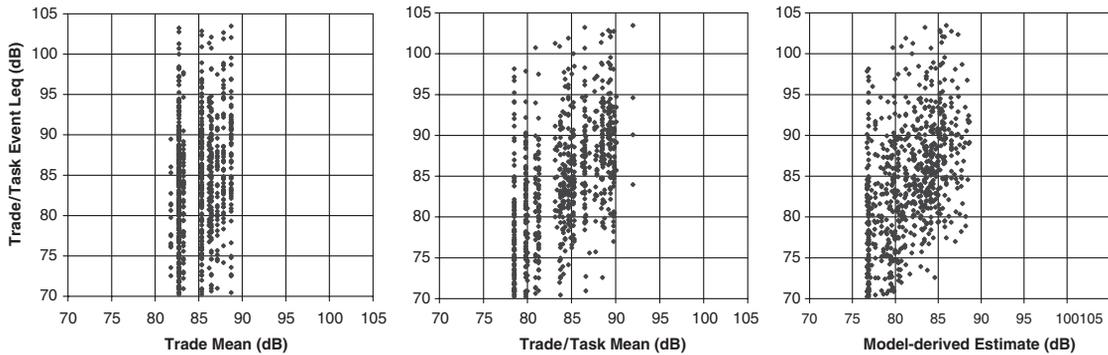


Fig. 2. Comparison of trade, trade/task and model-derived estimates of L_{eq} with observed trade/task event levels in validation data subset ($n = 768$).

Table 4. Performance of metrics in validation dataset

	Trade mean	Trade/task mean	Model predicted ^a
L_{eq}			
Bias	-0.72	0.12	-0.29
Precision	7.03	6.03	6.14
r^2	0.04	0.29	0.28
L_{avg}			
Bias	-0.79	0.09	-0.61
Precision	6.29	5.36	5.42
r^2	0.04	0.29	0.28
L_{max}			
Bias	-0.52	0.12	-0.36
Precision	7.98	6.98	7.21
r^2	0.07	0.29	0.24
L_{eq}/L_{avg}			
Bias	-0.06	-0.04	-0.05
Precision	0.58	0.56	0.55
r^2	0.03	0.12	0.12
L_{max}/L_{avg}			
Bias	-0.88	-0.19	-0.71
Precision	17.64	16.43	16.48
r^2	0.07	0.20	0.19

^aFrom linear model in which the estimate is corrected using the MSE, as per Seixas 2003.

other, while the two ratio metrics demonstrated little correlation with the other metrics. The ratio metrics were mildly correlated with each other ($r = 0.48$). L_{max} was somewhat correlated with L_{max}/L_{eq} (0.51), but not with L_{eq}/L_{avg} (-0.08).

DISCUSSION

Summarizing exposure over time to capture the important characteristics associated with risk continues to constitute a challenging area of occupational exposure assessment and epidemiology. Even in the area of noise-induced hearing loss, in which basic exposure-response relationships for steady-state

noise have been known for many years, and in which the importance of peak exposures is increasingly recognized, a summary metric for characterizing risk has not been adequately demonstrated. Sophisticated dosimeters are available which log data on multiple channels at relatively short time intervals, but the best method for integrating these data into risk evaluation remains unclear. In this analysis, we have taken advantage of a large and well-documented dataset of noise dosimetry measurements in the construction industry—an industry with high and extremely variable exposures—to explore relations between various exposure metrics.

During exploration of the extensive accumulated dataset, numerous problems were identified, and the apparently erroneous data were either discarded or adjusted to conform to expected relationships. The edits made to these data were designed to be conservative and minimize any possible biases and only 3.5% of the minutes were removed. Nevertheless, dosimeter manufacturers and end users should consider these occasional problems whenever large noise dosimetry datasets are being collected as erroneous data can lead to incorrect hearing conservation decisions.

Two standard measures of average exposure level, the L_{eq} and L_{avg} , were compared. Previous studies in construction (Neitzel *et al.*, 1999), trucking (Seshagiri, 1998) and fixed industry with variable noise levels (Sriwattanatamma and Breyse, 2000) have shown average differences in mean L_{avg} and L_{eq} levels of over 6 dBA. The fraction of the monitored population with exposures >85 dBA was also substantially increased using the L_{eq} metric in these previous studies. The primary difference between L_{eq} and L_{avg} is the exchange rate of 3 and 5 dB, respectively. Reviews of the literature on exchange rates generally indicate that although some hearing damage recovery may occur during intermittent quiet periods, there is no adequate scientific basis for the use of a 5 dB ER (Suter, 1992; NIOSH, 1998; ACGIH, 2001). While the 3 dB ER may be

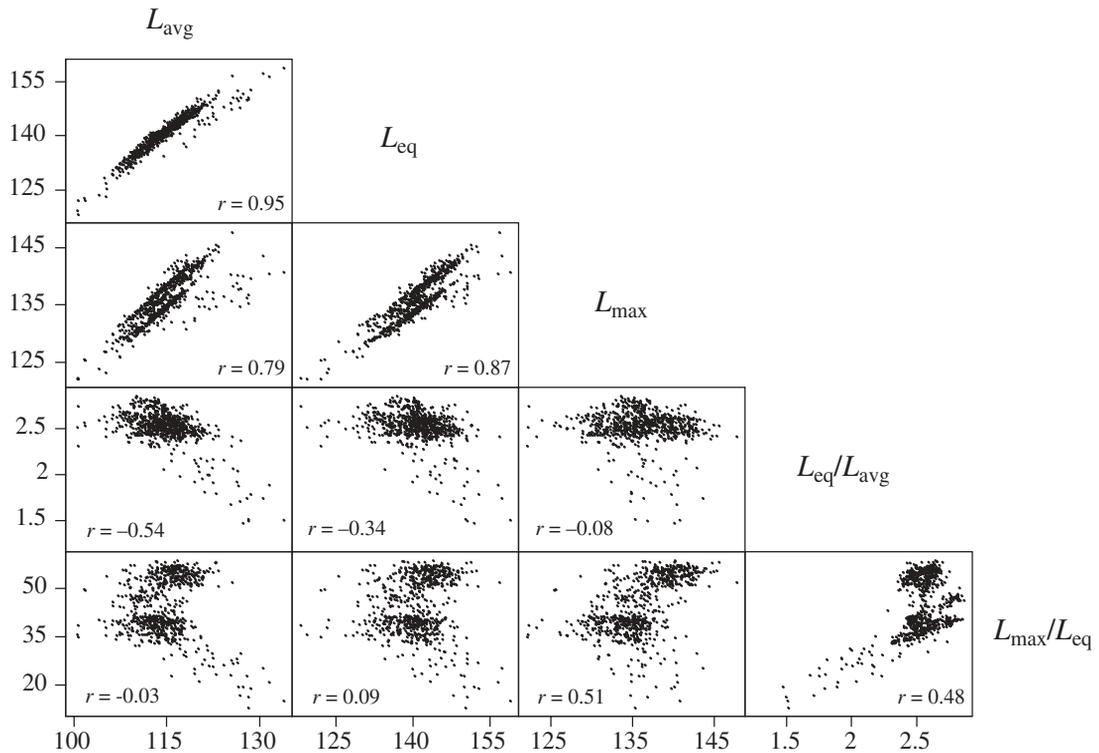


Fig. 3. Scatter plot matrix of alternative noise metrics applied to the apprentices cohort on the basis of trade/task activities over the previous study interval ($n = 700$ intervals).

over-protective under some circumstances, the 5 dB ER will often underestimate risk. Nevertheless, the US OSHA continues to use a 5 dB ER and has rejected addressing this important issue in recent rule-making proceedings for the construction industry (OSHA, 2002).

In the current analysis, large differences in the absolute levels of L_{avg} and L_{eq} were observed; however the two metrics were very highly correlated ($r > 0.9$). As a result, use of the L_{eq} or L_{avg} in an epidemiologic study should make very little difference in the slope of the exposure–response relationship although the risk at any particular exposure magnitude would be substantially different using one measure or the other. In adopting a standard based on one metric, the data used to set the standard must be based on the same metric—especially when applied to industries with variable noise patterns.

While the L_{eq} integrates short-term high levels on an equal energy basis, animal studies have demonstrated that peak levels above a ‘critical level’ of ~ 120 – 135 dB (Roberto *et al.*, 1985; Danielson *et al.*, 1991; Henderson *et al.*, 1994; Levine *et al.*, 1998) produce more damage than expected using the EEH. This critical level, which is not well established in humans, marks the transition from metabolic to direct mechanical damage to the hearing mechanism (Price, 1981; Lataye and Campo, 1996). Construction workers are regularly exposed to peak noise levels

>120 dB; in an earlier analysis of a subset of these data, an average of 18 min per shift included peaks >140 dB (Neitzel *et al.*, 1999). Even if the L_{eq} underestimates the damage risk from high-level noise exposures, and needs augmentation with a peak metric, it is still more protective than the L_{avg} used by OSHA for all exposure scenarios.

Although peak exposures are common in some work settings, and exposure to impulse noise may be more damaging than longer exposure to lower-level sounds of the same total energy, methods of integrating peak characteristics into exposure metrics for epidemiologic purposes have not been demonstrated. Noise peaks are defined in terms of their amplitude, duration, rise time, number of impulses and repetition rate, each of which affects the risk of hearing damage (Henderson and Hamernik, 1986). Existing exposure standards almost uniformly specify a peak limit of 140 dB, but the scientific evidence supporting this choice of levels is scant (NIOSH, 1998) and peak limits based solely on amplitude may be inadequate (McRobert and Ward, 1973; Eiber, 1999).

Two models of hearing risk associated with peak noise exposures have been suggested. Price’s Auditory Hazard Assessment Algorithm for the Human ear (AHAH) is a mathematical model of the ear that incorporates spectral, amplitude and duration data into the exposure evaluation process and that has

been validated for human exposures >130 dB (Price, 2003). Time or frequency domain kurtosis (the ratio of fourth moment to the second moment squared) (Erdreich, 1986) has been used as a predictor of the magnitude and frequency distribution of noise-induced hearing damage in animal models (Lei *et al.*, 1994). Kurtosis is attractive as an exposure metric because it represents a unified measure of the 'peakiness' of the distribution of a noise signal. Despite the apparent utility of these two peak exposure models, however, both are currently most applicable to short term exposures—i.e. single gun shots or laboratory-created exposure patterns—and require high-speed digital sampling. Their use in assessing long-term real-world occupational exposures is therefore limited.

L_2 , the level exceeded for 2% of the exposure time, was also evaluated on a subset of the full-shift exposures assessed here. Burns and Robinson found that the linear fit of hearing loss over time was slightly improved when subjects were grouped by their average L_2 rather than L_{eq} levels (Burns and Robinson, 1970), presumably because L_2 better represented exposure profiles with a high peak component. In the current study, L_2 was measured by dosimetry for 350 full-shift exposure measurements across all nine trades. Although L_2 was expected to address the 'peakiness' in the data, it was very highly correlated with L_{avg} and L_{eq} (r of 0.91 and 0.88, respectively) and not correlated with the ratio metrics. Although this result was somewhat unexpected, it can be explained by the fact that the all three metrics are largely driven by the higher end of the exposure distribution due to exponential averaging [e.g. equation (1)].

To address the shortcomings of simple average metrics in a variable noise environment, two novel exposure metrics were derived. L_{avg}/L_{eq} was designed to address variability over time by exploiting the difference between levels measured using the 3 and 5 dB exchange rates. L_{max}/L_{eq} also addresses variability, but was designed to be more sensitive to high exposure levels within each minute. While the two metrics are defined similarly, they appear to capture different information as they are poorly correlated with each other and uncorrelated with the average metrics. Trade/tasks with the highest L_{max}/L_{eq} are generally those with high exposure to impact-type noise, i.e. carpenters setting forms and conducting demolition work (L_{max}/L_{eq} of 71 and 77), or ironworkers laying metal decking (L_{max}/L_{eq} of 61). Trade tasks with the highest levels of L_{eq}/L_{avg} are those which typically have highly intermittent exposures, such as labourers manually moving materials (L_{eq}/L_{avg} of 3.0) and electricians installing cable trays (L_{eq}/L_{avg} of 2.9).

In order to assign exposure levels to subjects within an occupational cohort, exposure levels must be linked to the subjects' work histories via some intermediate variable. Linkage is commonly made using

job title, as with a job exposure matrix (Goldberg *et al.*, 1993). More recently, exposures have been modelled based on a set of exposure determinants (i.e. job, factory, production rate) and linked via these same determinants in the work history (Burstyn and Teschke, 1999). A third approach, task-based exposure assessment, refines the job-based approach to capture the variability associated with varying distribution of time spent in individual tasks even by subjects within the same job title. While the task-based approach seems appropriate for an industry with highly variable work such as construction, estimation of individual exposure level was not especially accurate in a study of five construction trades (Seixas *et al.*, 2003). The approach has shown more promise when used to target high-hazard noise sources for exposure reduction efforts (Stephenson, 1995).

These three approaches were tested in the current dataset by splitting the data into model development and validation subsets, developing estimates in the training dataset and comparing these with the observed values in the validation subset. The results are most graphically demonstrated in Fig. 3. The very restrictive range of exposures using the trade mean—only nine values for all possible exposure scenarios—produced little relation with observed levels; only 4% of the variability in trade/task exposure levels was explained by the trade mean. Trade/task means did considerably better explaining ~30% of the variability. Multivariate modelling introduced a refinement of the predicted mean, but the correlation with observed levels did not improve and the non-linear averaging of noise levels necessitated an adjustment for variability. Despite this adjustment, a bias remained in these estimates, probably due to violations of the homoscedasticity (equal variance between groups) assumption of the linear model.

Having chosen the trade/task estimate as our most accurate method of linkage to individuals, the final assignment of exposure metrics to subjects required integration over work time using equations (3) and (4). While the metrics that described average exposure levels (L_{avg} , L_{eq} , L_{max}) were constructed so that they would increase with longer periods of time, the variability metrics (L_{eq}/L_{avg} , L_{max}/L_{eq}) were calculated as simple averages. As a result, exposure response modelling could incorporate one of each type of metric—one to express exposure intensity and one to adjust for variability.

CONCLUSION

It is well known that errors in estimation of exposure can produce an attenuation of exposure–response relations. In fact, the degree of attenuation (in a linear dose–response model) can be determined by the correlation between the exposure metric used and

the true, underlying metric (Armstrong *et al.*, 1992). In our setting, assuming L_{eq} was the correct metric, very little attenuation would be expected by the use of L_{avg} , or even L_{max} . On the other hand, if variability and peakiness are strong predictors of damage, beyond that associated with the average level, then these metrics may considerably alter the observed relation between construction noise exposure and hearing damage. Future application of these metrics to the longitudinal hearing levels observed in our cohort should provide insight into the importance of these different exposure metrics and the relative errors associated with their estimation.

Acknowledgements—This project was supported by the National Institute for Occupational Safety and Health of the US Centers for Disease Control and Prevention, Grant number 5 RO1 OH03912. The authors thank Karen Powers for her skillful development and analysis of the datasets the construction industry apprentices, apprenticeship coordinators and contractor representatives that made this project possible.

REFERENCES

- ACGIH. (2001) 2001–2002 threshold limit values for chemical substances and physical agents and biological exposure indices (BEIs): noise. Cincinnati, OH: American Conference of Governmental Industrial Hygienists.
- Armstrong B, White E, Saracci R. (1992) Exposure measurement error and its effects. In Kelsey JL, Marmot MG, Stolley PD, Vessey MP editors. Principles of exposure measurement in epidemiology. Vol. 21. New York: Oxford University Press.
- Atherley GR, Martin AM. (1971) Equivalent-continuous noise level as a measure of injury from impact and impulse noise. *Ann Occup Hyg*; 14: 11–23.
- Burns W, Robinson D. (1970) Hearing and noise in Industry. London: HMSO.
- Burstyn I, Teschke K. (1999) Studying the determinants of exposure: a review of the methods. *Am Ind Hyg Assoc J*; 60: 57–72.
- Checkoway H, Pearce N, Kriebel D. (2004) Research methods in occupational epidemiology. 2nd edn New York: Oxford University Press.
- Dancer A, Buck K, Parmentier G *et al.* (1998) The specific problems of noise in military life. *Scand Audiol Suppl*; 48: 123–30.
- Danielson R, Henderson D, Gartton MA *et al.* (1991) The importance of 'temporal pattern' in traumatic impulse noise exposures. *J Acoust Soc Am*; 90: 209–18.
- Earshen J. (2000) Sound measurement: instrumentation and noise descriptors. In Berger E, Royster L, Royster J, Driscoll D, Layne M, editors. The noise manual. Fairfax, VA: American Industrial Hygiene Association.
- Eiber A. (1999) Mechanical modeling and dynamical behavior of the human middle ear. *Audiol Neurootol*; 4: 170–7.
- Erdreich J. (1986) A distribution based definition of impulse noise. *J Acoust Soc Am*; 79: 990–8.
- Goldberg M, Kromhout H, Guenel P *et al.* (1993) Job exposure matrices in industry. *Int J Epidemiol*; 22(Suppl. 2): S10–S15.
- Henderson D, Hamernik RP. (1986) Impulse noise: critical review. *J Acoust Soc Am*; 80: 569–84.
- Henderson D, Spongr V, Subramaniam M *et al.* (1994) Anatomical effects of impact noise. *Hear Res*; 76: 101–17.
- Hornung R. (1991) Statistical evaluation of exposure assessment strategies. *Appl Occup Environ Hyg*; 6: 516.
- Lataye R, Campo P. (1996) Applicability of the L_{eq} as a damage-risk criterion: an animal experiment. *J Acoust Soc Am*; 99: 1621–32.
- Lei S-FA, Ahroon W, Hamernik RP *et al.* (1994) The application of frequency and time domain kurtosis to the assessment of hazardous noise exposures. *J Acoust Soc Am*; 96: 1435–44.
- Levine S, Hofstetter P, Zheng XY *et al.* (1998) Duration and peak level as co-factors in hearing loss from exposure to impact noise. *Scand Audiol Suppl*; 48: 27–36.
- McRobert H, Ward WD. (1973) Damage-risk criteria: the trading relation between intensity and the number of nonreverberant impulses. *J Acoust Soc Am*; 53: 1297–300.
- Neitzel R, Seixas N, Camp J *et al.* (1999) An assessment of occupational noise exposures in four construction trades. *Am Ind Hyg Assoc J*; 60: 807–17.
- Nieuwenhuijsen M. (2003) Exposure assessment in occupational and environmental epidemiology. Oxford: Oxford University Press.
- NIOSH. (1998) Criteria for a recommended standard: occupational noise exposure. Cincinnati, OH: US Department of Health and Human Services/Centers for Disease Control, National Institute for Occupational Safety and Health.
- OSHA. (1983) Occupational noise exposure: hearing conservation amendment; final rule. *Fed Reg*. Washington, DC, US Department of Labor, Occupational Safety and Health Administration. 48: 9738–85.
- OSHA. (2002) Advance notice of proposed rulemaking: hearing conservation program for construction workers. *Fed Reg* 67. Washington, DC, US Department of Labor, Occupational Safety and Health Administration: 50610–8.
- Price G. (1981) Implications of a critical level in the ear for assessment of noise hazard at high intensities. *J Acoust Soc Am*; 69: 171–7.
- Price G. (2003) Auditory hazard units an index of risk from intense sounds. National Institute for Occupational Safety and Health Best Practices Workshop: Impulsive Noise, Cincinnati, OH.
- Reeb-Whitaker CK, Seixas NS, Sheppard L *et al.* (2004) Accuracy of task recall for epidemiological exposure assessment to construction noise. *Occup Environ Med*; 61: 135–42.
- Roberto M, Hamernik R, Salvi RJ *et al.* (1985) Impact noise and the equal energy hypothesis. *J Acoust Soc Am*; 77: 1514–20.
- Seixas NS, Checkoway H. (1995) Exposure Assessment in industry-specific retrospective occupational epidemiology studies. *Occup Environ Med*; 52: 625–33.
- Seixas NS, Ren K, Neitzel R *et al.* (2001) Noise exposure among construction electricians. *Am Ind Hyg Assoc J*; 62: 615–21.
- Seixas NS, Sheppard L, Neitzel R *et al.* (2003) Comparison of task-based estimates with full-shift measurements of noise exposure. *Am Ind Hyg Assoc J*; 64: 823–9.
- Seixas NS, Kujawa S, Norton S *et al.* (2004) Predictors of hearing threshold levels and distortion product otoacoustic emissions among noise exposed young adults. *Occup Environ Med*; 61: 899–907.
- Seixas NS, Goldman B *et al.* (2005) Prospective noise induced changes to hearing among construction industry apprentices. *Occup Environ Med*; in press.
- Seshagiri B. (1998) Occupational noise exposure of operators of heavy trucks. *Am Ind Hyg Assoc J*; 59: 205–13.
- Sriwattanatamma P, Breyse P. (2000) Comparison of NIOSH noise criteria and OSHA hearing conservation criteria. *Am J Ind Med*; 37: 334–8.
- Stephenson M. (1995) Noise exposure characterization via task-based analysis. 23rd Annual National Hearing Conservation Association Conference, Cincinnati, OH, National Hearing Conservation Association.
- Suter AH. (1992) The relationship of the exchange rate to noise-induced hearing loss. Cincinnati, OH: Alice Suter and Associates.