HHS Public Access

Evaluation and Comparison of Three Exposure Assessment Techniques

R.L. Neitzel[1], **W.E. Daniell**[1], **L. Sheppard**[1,2], **H.W. Davies**[3], and **N.S. Seixas**[1]

[1]Department of Occupational and Environmental Health Sciences, University of Washington, Seattle Washington

[2]Department of Occupational and Environmental Health Sciences and Department of Biostatistics, University of Washington, Seattle, Washington

[3]School of Environmental Health, University of British Columbia, Vancouver, BC, Canada

Abstract

This study was conducted to verify the performance of a recently developed subjective rating (SR) exposure assessment technique and to compare estimates made using this and two other techniques (trade mean, or TM, and task-based, or TB, approaches) to measured exposures. Subjects (n = 68) each completed three full-shift noise measurements over 4 months. Individual measured mean exposures were created by averaging each subject's repeated measurements, and TM, TB, and SR estimates were created using noise levels from worksites external to the current study. The bias, precision, accuracy, and absolute agreement of estimates created using the three techniques were evaluated by comparing estimated exposures with measured exposures. Trade mean estimates showed little bias, while neither the TM nor the SR techniques produced unbiased estimates, and the SR estimates showed the greatest bias of the three techniques. Accuracy was essentially equivalent among the three techniques. All three techniques showed poor agreement with measured exposures and were not highly correlated with each other. Estimates from the SR technique generally performed similarly to the TM and TB techniques. Methods to incorporate information from each technique into exposure estimates should be explored.

Keywords

exposure; noise; subjective rating; task-based assessment

INTRODUCTION

This article represents part one of a two-part study to improve estimates of exposure to construction noise over a longer period (4 months). The current study had two goals: (1) to validate the initial performance of a subjective rating (SR) exposure assessment technique described previously,[1] and (2) to compare SR exposure estimates with measured exposures and with estimates developed using two more traditional techniques in the occupational

Address correspondence to: Richard Neitzel, University of Washington, Department of Environmental and Occupational Health Sciences, Box 354695, Seattle WA 98185-4695; rneitzel@u.washington.edu.

literature (use of trade mean exposure levels and a task-based approach). The second part of this study[2] evaluated the performance of exposure estimates based on combinations of data from the SR, trade mean (TM), and task-based (TB) techniques.

BACKGROUND

Assessment of occupational exposure for workers with highly variable exposure levels and work activities—conditions increasingly common in the United States and other modern societies—remains challenging. Repeated quantitative measurement of personal exposure on individual workers represents the gold standard of assessment for occupational epidemiologic studies.[3,4] However, direct measurements on individuals are not feasible in many situations due to fiscal or logistical limitations, necessitating the use of alternative approaches to assign individual exposures.

The primary alternative to direct measurements has been use of exposure groups[5–7] often based on job title or trade, work zone or department[3,8,9] The workers who form an exposure group share a given exposure characteristic (most commonly, job title)[10] and are assumed to have similar exposure distributions across their work periods. All workers within the group are assigned a group mean level based on exposures measured on a representative sample of group members.[11,12] For exposures where dose accumulates in a linear fashion, individuals' exposures are assumed to be randomly distributed around the group mean, with measurement errors—the differences between individuals' exposures and the group mean—assumed to average to zero, resulting in an unbiased estimate of group exposure.[13] It is unclear whether this assumption holds for exposures, such as noise, where dose accumulates exponentially.[14]

The utility of exposure groups may be reduced when workers who share the defining characteristic of the exposure group are mobile and involved in a diverse and changing array of work activities, which results in increasing within-group variability and greater overlap (i.e., less contrast)[15] between groups. This situation is common in industries like construction, transportation, and agriculture. In industries like these, within-worker exposure variability (e.g., variability within and between different work shifts) is high.[16,17] Although trade-based approaches have been used extensively to characterize exposures in the construction industry,[17–22] when within-worker and within-group variability is large, creating groups based on job title may produce groups of workers with very dissimilar exposures.[7,23–25]

Several alternative techniques are available for exposure assessment among workers with highly dynamic exposures. Task-based assessment may be useful where workers within a job title conduct a range of tasks that change both within and between work shifts.[26] In this approach, workers[27,28] or their supervisors[29] report their tasks for a specific day, a "typical"[29] or "worst case"[30] day, or a longer period,[31,32] or tasks are observed by researchers.[16,33] Time-at-task information is then combined with task-specific exposure levels measured on the same individual or, more commonly, on other individuals conducting the same task, to create an average exposure estimate.[10]

TB assessment has been used in construction,[17,34,35] manufacturing,[29] repair services,[36] and agriculture and forestry,[37,38] to assess exposures such as noise,[39,40] air contaminants,[21,31,36,41] vibration,[28] and ergonomic hazards.[27,42] Benefits of TB assessment include assignment of exposure at the individual (rather than group) level, the ability to account for within- and between-subject variation in work activities, and the ability to incorporate exposure attenuation resulting from use of personal protective equipment (presuming use of such equipment is evaluated at the task level). However, TB assessment requires repeated measurements of all tasks, as well as collection of time-at-task data, which presents an additional and potentially quite important source of error.[43,44] Although TB assessment can effectively identify tasks with high exposure potential,[34,45–47] the accuracy of full-shift average TB estimates has not been adequately explored.[48] Finally, task definitions are of critical importance.[26] Some authors have used simple binary measures of task,[33] while others have used limited[16] or more comprehensive[27] sets of tasks, or allowed workers to report tasks in their own words.[49]

Subjective rating offers another alternative for exposure assessment. This method may involve evaluation of the presence, intensity, frequency, and/or duration of exposure by expert raters, workers, or supervisors.[50] Subjective rating data concerning the mere presence or absence of an exposure by workers or supervisors provides information of limited use for quantitative exposure-response studies; data on perceived exposure frequency, intensity, or duration ratings are possible for exposures that are easily sensed by workers,[50] and provide more useful information for developing quantitative exposure estimates. Perceived exposure intensity reported by workers and supervisors has been used to evaluate exposures to dust,[51,52] chemicals,[8] ergonomic hazards,[53,54] and noise.[1,55]

The benefits of SR by workers or supervisors include relative ease of data collection (via survey or interview), the possibility of evaluating multiple aspects of exposure (i.e., frequency, duration, and intensity simultaneously), and the possibility that potential determinants of exposure not evaluated by group-mean or TB approaches (e.g., work site factors, behavioral differences, and so on) are incorporated into workers' estimates. SR approaches may also incorporate exposure attenuation resulting from use of personal protective equipment.

The performance of different exposure assessment approaches can be evaluated in various ways. From an epidemiologic perspective, the most important performance measures evaluate the average size and variability of errors in estimated exposures—that is, differences between estimates and individuals' "true" exposures. Quantification of measurement error involves evaluation of both the bias and precision of estimated exposures —that is, systematic average differences between pairs of measurements made using two methods and the degree of variability in those differences, respectively.[56] Bias and precision can also be summed, resulting in a measure referred to as accuracy.[56] An important assumption underlying evaluations of bias and precision is that treatment of one set of exposure values as the "true" exposure against that the other can be compared ignores the uncertainty inherent in any estimate of exposure.

Analyses that evaluate the degree of absolute agreement between measured and estimated values are also informative. A variety of methods have been used, often inappropriately, to evaluate absolute agreement; these include the Pearson correlation coefficient, *t*-tests, and slope coefficients from linear regression analyses.[57–60] Intraclass correlation coefficients (ICCs), which evaluate the amount of overall variance in observed data due to variability in assigned ratings or levels between subjects,[61–63] are a preferred measure of absolute agreement. ICC values can be computed between measured exposures and exposures estimated using various techniques, allowing for identification of the technique that results in the greatest absolute agreement with measured levels. A variety of ICCs are available,[61,63] but for evaluating exposure estimates from different assessment techniques, one particular form of ICC (ICC(A,1), Case 3A, referred to subsequently as ICC)[61] is appropriate. This ICC summarizes absolute agreement between paired values assigned to objects of measurement (individual subjects, in this case) by different observers (or exposure assessment techniques, as in the current study). In this study, we present both absolute agreement as assessed by ICC and an analysis of predictive quality of our estimates (e.g., bias, precision, and accuracy) to thoroughly examine the performance of SR, TB, and TM exposure assessment techniques in assessing occupational construction noise exposures.

METHODS

Site and Subject Recruitment

Three large, commercial construction sites enrolled in a University of Washington (UW) hearing conservation study in the Seattle, Washington area, participated in this sub-study. Subjects were recruited from among the workers at each site who were expected to be on site for 6 months, the duration of the hearing conservation study. Non-English speakers were excluded. Research staff described the study methods and procedures to potential subjects; volunteers were enrolled after signing a consent form. All study procedures were approved by the UW Institutional Review Board, and all subjects were advised of their noise exposures at the conclusion of their participation. Each subject participated during three work shifts over a 4-month period, with each of these work shifts being separated by approximately a 2-month period. Subjects received a small monetary incentive for each work shift in which they participated.

Survey Data Collection

Subjects completed a self-administered written survey during each of the three measured work shifts.[64,65] The 63-item survey was written at a sixth-grade level (Flesch-Kincaid scale) Three SR survey items (Appendix 1[1]) related to subjects' current perceived noise exposure intensity. The first of these items ("Noise frequency") asked subjects to rate on a five-point ordered scale how often they were exposed to "high" noise. The second ("Raise voice") asked subjects to rate on a five-point ordered scale how often they had to raise their voice to be heard by someone at arm's length. The third ("Percent time in noise") asked subjects to rate the percentage of time they spent in each of five ordered noise intensity categories. Responses for the "Noise frequency" and "Raise voice" items were collapsed post hoc from five categories to three by combining the middle three categories into one. We collapsed responses for the "Percent time in noise" item from five categories to three by

summing the percent of time reported for the three middle categories and then selecting as the item response the post-collapse category with the greatest percent of time. These post hoc treatments of the SR items were based on our earlier work.[1]

An additional three SR survey items evaluated variability in subjects' exposures. The first ("Noise bursts") evaluated exposures to brief but intense impulse noise on a five-point scale. The second ("Noise variability") evaluated changes in exposure levels over time on a four-point scale. The third ("Task variability") assessed between-shift changes in tasks performed on a four-point scale. As with the perceived intensity SR items, we collapsed the three exposure variability items post hoc.[1] The "Noise bursts" item was collapsed from five categories to three by combining the three middle categories, and the "Noise variability" and "Task variability" items were collapsed into two categories by combining the two lower categories together and two upper categories together.

Dosimetry and Activity Card Data Collection

Subjects wore a noise dosimeter during each of the three work shifts in which they completed a survey. The noise measurement methodology has been described previously.[65,66] Briefly, subjects wore a datalogging dosimeter (Q-300 or Noise-Pro DLX; Quest Technologies, Oconomowoc, Wisc.) fit by researchers for the entire work shift. The dosimeters logged noise levels at 1-min intervals throughout the shift. Metrics logged during the 1-min intervals included the equivalent continuous level (L_{EQ}) recommended by the National Institute for Occupational Safety and Health (NIOSH),[67] the average level (L_{AVG}) required by the Occupational Safety and Health Administration (OSHA),[68] and the maximum level (L_{MAX}) measured using a slow response time. Dosimeters were calibrated pre- and post-measurement.

During measured shifts, subjects completed an activity card reporting the timing of each of the various tasks they performed, their use of hearing protection devices (HPDs), and various environmental factors. Information from this type of activity card has shown good agreement with researcher observation.[17,32,39] Subjects reported the timing of their activities with approximately 15-min time resolution.

External Noise Level Data

Supplementary noise exposure data from six other large, commercial construction sites not enrolled in (i.e., "external" to) the current study were used to validate the SR technique. Noise levels associated with the SR survey items we consider here were available from 19 full-shift measurements at one of these six sites.[1] TM levels for each of the six trades and mean L_{EQ} noise levels for all 102 trade tasks reported by subjects in the current study were available from 323 full-shift measurements at the other five external sites. External noise data are important because exposure estimates developed from data collected from our subjects and obtained at the same time as the time-activity and subjective questionnaire responses are optimistic.

We compared the external SR noise levels with the SR noise levels from the three current (internal) study sites to validate the SR technique. For the purposes of evaluating the performance of exposures estimated using the TM, TB, and SR techniques, we combined

information collected on subjects in the current study (e.g., trade, time spent at various tasks, and SR perceptions of exposure) with noise levels measured at the external sites.

Data Analysis

We checked dosimetry data for errors and corrected or removed data using previously published criteria.[66] One-minute noise levels were merged with task information from the activity card and with corresponding survey data in an MS Access (Microsoft, Redmond, Wash.) database and exported for statistical analysis (Intercooled Stata 10.0, Statacorp LP, College Station, Texas, and SPSS for Windows 15.0, SPPS Inc., Chicago, Ill.).

Computation of Noise Metrics—We assessed full-shift dosimetry exposure levels (dBA) using the L_{EQ} metric, computed for individual i on shift j as:

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

where L_{ijk} are the 1-min average L_{EQ} levels measured over $k = 1$ to n_{ij} time periods, and M_{ij} is the total number of minutes measured in the shift.[14] We computed individuals' measured mean exposure levels L_{EQi} as the scaled logarithm of the average of their three measured full-shift exposures:

$$L_{EQi} = 10\log_{10}\left[\frac{1}{3}\sum_{j=1}^{3}\left(10^{L_{EQ_{ij}}/10}\right)\right] \quad (2)$$

We assessed variability in measured noise levels using two variability metrics[66] evaluated in our previous work.[1] Variability at the work shift level was summarized as the average ratio of the L_{EQ} to L_{AVG} levels across the minutes within the shift, as shown in Eq. 3.

$$L_{EQ}/L_{AVG_{ij}} = \frac{1}{M_{ij}}\sum_{k=1}^{n_{ij}}\frac{10^{L_{EQ_{ijk}}/10}}{10^{L_{AVG_{ijk}}/10}} \quad (3)$$

We computed a metric evaluating the impulsiveness of exposure (ratio of L_{MAX} to L_{EQ}) similarly, substituting L_{MAX} for L_{EQ} and L_{EQ} for L_{AVG}. We also computed work shift L_{EQij} standard deviations (SDs) to assess work shift-level variability.

Computation of SR Noise Levels—We assigned each SR perceived intensity survey item response category the mean full-shift L_{EQij} level measured across all individuals reporting that SR response level from the internal sites (e.g., sites enrolled in the current study). For validation purposes, we compared these SR levels with SR levels measured at the single external site from which SR data were available.[1]

We also computed exposure variability metrics (L_{MAX}/L_{EQij} ratio, L_{EQ}/L_{AVGij} ratio, and the SD of L_{EQij} levels) across all measured work shifts for each response category of the SR exposure variability items. We then compared these values with those from the single

external site. As with the SR perceived intensity computations, analysis of the exposure variability metrics occurred only at the work shift level and not averaged across shifts within worker.

Development of SR, TM, and TB Exposure Estimates—We estimated exposures to individual subjects using noise levels measured at the external sites combined with the following: subjects' responses to each of the three SR survey items (Noise frequency, Raise voice, and Percent time in noise), their reported trade (e.g., assignment of a TM level), and a TB approach (e.g., a combination of subjects' reported time at various tasks with task-specific noise levels), as described below.

We computed individual SR exposure estimates, $L_{EQ,SRi}$ for each of the three perceived intensity items as the logarithmic average (Eq. 2) of the mean full-shift L_{EQ} associated with the survey response categories reported for the three measured shifts. We developed trade-based exposure estimates by computing the mean full-shift L_{EQ} for all individuals within each trade at the external sites and applying this mean to each individual in that trade at the internal sites, resulting in a $L_{EQ,TMi}$ estimate for each individual.

We developed TB estimates using "trade/task events" as the basic unit of analysis.[66] We defined trade/task events as the mean exposure during all periods of time during a single work shift that an individual subject reported a single task. We computed mean L_{EQ} noise levels L for each trade/task as the arithmetic average of the L_{EQ} for each trade/task event across all subjects reporting that trade/task at the external sites. We created TB exposure predictions for individual i on shift j using Eq. 4:[66]

$$L_{ij,TB}=10\log_{10}\left[\frac{1}{M_{ij}}\sum_{t=1}^{T}M_{ijt}\,10^{L_t/10}\right] \quad (4)$$

where L_t, the mean noise level for trade/task t, is applied to the period M_{ijt} in which that trade/task was reported by individual i on shift j. We then computed $L_{EQ,TBi}$ estimates for each individual as the logarithmic average of subjects' full-shift TB level for each of the three work shifts, as in Eq. 2.

Comparison of Exposure Measures—We computed descriptive statistics for all survey items and for full-shift dosimetry L_{EQ} levels within and across the three sites. To evaluate reporting patterns over the course of the study, we computed the percentage of subjects reporting the same response to the three SR perceived intensity survey items and the same primary task for all three measured time periods, as well as for two of the three time periods. We evaluated differences in full-shift L_{EQ} levels between sites and by trade and SR perceived intensity survey item response category within and between sites, via one-way analysis of variance (ANOVA).

For validation purposes, we compared differences between internal and external noise levels associated with the SR survey item response categories using two-sample t-tests. We evaluated trends in measured full-shift L_{EQ} levels associated with response categories for each SR survey item using both internal and external noise levels via Cuzik's nonparametric

test for trend.[69] Cuzik's test was also used to evaluate trends in L_{EQ}/L_{AVG} and L_{MAX}/L_{EQ} ratio levels for the SR "Noise variability" and "Noise bursts" survey item response categories, respectively, as well as to assess trends in the SDs of L_{EQij} levels associated with the SR "Task variability" survey item responses.

We evaluated the appropriateness of treating subjects from all sites as a single group for exposure estimation purposes using the results of the χ^2 analyses of demographic and SR survey items. In addition, we ran a mixed effects model (Stata "xtmixed") with full-shift L_{EQ} as the dependent variable and random effects for site and subject to evaluate the relative magnitude of between-site, between-subject, and within-subject SDs.

We evaluated the bias, precision, and accuracy of the exposure estimates from each of the exposure assessment techniques. Bias in estimated exposures was computed as the mean difference between the measured and estimated exposure levels.[56] Precision of the estimates was computed as the SD of the differences between the measured and estimated exposures.[56] Accuracy was computed as[56] $\sqrt{(bias)^2 + (precision)^2}$, also referred to as the root mean squared error. Smaller bias, precision, and accuracy values indicate better estimate performance, e.g., closer proximity of estimated exposures to measured values. For comparison purposes, a 3-dB difference in levels corresponds to a doubling of noise intensity, and a presumed doubling in risk of noise-induced hearing loss.[67] We evaluated absolute agreement among TM, TB, and SR exposure estimates and measured exposure levels using intraclass correlation coefficient ICC3(A,1)[61] (SPSS "Reliability" tool). Finally, we computed nonparametric Spearman correlation coefficients among the exposure estimates from the different techniques to evaluate consistency in exposure rankings, and used scatterplots to visually compare measured and estimated exposures.

RESULTS

Sixty-eight subjects participated in the study. Table I displays demographic and measurement information by site and overall. Nearly all subjects (97%) were male, and the largest fraction was under the age of 30 (41%), had less than 10 years of experience in construction (45%), had a high school education or equivalent (61%), and felt their hearing ability was good (53%). Of the six trades represented, carpenters and laborers were the largest groups (50 and 18%, respectively), with other trades representing 4–12% of subjects. Trade distribution differed significantly by site (χ^2, p < 0.0001); every site had carpenters and laborers, but no site had subjects from all six trades, and two sites had only three of the six trades. The average time between the first and second measured shifts was 73 ± 9 days and between the first and third shifts was 122 ± 16 days.

For comparison purposes, the distribution of trades across 123 workers at the five external sites (data not shown) was comparable to that of the internal sites, with carpenters and laborers again being the largest groups (50% and 20%, respectively), and other trades representing 3–20% of subjects. As with the internal sites, the distribution of trades across the external sites differed significantly (χ^2, p < 0.0001), with only carpenters and laborers present at all five external sites. One of the external sites had five of the six trades, two had four trades, and the other two had three trades.

Measured Exposures and SR Noise Levels

Table II summarizes the 204 measured work shift-level L_{EQij} exposures at the three internal study sites at the work shift level overall and by SR perceived intensity and variability response category. Work shift-level L_{EQij} exposures are also shown for the external study site with SR data[1] for comparison. The mean full-shift L_{EQij} of 88.8 ± 4.3 dBA was above the recommended 85 dBA limit.

All three subjective response items showed increasing noise exposure with higher subjective ratings, and the linear trends in the internal mean full-shift L_{EQij} levels across survey item response categories for the "Percent time in noise" and "Noise frequency" survey items were significantly different from zero (Table II). The external L_{EQij} levels were higher and generally more variable for the response categories from all three SR perceived intensity survey items. The differences between internal and external estimates ranged from −0.7 to −4.0 dBA. For all three SR items, the greatest difference was associated with the lowest reporting category, and the smallest difference with the highest category. We found increasing trends across response categories in L_{MAX}/L_{EQ} ratios for the "Noise bursts" SR item at both the internal and external sites, and in L_{EQ}/L_{AVG} ratios for the "Noise variability" SR item and mean full-shift L_{EQij} SDs for the "Task variability" item at the internal sites. We observed some differences between the internal and external ratio estimates, but none were statistically significantly different.

Appropriateness of Combining Sites for Exposure Estimation—No significant differences in the distribution of reported response categories among sites for any of the three SR survey items were identified by χ^2 analyses (data not shown). Site-specific internal estimates of L_{EQij} levels associated with the SR survey items were also generally comparable. These results, combined with four other findings—the sites were all of comparable size and complexity, there were no significant demographic differences between subjects at these sites, trade mean levels did not differ among sites for the three trades represented at multiple sites, and noise levels associated with the majority of response categories for the SR items also did not differ among sites—suggested that analyzing data combined across all three sites was reasonable. The results of the mixed-model assessment of between-site, between-subject, and within-subject SDs provided further support for this approach. The between-site SD (1.0 dBA) in measured full-shift L_{EQij} levels was smaller than the between-subject (2.5 dBA) SD, which was in turn smaller than the within-subject across worksoft SD (3.4 dBA). Based on these findings, we combined subjects from all three sites for exposure estimation purposes.

Stability of Task and SR Item Reporting and of Full-Shift L_{EQ} Levels over Time

When we assessed subject reporting patterns across the three measured work shifts, no subjects reported the same primary task across all three work shifts, but nearly half (46%) the subjects reported the same primary task for two of the three work shifts. Four, ten, and twenty-seven percent of subjects reported the same responses to the "Noise frequency," "Raise voice," and "Percent time in noise" SR survey items, respectively, across all three work shifts, and at least 90% of subjects reported the same responses on two of three work shifts for each of the items. Mean full-shift L_{EQ} levels did not differ among the three

measured work shifts for each subject, either within individual sites, or overall, indicating no time trend in exposure levels (data not shown).

Performance of Exposure Estimates Compared with Measured Exposure Levels

Table III shows the mean and SD of subjects' measured levels, as well as the mean, SD, bias, precision, accuracy, and absolute agreement (ICC) of the estimated and measured exposures. The measured mean L_{EQi} for the 68 subjects was 89.6 ± 3.4 dBA, well in excess of the 85 dBA exposure limit recommended by NIOSH.[67] Neither the TM nor SR group mean approaches produced unbiased exposure estimates. The bias in the SR estimates was an order of magnitude larger than that of the TM and TB estimates, most likely driven by site differences and the availability of only one site for SR noise level estimates. Overall, the TB estimates showed the best precision, accuracy, and absolute agreement with measured exposures, followed by the SR "Percent time in noise" estimates. The SDs of the SR estimates were in almost all cases smaller than those of the TM and TB measures due to the small number of possible exposure values (three) for these items and to the fact that most SR responses were in the middle response category. The TM and SR "Raise voice" and "Noise frequency" estimates generally showed the worst precision, accuracy, and absolute agreement with measured exposures.

We previously explored several factors that could conceivably influence SR evaluations of perceived noise intensity and did not find any relationships between these factors and perceptions of intensity.[1] To confirm these earlier findings, we evaluated the effects of subjects' perceived hearing sensitivity (good vs. less than good) and reported use of hearing protection (50% of time spent in high noise vs. <50% time) on the performance of the SR estimates. SR exposure estimates for subjects with good perceived hearing sensitivity and subjects who reported using hearing protectors <50% of the time in high noise had slightly better bias, precision, and accuracy than those who did not. The differences were very small (0.2 dBA or less) and not statistically significant, so these factors were not incorporated into further analyses.

Correlations among the estimated and measured exposures for the three approaches were all positive but were generally low, ranging from weak to moderate (Spearman r 0.09 to 0.64, Figure 1). Inter-estimate correlations were also low (Spearman r 0.18 to 0.47). The TM estimates showed poor correlation with measured exposures and with all other estimates except TB. The generally low correlations among the estimates from the different techniques suggest that they rank exposures differently.

To better understand the unexpectedly poor performance of the TM exposure estimates, we developed a separate set of exposure estimates using the internal TM noise levels and, for comparison purposes, the internal TB and SR "Percent time in noise" noise levels as well. These internal estimates were then compared with the estimates described in Table III. Figure 2 shows scatterplots of the internal and external estimates and also displays for each plot the R^2 value from a linear regression model using the internal estimates as the dependent variable and the external estimates as the independent variable.

The overall mean of the internal and external full-shift measurements differed by only 1 dBA (p = 0.02, data not shown). However, the rank ordering of trades differed between the internal and external sites, and substantial differences between TM estimates based on internal and external noise levels (−4.2 to +3.6 dBA, data not shown) were found for three of the six trades evaluated. The effect of the different TM rank orderings is demonstrated in the poor association in the TM scatterplot in Figure 2.

The range of differences between the internal and external TB estimates was similar, but there were notable differences between the underlying trade/task levels from the internal and external sites. Although the mean difference between internal and external trade/task levels was small (0.8 ±4.9 dBA, data not shown), 40% of trade/tasks had differences between internal and external levels that equaled or exceeded an absolute difference of 3 dBA, 23% had differences of 5 dBA or greater, and four tasks showed very large differences of 10 dBA or more. The squared correlation of the internal and external TB estimates was somewhat higher than that of the TM estimates, due at least in part to the larger range of TB estimates. Thus, the striking effect of rank ordering in the TM estimates is less apparent in the TB data at face value, but certainly there are differences in internal and external trade/task levels and in ranks across the two sets of TB estimates. In contrast to the poor to moderate linear association in the TM and TB scatterplots, the strong association for the SR "Percent time in noise" measure reflects the consistent rank ordering of noise levels associated with the response categories, as well as the consistent inter-category difference in noise levels, for both internal and external levels. This strong association was also seen for the SR "Noise frequency" and "Raise voice" measures (data not shown).

DISCUSSION

We evaluated the performance of noise exposures estimated using three different exposure assessment techniques—trade mean, task-based, and subjective rating—against measured exposures over a 4-month period. Our results demonstrate that perceptions of noise exposure intensity reported by workers provide information that can be used to distinguish and estimate exposure levels. Subjects' perceptions of increasing noise intensity were associated with increasing measured L_{EQij} levels for three SR survey items. Estimates developed from the "Percent time in noise" item consistently had the best performance among the three SR items. Taken together, these results suggest the SR technique developed by Neitzel et al.[1] is a valid tool for use in assessing noise on commercial construction sites, though this technique cannot necessarily be applied to other sectors of construction or other industries without additional validation. Our results also support the previous finding[1] that perceptions of noise exposure variability may be used to divide subjects into groups with different variability in exposure and exposure to impulsive noise.

We found that the TB exposure assessment technique had the best agreement with measured exposure levels. In assessing agreement between estimated and measured mean levels, the uncertainty in the measured exposures, which are based on only three measurements per individual, must be considered. For epidemiologic purposes, larger variability in estimated exposures, as measured by the SD of the exposure estimates, is desirable, as long as the

exposure estimates capture variability in the true exposure, rather than simply representing random error.

The TB technique consistently produced the largest SDs. The estimated accuracy of the estimates developed using the three techniques was 3–4 dBA. The TB estimates showed the best accuracy and absolute agreement with measured levels. SR estimates showed absolute agreement with measured levels and greater accuracy than did the TM estimates. The associations among estimates developed using the three techniques were poor, suggesting that they measured different aspects of exposure.

SR techniques can be useful in assessing exposures for which human senses provide good detection and discrimination. A number of studies have documented moderate to high correlations between measured exposure levels and SR perceptions of exposure to dust,[51,52] noise,[55] and ergonomic hazards.[70] However, caution must be exercised in the incorporation of SR assessment of noise in the workplace. Hearing loss prevention programs commonly use a guideline that suggests if workers must raise their voice to speak to someone at arm's length, noise levels are likely over 85 or 90 dBA.[65,71] The "Raise voice" SR item evaluated here was modeled after this guideline, but the performance of the item was poor, suggesting that this guideline may not relate well to workers' actual exposures and may not be appropriate for use in hearing loss prevention efforts.

Although SR assessment has traditionally been considered to be less accurate than other approaches,[50,72] in our dataset the "Percent time in noise" SR estimates showed greater accuracy and absolute agreement with measured levels than did TM estimates. The rank ordering of SR levels was consistent across sites, while TM levels showed different rank ordering across sites, violating the assumption of consistent rank ordering of exposure groups, which is fundamental to group-based assessment techniques. There are several possible sources of this difference between techniques that may have resulted from our study design and sample. Our sample sizes were small, and the external sites differed between the SR and TM in both number and specific sites. None of our three internal or five external sites had all six trades represented, and the absence of certain trades at individual sites could have affected the rank ordering of our internal estimates.

There are several caveats to the SR results presented here. First, SR estimates developed over long exposure periods (e.g., many months to years) will likely be less accurate than those seen here due to the effects of recall bias and misreporting. Second, the SR results were systematically higher at the single external site (described in detail in earlier paper[1]) vs. the internal sites from the current study. The single site used to provide external SR data was unusually noisy, with substantially higher full-shift L_{EQij}, trade mean, and TB levels than those found at all other internal and external sites. Thus the internal SR noise levels from the current study should be considered more robust and generalizable, and the strong association between the internal and external SR noise levels should be interpreted with caution.

Third, an interesting trend was apparent in noise levels associated with SR item response categories in the current study vs. the external site, with differences across all three items

being largest for the lowest category, and smallest in the highest category. This suggests that inter-worker variability in perceptions of lower levels differs more than perceptions of high levels; in other words, workers may not share common perceptions of what a "low" noise level is, but perceptions are more uniform concerning "high" noise.

We found that TB estimates were somewhat more accurate than SR or TM estimates although all our estimates of accuracy were fairly similar (with a range of 3.1 to 4.1 dB). Results of previous studies suggest that the performance of TB techniques may vary by industry and exposure. TB estimates of noise exposure based on tasks reported by industrial workers[29] and construction workers[32,48,66] had similar bias and worse precision than seen here. TB noise estimates for industrial workers showed somewhat better absolute agreement with measured levels than seen here,[29] with the best agreement associated with jobs featuring low exposure variability, mobility, and complexity[29]—jobs for which TB assessment is unnecessarily complicated.

The precision of our TB noise exposure estimates was better than those seen by Seixas and colleagues,[32,48,66] due at least in part to use of a richer set of trade/task definitions (and therefore reduced task misclassification) here. Seixas et al.[66] also found that TB estimates of noise exposure had lower bias and better precision than did TM estimates. Conversely, two studies of ergonomic exposures found that TB assessment provided only minimal gains in precision over estimates based on job title[27] and had equivalent or worse bias.[33] As with SR estimates, long-term TB estimates will likely be less accurate than those developed here due to increased misreporting of the types and durations of tasks performed.

Although trade has been used as a predictor of occupational injuries[73–75] and fatalities,[76–78] silica and lead exposures,[22,46,79,80] chemical exposures,[21] and ergonomic hazards,[81–83] we found trade-based estimates of noise exposure to be less accurate than TB and SR estimates. There are at least four potential reasons for this finding. First, our dataset was relatively small and, in particular, had only three sites with workers that were not balanced with respect to trade, possibly leading to less accurate TM estimates. Second, many of the exposures previously assessed using trade, for example, work at height,[84] and slips and trips,[85] are highly localized and result in minimal exposure to nearby workers, whereas noise from a single high-level source often propagates widely. High noise activities can therefore result in substantial exposures to nearby and distant workers in a variety of trades, a likely explanation for the very small range (5 dBA) between TM levels in our study.

Third, although TM approaches are thought to produce unbiased estimates of group-mean exposures, in the case of nonlinear exposures such as noise, TM techniques may result in equal or greater bias than other techniques. Finally, trade may simply be too broad a grouping strategy to assess construction noise, since workers within a trade may perform vastly different work activities, there is large variability in noise across construction work sites and over time within work sites, and some activities overlap across trades.

The current study had a number of limitations that may reduce the generalizability of our findings. The first limitation is that treatment of three measured full-shift exposure levels

over a 4-month period as a worker's "true" exposure level is unrealistic. The measured average exposure levels used as a gold standard exposure for subjects in this study are in fact a highly uncertain estimate of the 4-month long-term average exposure. Better estimates of subjects' true exposures would have required many more repeated measurements over the study period. Nevertheless, three measurements per subject is nevertheless greater than what is typically available in occupational settings.[12]

The second limitation is the small number of sites and subjects included in this study and the fact that different trades were represented at different sites. Although our ANOVA analyses indicated that data from different sites could reasonably be combined, these analyses were underpowered with respect to trade, and imbalances in the distribution of trades may have resulted in between-site differences that affected our results. In particular, imbalances in the distribution of trades across our three sites probably reduced the performance of the TM technique we evaluated and, to a lesser extent, the TB technique, which was based on trade/tasks. The most prevalent trades at the internal and external sites were similar, i.e., the sites were dominated by the two most common trades in the industry,[86] and had far fewer workers from specialty trades, the distribution of which depends largely on the stage of construction at a particular site.

To further explore the effects of trade distribution on TM estimates, we conducted two additional analyses: one was based on TM levels from a larger external dataset presented by Neitzel et al.,[18] and the other created site-adjusted TM levels for each of the internal sites. Use of TM levels from the larger dataset provided negligible improvements overall in the accuracy of the TM estimates for all trades. Site adjustment also provided negligible improvements overall in accuracy for the prevalent trades represented across all sites; however, greater improvements (1–2 dBA) were seen for trades that were poorly distributed across sites. This suggests that site-specific effects may have important effects on TM and TB exposure estimates when trade-specific data are sparse or when imbalances exist in the distribution of trades across different sites, a finding that highlights the benefits of site-specific measurement data in epidemiologic studies. These imbalances should not affect the results of the SR technique.

The third limitation is that the information on which exposure estimates were based—trade, reported tasks, and SR item responses—was not on the same time scale. Trade was invariant within subject over the three measured work shifts. The three SR perceived intensity survey items completed on each of the three measured work shifts inquired about "current" exposure, with "current" not assigned a specific time period. The full-shift TB exposure levels estimated for each worker were based on reported time-at-task in three work shifts. Comparison of the TB and SR item estimates assumes that a TB estimate from a single work shift represents an accurate "snapshot" of activities during the period represented by subjects' survey item responses. The fact that nearly half the subjects reported the same primary task on two of their three measured shifts does lend some support to the notion that single-day TB assessments may be representative of a period of time surrounding the measured shift. However, perceived exposures reported via the SR survey items likely represent different periods of exposure than do the trade or TB information, introducing an additional source of potential error.

The final limitation of this study is that estimation of TB exposures at the shift level, as was done here, is more optimistic —that is, more likely to result in good agreement with measured levels than would be expected for a longer-term assessment, since focusing on single work shifts largely ignores inter-work shift and inter-job variability in tasks and TB noise levels.

No comparison of an assessment of accuracy and absolute agreement (i.e., ICC) analyses appears to have been presented in the literature to date. Our study found that the information provided by these two measures of performance was consistent—that is, the estimates with the best accuracy always showed the best agreement, and the estimates with the worst accuracy always had the worst agreement. This suggests that in future comparisons of exposure assessment techniques it may be reasonable to focus on one measure of performance or the other, rather than assessing both simultaneously, although this judgment will depend on the goals of the exposure assessment study.

CONCLUSIONS

Based on the available data, and given the sampling limitations described above, the best performing exposure assessment technique in this study was one that used a task-based approach, followed by a subjective rating technique, and finally, a technique based on use of trade mean exposure levels. Each of the techniques evaluated here displayed substantial (3–4 dBA) measurement error. Methods to combine estimates made using different techniques have the potential to reduce the measurement error associated with each of these individual assessment techniques, and these methods are explored in the second part of our study effort.[2]

Acknowledgments

References

1. Neitzel R, Daniell W, Sheppard L, Davies H, Seixas N. Comparison of perceived and quantitative measures of occupational noise exposure. Ann Occup Hyg. 2009; 53(1):41–54. [PubMed: 18984805]

2. Neitzel R, Daniell W, Sheppard L, Davies H, Seixas N. Improving exposure assessment by combining exposure information. Ann Occup Hyg. [In press].

3. Nieuwenhuijsen MJ. Exposure assessment in occupational epidemiology: Measuring present exposures with an example of a study of occupational asthma. Int Arch Occup Environ Health. 1997; 70:295–308. [PubMed: 9352332]

4. Nieuwenhuijsen M, Paustenbach D, Duarte-Davidson R. New developments in exposure assessment: The impact on the practice of health risk assessment and epidemiological studies. Environ Int. 2006; 32(8):996–1009. [PubMed: 16875734]

5. Drummond I, Murray N, Armstrong T, Schnatter AR, Lewis RJ. Exposure assessment methods for a study of mortality and cancer morbidity in relation to specific petroleum industry exposures. J Occup Environ Hyg. 2006; 3:513–520. [PubMed: 16908452]

6. Hawkins, NC.; Norwood, SK.; Rock, JC. A Strategy for Occupational Exposure Assessment. Akron, Ohio: American Industrial Hygiene Association; 1991. p. 2-26.

7. Rappaport SM, Kromhout H, Symanski E. Variation of exposure between workers in homogeneous exposure groups. Am Ind Hyg Assoc J. 1993; 54:654–662. [PubMed: 8256689]

8. Kromhout H, Oostendorp Y, Heederik D, Boleij JS. Agreement between qualitative exposure estimates and quantitative exposure measurements. Am J Ind Med. 1987; 12(5):551–562. [PubMed: 3687951]

9. Corn M, Esmen NA. Workplace exposure zones for classification of employee exposures to physical and chemical agents. Am Ind Hyg Assoc J. 1979; 40:47–57. [PubMed: 484450]

10. Smith TJ, Hammond K, Hallock M, Woskie SR. Exposure assessment for epidemiology: Characteristics of exposure. Appl Occup Environ Hyg. 1991; 6:441–447.

11. Rappaport SM. Selection of the measures of exposure for epidemiology studies. Appl Occup Environ Hyg. 1991; 6:448–457.

12. Loomis D, Kromhout H. Exposure variability: Concepts and applications in occupational epidemiology. Am J Ind Med. 2004; 45(1): 113–122. [PubMed: 14691975]

13. Seixas N, Sheppard L. Maximizing accuracy and precision using individual and grouped exposure assessments. Scand J Work Environ Health. 1996; 22:94–101. [PubMed: 8738886]

14. Earshen, J. Chapter 3: Sound measurement: Instrumentation and noise descriptors. In: Berger, E.; Royster, L.; Royster, J.; Driscoll, D.; Layne, M., editors. The Noise Manual. 5. Fairfax, Va: American Industrial Hygiene Association; 2000. p. 41-100.

15. Kromhout H, Heederik D. Occupational epidemiology in the rubber industry: Implications of exposure variability. Am J Ind Med. 1995; 27(2):171–185. [PubMed: 7755008]

16. Verma DK, Cheng WK, Shaw DS, et al. A simultaneous job- and task-based exposure evaluation of petroleum tanker drivers to benzene and total hydrocarbons. J Occup Environ Hyg. 2004; 1:725–737. [PubMed: 15673093]

17. Neitzel R, Seixas NS, Camp J, Yost M. An assessment of occupational noise exposures in four construction trades. Am Ind Hyg Assoc J. 1999; 60:807–817. [PubMed: 10635548]

18. Neitzel R, Seixas N. The effectiveness of hearing protection among construction workers. J Occup Environ Hyg. 2005; 2:227–238. [PubMed: 15788384]

19. Legris M, Poulin P. Noise exposure profile among heavy equipment operators, associated laborers, and crane operators. Am Ind Hyg Assoc J. 1998; 59:774–778. [PubMed: 9830084]

20. Flanagan ME, Seixas N, Becker P, Takacs B, Camp J. Silica exposure on construction sites: Results of an exposure monitoring data compilation project. J Occup Environ Hyg. 2006; 3:144–152. [PubMed: 16464818]

21. Verma DK, Kurtz LA, Sahai D, Finkelstein MM. Current chemical exposures among Ontario construction workers. Appl Occup Environ Hyg. 2003; 18:1031–1047. [PubMed: 14612300]

22. Rappaport SM, Goldberg M, Susi P, Herrick RF. Excessive exposure to silica in the US construction industry. Ann Occup Hyg. 2003; 47(2):111–122. [PubMed: 12581996]

23. Kromhout H, Symanski E, Rappaport SM. A comprehensive evaluation of within- and between-worker components of occupational exposure to chemical agents. Ann Occup Hyg. 1993; 37(3): 253–270. [PubMed: 8346874]

24. Symanski E, Maberti S, Chan W. A meta-analytic approach for characterizing the within-worker and between-worker sources of variation in occupational exposure. Ann Occup Hyg. 2006; 50(4): 343–357. [PubMed: 16513810]

25. Tielemans E, Kupper L, Kromhout H, et al. Individual based and group based exposure assessment: Some equations to evaluate different strategies. Ann Occup Hyg. 1998; 42:115–119. [PubMed: 9559571]

26. Susi P, Schneider S. Database needs for a task-based exposure assessment model for construction. Appl Occup Environ Hyg. 1995; 10:394–399.

27. Svendsen SW, Mathiassen SE, Bonde JP. Task based exposure assessment in ergonomic epidemiology: A study of upper arm elevation in the jobs of machinists, car mechanics, and house painters. Occup Environ Med. 2005; 62(1):18–27. [PubMed: 15613604]

28. Neitzel R, Yost M. Task-based assessment of occupational vibration and noise exposures in forestry workers. AIHAJ. 2002; 63:617–627.

29. Virji MA, Woskie SR, Waters M, et al. Agreement between task-based estimates of the full-shift noise exposure and the full-shift noise dosimetry. Ann Occup Hyg. 2009; 53(3):201–214. [PubMed: 19282390]

30. Hager L. Sound exposure profiling: A noise monitoring alternative. Am Ind Hyg Assoc J. 1998; 59:414–418.

31. Nicas M, Spear RC. A task-based statistical model of a worker's exposure distribution: Part I—Description of the model. Am Ind Hyg Assoc J. 1993; 54:211–220. [PubMed: 8498356]

32. Reeb-Whitaker CK, Seixas NS, Sheppard L, Neitzel R. Accuracy of task recall for epidemiological exposure assessment to construction noise. Occup Environ Med. 2004; 61(2):135–142. [PubMed: 14739379]

33. Mathiassen SE, Nordander C, Svendsen SW, Wellman HM, Dempsey PG. Task-based estimation of mechanical job exposure in occupational groups. Scand J Work Environ Health. 2005; 31:138–151. [PubMed: 15864908]

34. Virji MA, Woskie SR, Pepper LD. Task-based lead exposures and work site characteristics of bridge surface preparation and painting contractors. J Occup Environ Hyg. 2009; 6:99–112. [PubMed: 19065390]

35. Greenspan CA, Moure-Eraso R, Wegman D, Oliver LC. Occupational hygiene characterization of a highway construction project: A pilot study. Appl Occup Environ Hyg. 1995; 10:50–58.

36. Woskie SR, Bello D, Gore RJ, et al. Comparison of task-based exposure metrics for an epidemiologic study of isocyanate inhalation exposures among autobody shop workers. J Occup Environ Hyg. 2008; 5:588–598. [PubMed: 18615291]

37. Achutan C, Tubbs RL. A task-based assessment of noise levels at a swine confinement. J Agromed. 2007; 12(2):55–65.

38. Humann MJ, Donham KJ, Jones ML, Achutan C, Smith BJ. Occupational noise exposure assessment in intensive swine farrowing systems: Dosimetry, octave band, and specific task analysis. J Agromed. 2005; 10(1):23–37.

39. Seixas NS, Ren K, Neitzel R, Camp J, Yost M. Noise exposure among construction electricians. Am Ind Hyg Assoc J. 2001; 62:615–621.

40. Kerr M, Brosseau L, Johnson C. Noise levels of selected construction tasks. Am Ind Hyg Assoc J. 2002; 63:334–339.

41. Susi P, Goldberg M, Barnes P, Stafford E. The use of a task-based exposure assessment model (T-BEAM) for assessment of metal fume exposures during welding and thermal cutting. Appl Occup Environ Hyg. 2000; 15:26–38. [PubMed: 10660986]

42. Buchholz B, Paquet V, Wellman H, Forde M. Quantification of ergonomic hazards for ironworkers performing concrete reinforcement tasks during heavy highway construction. AIHAJ. 2003; 64:243–250.

43. Unge J, Hansson GA, Ohlsson K, et al. Validity of self-assessed reports of occurrence and duration of occupational tasks. Ergonomics. 2005; 48:12–24. [PubMed: 15764303]

44. Burstyn I. Measurement error and model specification in determining how duration of tasks affects level of occupational exposure. Ann Occup Hyg. 2009; 53(3):265–270. [PubMed: 19188265]

45. Methner M, McKernan J, Dennison J. Task-based exposure assessment of hazards associated with new residential construction. Appl Occup Environ Hyg. 2000; 15:811–819. [PubMed: 11062925]

46. Johnson JC, Reynolds SJ, Fuortes LJ, Clarke WR. Lead exposure among workers renovating a previously de-leaded bridge: Comparison of trades, work tasks. AIHAJ. 2000; 61:815–819. [PubMed: 11192213]

47. Goldberg M, Levin SM, Doucette JT, Griffin G. A task-based approach to assessing lead exposure among iron workers engaged in bridge rehabilitation. Am J Ind Med. 1997; 31(3):310–318. [PubMed: 9055954]

48. Seixas NS, Sheppard L, Neitzel R. Comparison of task-based estimates with full-shift measurements of noise exposure. AIHAJ. 2003; 64:823–829.

49. Neitzel R, Seixas N, Olson J, Daniell W, Goldman B. Nonoccupational noise: Exposures associated with routine activities. J Acoust Soc Am. 2004; 115(1):237–245. [PubMed: 14759016]

50. Teschke K, Olshan AF, Daniels JL, et al. Occupational exposure assessment in case-control studies: Opportunities for improvement. Occup Environ Med. 2002; 59(9):575–593. Discussion 594. [PubMed: 12205230]

51. Fonn S, Groeneveld HT, deBeer M, Becklake MR. Relationship of respiratory health status to grain dust in a Witwatersrand grain mill: Comparison of workers' exposure assessments with industrial hygiene survey findings. Am J Ind Med. 1993; 24(4):401–411. [PubMed: 8250060]

52. Nieuwenhuijsen MJ, Noderer KS, Schenker MB. The relation between subjective dust exposure estimates and quantitative dust exposure measurements in California agriculture. Am J Ind Med. 1997; 32(4):355–363. [PubMed: 9258389]

53. Winnemuller LL, Spielholz PO, Daniell WE, Kaufman JD. Comparison of ergonomist, supervisor, and worker assessments of work-related musculoskeletal risk factors. J Occup Environ Hyg. 2004; 1:414–422. [PubMed: 15238331]

54. Dimov M, Bhattacharya A, Lemasters G, Atterbury M, Greathouse L, Ollila-Glenn N. Exertion and body discomfort perceived symptoms associated with carpentry tasks: An on-site evaluation. AIHAJ. 2000; 61:685–691. [PubMed: 11071420]

55. Ising H, Babisch W, Kruppa B, Lindthammer A, Wiens D. Subjective work noise: A major risk factor in myocardial infarction. Soz Praventivmed. 1997; 42(4):216–222. [PubMed: 9297760]

56. Hornung RW. Statistical evaluation of exposure assessment strategies. Appl Occup Environ Hyg. 1991; 6:516–520.

57. White SA, Van Den Broek NR. Methods for assessing reliability and validity for a measurement tool: A case study and critique using the WHO haemoglobin colour scale. Stat Med. 2004; 23(10): 1603–1619. [PubMed: 15122740]

58. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986; 1:307–310. [PubMed: 2868172]

59. Altman DG, Bland JM. Measurement in medicine: The analysis of method comparison studies. Statistician. 1983; 32:307–317.

60. Bland JM, Altman DG. Applying the right statistics: Analyses of measurement studies. Ultrasound Obstet Gynecol. 2003; 22(1):85–93. [PubMed: 12858311]

61. McGraw KO, Wong SP. Forming inference about some intraclass correlation coefficients. Psychol Meth. 1996; 1(1):30–46.

62. Carrasco JL, Jover L. Estimating the generalized concordance correlation coefficient through variance components. Biometrics. 2003; 59:849–858. [PubMed: 14969463]

63. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull. 1979; 86(2):420–428. [PubMed: 18839484]

64. Trabeau M, Neitzel R, Meischke H, Daniell WE, Seixas NS, et al. A comparison of "Train-the-Trainer" and expert training modalities for hearing protection use in construction. Am J Ind Med. 2008; 51(2):130–137. [PubMed: 18067179]

65. Neitzel R, Meischke H, Daniell WE, Trabeau M, Somers S, Seixas NS. Development and pilot test of hearing conservation training for construction workers. Am J Ind Med. 2008; 51(2):120–129. [PubMed: 18067178]

66. Seixas N, Neitzel R, Sheppard L, Goldman B. Alternative metrics for noise exposure among construction workers. Ann Occup Hyg. 2005; 49(6): 493–502. [PubMed: 15797894]

67. National Institute for Occupational Safety and Health (NIOSH). Occupational Noise Exposure. Cincinnati, Ohio: US Department of Health and Human Services, Centers for Disease Control and Prevention, NIOSH; 1998.

68. Occupational Safety and Health Administration (OSHA). Occupational Noise Exposure; Hearing Conservation Amendment; Final Rule. Code of Federal Regulations Title. 1983; 29(Part 1910.95)

69. Cuzick J. A Wilcoxon-type test for trend. Stat Med. 1985; 4(1):87–90. [PubMed: 3992076]

70. Buchholz B, Park JS, Gold JE, Punnett L. Subjective ratings of upper extremity exposures: Inter-method agreement with direct measurement of exposures. Ergonomics. 2008; 51(7):1064–1077. [PubMed: 18568965]

71. Daniell WE, Swan SS, McDaniel MM, Camp JE, Cohen MA, Stebbins JG. Noise exposure and hearing loss prevention programmes after 20 years of regulations in the United States. Occup Environ Med. 2006; 63(5):343–351. [PubMed: 16551755]

72. Stewart P, Stenzl M. Exposure assessment in the occupational setting. Appl Occup Environ Hyg. 2000; 15(5):435–444. [PubMed: 10808266]

73. Welch LS, Hunting KL, Murawski JA. Occupational injuries among construction workers treated in a major metropolitan emergency department in the United States. Scand J Work Environ Health. 2005; 31(Suppl 2):11–21. [PubMed: 16366004]

74. Dong X. Long work hours, work scheduling and work-related injuries among construction workers in the United States. Scand J Work Environ Health. 2005; 31(5):329–335. [PubMed: 16273958]

75. Horwitz IB, McCall BP. Disabling and fatal occupational claim rates, risks, and costs in the Oregon construction industry 1990–1997. J Occup Environ Hyg. 2004; 1:688–698. [PubMed: 15631060]

76. Buskin SE, Paulozzi LJ. Fatal injuries in the construction industry in Washington State. Am J Ind Med. 1987; 11(4):453–460. [PubMed: 3578298]

77. Dong X, Platner JW. Occupational fatalities of Hispanic construction workers from 1992 to 2000. Am J Ind Med. 2004; 45(1):45–54. [PubMed: 14691968]

78. McCann M, Hunting KL, Murawski J, Chowdhury R, Welch L. Causes of electrical deaths and injuries among construction workers. Am J Ind Med. 2003; 43(4):398–406. [PubMed: 12645095]

79. Reynolds SJ, Seem R, Fourtes LJ, et al. Prevalence of elevated blood leads and exposure to lead in construction trades in Iowa and Illinois. Am J Ind Med. 1999; 36(2):307–316. [PubMed: 10398939]

80. Dement JM, Welch L, Bingham E, et al. Surveillance of respiratory diseases among construction and trade workers at Department of Energy nuclear sites. Am J Ind Med. 2003; 43(6):559–573. [PubMed: 12768606]

81. Tak S, Paquet V, Woskie S, Buchholz B, Punnett L. Variability in risk factors for knee injury in construction. J Occup Environ Hyg. 2009; 6:113–120. [PubMed: 19085603]

82. Paquet V, Punnett L, Woskie S, Buchholz B. Reliable exposure assessment strategies for physical ergonomics stressors in construction and other non-routinized work. Ergonomics. 2005; 48(9): 1200–1219. [PubMed: 16251156]

83. Holmstrom E, Engholm G. Musculoskeletal disorders in relation to age and occupation in Swedish construction workers. Am J Ind Med. 2003; 44(4):377–384. [PubMed: 14502765]

84. Kaskutas V, Dale AM, Nolan J, Patterson D, Lipscomb HJ, Evanoff B. Fall hazard control observed on residential construction sites. Am J Ind Med. 2009; 52(6):491–499. [PubMed: 19363784]

85. Lipscomb HJ, Glazner JE, Bondy J, Guarini K, Lezotte D. Injuries from slips and trips in construction. Appl Ergon. 2006; 37(3):267–274. [PubMed: 16212931]

86. [Accessed July 12, 2010] Career Guide to Industries: Occupations in the Industry. [Online] Available at http://www.bls.gov/oco/cg/cgs003.htm#related
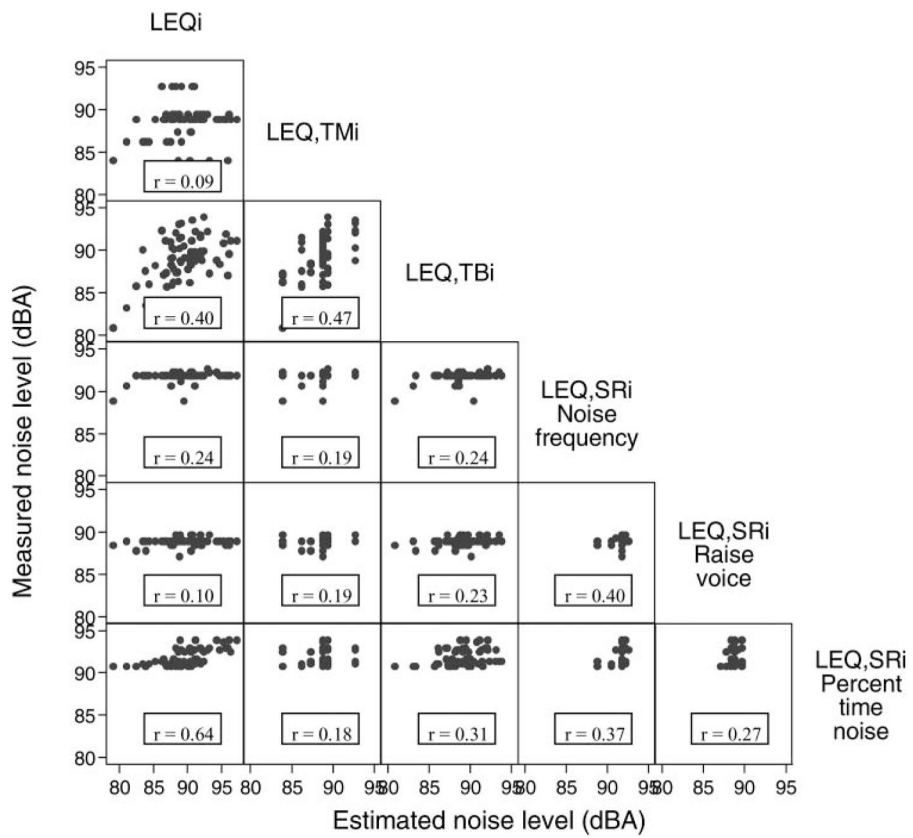
FIGURE 1.
Measured mean exposure level and individual exposure estimates (dBA) for five measures
(n = 68 subjects)
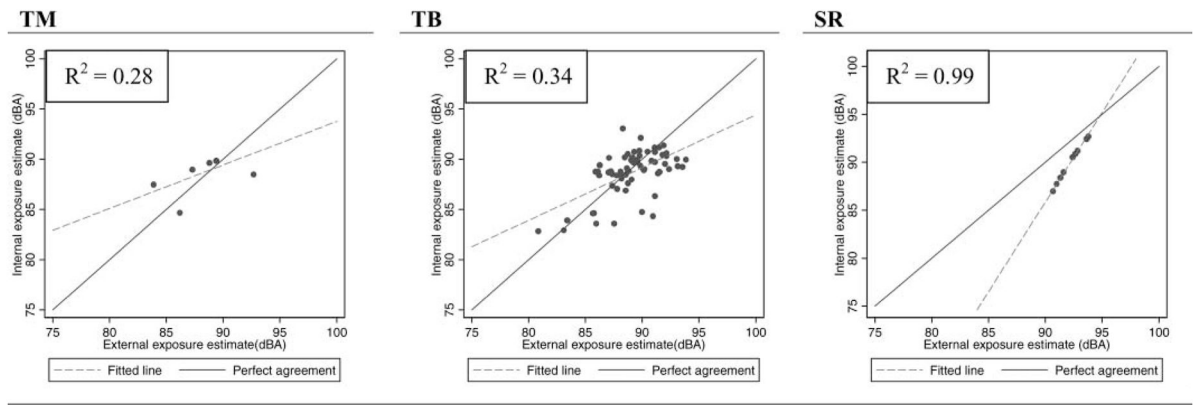
FIGURE 2.

Exposure estimates (dBA) at worker level based on internal (y axis) and external (x axis) noise levels for the TM, TB, and "Percent time in noise" SR measure (n = 68 subjects)

TABLE I

Internal Site Summary Statistics: Demographic and Measurement Information

Variable	Overall		Site 1		Site 2		Site 3	
	N	(%)	N	(%)	N	(%)	N	(%)
Total subjects	68	100	21	100	16	100	31	100
Gender–male	66	97	21	100	16	100	29	94
Age								
<30	28	41	8	38	8	50	12	39
30–40	20	29	6	29	5	31	9	29
>40	20	29	7	33	3	19	10	32
Education								
< High School (HS)	8	12	1	5	4	25	3	10
HS/GED	41	61	12	57	8	53	22	72
> HS	18	26	8	38	4	23	6	18
Trade[A]								
Carpenter	34	50	8	38	13	81	13	42
Electrician	5	7	—	—	—	—	5	16
Ironworker	6	9	—	—	1	6	5	16
Laborer	12	18	5	7	2	13	5	16
Operating engineer	3	4	—	—	—	—	3	4
Plumber/pipefitter	8	12	8	38	—	—	—	—
Years in construction								
<10 years	31	45	10	46	10	63	10	32
10–20 years	20	30	6	30	4	25	11	35
>20 years	17	25	5	24	2	13	10	32
Hearing ability								
Good	36	53	10	49	10	63	16	51
Fair	28	41	10	46	4	25	14	46
Poor	4	6	1	5	2	13	1	3
Tinnitus								
Never	14	21	5	22	3	19	7	21

Variable	Overall		Site 1		Site 2		Site 3	
	N	(%)	N	(%)	N	(%)	N	(%)
Rarely	28	41	7	31	7	44	15	49
Sometimes	18	27	8	39	3	19	7	21
Frequently	5	8	1	6	2	13	2	6
Always	2	3	1	3	1	9	1	3
Family hearing loss: Yes	14	21	5	24	4	22	7	21
	Days		Days		Days		Days	
Time between measurements	Mean	SD	Mean	SD	Mean	SD	Mean	SD
First and second	73	9	75	4	70	3	74	7
First and third	122	16	125	6	120	2	123	6

[A] Statistically significant differences across sites (χ^2, p 0.05).

TABLE II

Noise Exposures at the Work Shift Level Overall and by Trade and Subjective (SR) Item Responses for Internal (i.e., Study) and External (i.e., Non-Study) Sites

Variable	Internal (three sites)			External (one site)[A]			
	n	**Mean L_{EQij} (dBA)**	**SD L_{EQ} (dBA)**	**n**	**Mean L_{EQij} (dBA)**	**SD L_{EQ} (dBA)**	**Diff. (dBA)**
Measured exposure	204	88.8	4.3	19	90.2	3.0	−1.4
SR Perceived Intensity Items							
Noise frequency	204			19			
Never	8	81.7[B]	6.1	3	85.7	4.1	−4.0
Sometimes	185	88.9[B]	4.0	15	91.8	3.8	−2.9
Always	11	91.8[B]	4.0	1	92.9	—	1.1
Raise voice	204			19			
Never	21	87.0	6.2	3	90.1	7.5	−3.1
Sometimes	174	88.9	4.1	13	91.6	3.7	−2.7
Always	9	90.9	3.5	3	93.0	2.8	−2.1
Percent time in noise	204			19			
Speaking voice	97	87.1[B]	4.2	7	90.7	5.6	−3.6
Vacuum-chainsaw	78	89.0[B]	2.9	9	91.6	3.0	−2.7
Siren	29	93.9[B]	4.1	3	94.6	2.2	−0.7

Variable	**n[B]**	**Mean L_{MAX}/L_{EQij}**	**SD L_{MAX}/L_{EQij}**	**n**	**Mean L_{MAX}/L_{EQij}**	**SD L_{MAX}/L_{EQij}**	**Diff. (dBA)**
SR Perceived Variability Items							
Noise bursts	204						
Never	11	52.0[C]	29.7	—	—	—	
Sometimes	187	60.4[C]	18.3	13	56.6[C]	8.2	3.8
Always	6	135.8[C]	36.7	6	64.7[C]	5.6	71.1

Variable	**n**	**Mean L_{EQ}/L_{AVGij}**	**SD L_{EQ}/L_{AVGij}**	**n**	**Mean L_{EQ}/L_{AVGij}**	**SD L_{EQ}/L_{AVGij}**	**Diff.**
Noise variability	204						
Usually steady	68	2.2[C]	0.7	6	2.0	0.1	0.2
Usually variable	136	2.7[C]	1.2	13	2.1	0.3	0.6

	Internal (three sites)			External (one site)[A]			
Variable	n	Mean L_{EQij} (dBA)	SD L_{EQ} (dBA)	n	Mean L_{EQij} (dBA)	SD L_{EQ} (dBA)	Diff. (dBA)
	n	Mean L_{EQij} SD	SD Mean L_{EQij} SD	n	Mean L_{EQij} SD	SD Mean L_{EQij} SD	Diff. (dBA)
Task variability	204						
Usually same	107	2.3[C]	1.4		—	—	—
Usually changing	97	3.3[C]	2.6		—	—	—

Note: n = three measurements each on 68 subjects.

[A] Source, Reference 1.

[B] Significant trend across response categories (Cuzik's test for trend, p < 0.05).

[C] Significant difference in levels across response categories (one-way ANOVA, p < 0.05).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE III

Performance of Exposure Estimates at the Worker Level Compared with Individuals' Measured Mean Exposure

Measure	n	Exposure Estimates (dBA)					
		Mean	SD	Bias	Precision	Accuracy	ICC
Measured mean L_{EQi}		89.6	3.4	—	—	—	—
Trade[A] ($L_{EQ,TMi}$)	68	88.5	2.0	−1.1	3.8	4.0	0.14
Task-based[A] ($L_{EQ,TBi}$)	68	89.1	2.5	−0.5	3.1	3.1	0.48
Subjective[B] ($L_{EQ,SRi}$)							
Noise frequency	68	91.7	1.0	2.5^C	3.0	3.9	0.11
Raise voice	68	91.6	0.5	2.6^C	3.2	4.1	0.04
Percent time in noise	68	91.8	1.0	2.2^C	3.0	3.7	0.25

Note: n = three measurements per subject.

[A]External noise levels developed from five sites.

[B]Significant difference between estimated and measured levels (paired-sample t-test, $p < 0.05$).

[C]External subjective noise levels developed from a single site.