January 2022

Occupational And Socioeconomic Risk Factors Of Noise-Induced Hearing Loss

Rongcan Sun
phoenix.rc.sun@gmail.com

Occupational and Socioeconomic Risk Factors of Noise-Induced Hearing Loss

Phoenix Rongcan Sun

Department of Environmental Health Sciences

Yale School of Public Health

Primary Advisor

Nicole Deziel, PhD, MHS

Department of Environmental Health Sciences

Yale School of Public Health

Secondary Advisor

Benjamin Roberts, PhD, MPH, CIH

Cardno ChemRisk

A Thesis

Submitted to the Department of Environmental Health Sciences

Yale School of Public Health

In Partial Fulfilment of the Requirements

for the Degree of Master of Public Health

May 2022

Abstract

Hazardous noise (>85 dBA) is among the most common occupational exposures and responsible for the majority of noise-induced hearing loss. Noise-induced hearing loss impacts many aspects of a person's life, including workplace safety, daily routine, economic consequences, and quality of life. However, studies on the relationship between occupational noise exposure and hearing loss were usually done in specific industries rather than the general population. In addition, studies focused on associations between socioeconomic factors and hearing loss rarely take the type of occupation into consideration. In this study, we aim to evaluate both occupational noise exposure and socioeconomic status as risk factors for noise-induced hearing loss among a general United States adult population. Data from the National Health and Nutrition Examination Survey and Occupational Information Network were used to evaluate associations between multiple demographical, socioeconomic, occupational factors and hearing loss. Results from Chi-square, linear regression, and logistic regression models showed that occupational noise exposure was associated with higher hearing thresholds at speech frequencies, and low education attainment was associated with increased odds of hearing loss in the United States adult populations.

Acknowledgments

I would like to start by thanking Dr. Benjamin Roberts for his brilliant guidance through my internship and this thesis, as well as introducing me to the fantastic world of noise research. Similarly, I would like to express my appreciation towards Dr. David Paltiel, who opened the door for me to modeling and operations research in public health. Without him, my perspectives of the field would not have been the same. I would also like to thank Dr. Joshua Wallach and Dr. Nicole Deziel for advising me through my time at YSPH and helping me navigate all the nooks and crannies of environmental health sciences. Finally, my gratitude goes to the Class of 2022 EHS MPH cohort, who have been the best support group I have ever had in my entire academic career.

On a more personal note, I must thank my mother for all the courage and work ethics that she has passed onto me, the vision she had for my education, and the unconditional support she always has had for me. Lastly, with all my heart, my biggest "thank you" goes to my partner for all his incomparable company and tolerance through both the finest and the roughest times in the past five years. Thank you for being the person you are and for having faith in me and everything that I do.

Table of Contents

1. Introduction

Hearing Loss (HL) is more prevalent than either diabetes or cancer and is the third most common chronic physical condition in the United States (U.S.).[1] Hearing loss can be broadly divided into conductive (outer and middle ear), sensorineural (inner ear) hearing loss, the latter of which is commonly caused by noise-induced hearing loss (NIHL), and presbycusis, which is hearing loss related to age.[2] NIHL is predominantly due to occupational exposures.[3] Hazardous noise exposure, defined by exposure greater than 85 dBA, is one of the most common occupational exposures globally.[3] Studies indicate that about 22 million U.S. workers are currently exposed to noise at a hazardous level, and about 25% of U.S. workers have history of such exposure.[4] NIHL can result in different levels of hearing impairment and can have severe consequences both in the workplace and in daily life. For example, workers with NIHL are more likely to miss auditable warnings, which in turn would lead to higher risks of occupational injuries even deaths.[5] It also has detrimental effects on one's daily routine due to communication difficulty, social isolation, stress, and fatigue.[6]

In addition to health and safety concerns, significant direct and indirect costs are also associated with NIHL.[7] A 2016 study showed that for every 1000 noise-exposed U.S. workers, 2.5 healthy working years are lost each year due to hearing impairment, which leads to reduced income.[1] A 2017 study estimated that the U.S. economy would save between $58 to $152 billion dollars if 20% of NIHL can be prevented.[7] Though there has been a consensus on the fact that NIHL leads to economic loss, the extend of it varies based on each individual study's sample population and design.[8] Further, there is evidence that those suffering from hearing impairment tend to be of lower socioeconomic status (SES) and are more vulnerable to additional direct costs and reduction in future earnings.[9]

To broadly assess the health status of the general population broken down by demographic groups, data from the National Health and Nutrition Examination Survey (NHANES) are often used. This cross-sectional study began in the 1960s, and collects data regarding population health and nutrition in the U.S. under the National Center for Health Statistics (NCHS), which is part of the Centers for Disease Control and Prevention (CDC).[10] Each NHANES cycle consists of two years of survey data (i.e. 1999-2000, 2001-2002, etc.) that cover a wide range of the non-institutionalized population. The NHANES program focuses on many aspects of population health, ranging from demographical and SES factors like race/ethnicity, occupation, and income to various physical and laboratory examinations, including audiometry, infectious diseases, and many more.[10] Since there is no large national cohort data for occupational study, NHANES program provides more than adequate datasets for evaluating hearing ability on a population level.

Previous studies have shown an association between occupational noise exposure and hearing loss in workers; however, less attention is paid to such association on a more generalized population. In 2012, Choi and colleagues used data from multiple sources to examine hearing levels in the general population.[11] The authors incorporated data from the Occupational Information Network (O*NET) to estimate the level of hearing impairment based on the provided occupational information.[11] They estimated a noise score for each occupation category based on workplace noise exposure frequency, where higher scores indicated more occupational noise exposure.[12] Hearing loss was defined as pure tone average (PTA) being greater than 25 dB at the speech frequencies (0.5, 1, 2, and 4 kHz), which indicates that a person would need the tone to be 25 dB louder than someone without hearing loss in order to detect the sound stimuli.[13] Choi and colleagues gathered audiometry data from the NHANES (1999-2004) and matched O*NET scores based on the provided occupational information. They showed that the higher the score, the higher the odds of hearing loss in the general U.S. population.[11] Similarly, Emmett and Francis used 1999-2002 NHANES data to show associations between hearing loss and multiple factors regrading SES. The results indicated that individuals with hearing loss had higher odds of being low income, low education, and/or unemployment/underemployment.[9]

The aim of this study is to explore whether occupational noise exposure and other socioeconomic elements are risk factors for hearing loss in the general U.S. population. Previous studies on hearing loss usually only focuses on either occupational or SES factors, yet occupations are often linked with a person's SES. To counter this dissonant in the field, this work builds on previous literatures by incorporating both occupational and SES factors within the same analysis while including data from additional NHANES cycles (1999-2018). This paper will 1) examine if hearing thresholds are significantly different when stratified by occupation and SES factors, and 2) assessing whether the existing associations between hearing loss and its risk factors change with a larger and properly weighted sample population. It is hypothesized that there would not be a significant difference in hearing thresholds between different occupational, demographic, and SES groups.

2. Methods

2.1 NHANES Database Methods

2.1.1 Study Cohort

NHANES program is a cross-sectional study that uses a multistage probability design to sample the US population. Each year, NHANES sample about 5,000 persons in designated mobile locations to perform interviews and examinations.[10] The survey intentionally oversamples groups of individuals, such as people with lower SES, people in minority groups, and older adults.[14] Additionally, surveyed individuals participated in different questionnaires and examinations, so the sample sizes would vary greatly between different datasets (see Supplementary Figure 1).[15] To account for the effect of uneven sampling and make the population more representative for the entire U.S. population, NHANES has a specific system to employ survey weights to constitute the complex survey design.[16]

2.1.2 Audiometric Assessment

Audiometric data were obtained through physical examinations. Trained examiners with the NHANES program perform audiometric exams in dedicated, sound-isolating rooms in their mobile examination centers based on standard NHANES protocols.[17] Hearing thresholds are measured from 0.5 to 8 kHz over an intensity range of -10 to 120 dB. The equipment used is calibrated daily, and the rooms' ambient noise levels are monitored during exam.[17]

2.1.3 Demographic, Occupation, and Socioeconomic Information

Demographic, occupation, and SES data were obtained through structured interviews. The race/ethnicity was coded as 1) non-Hispanic White, 2) non-Hispanic Black, 3) Mexican-American/Other Hispanic, and 4) Other. The NHANES contains occupational information such as working hours per week, type of occupation, length of current job, and occupational noise exposure, and SES information like family income and education.[14]

2.2 Data Collection, Analysis, and Modeling Methods

2.2.1 Data Collection from NHANES Database

Using the R package "nhanesA" (Endres 1/30/21),[18] Demographics, Audiometry Examination, Audiometry Questionnaire, Occupation, and Income datasets from 1999 to 2018 (excluding the 2013-2014 cycle), a total of 9 cycles and 18 years, were extracted through RStudio (RStudio, Boston, MA). Audiometry was not evaluated in the 2013-2014 cycle, and thus that cycle was excluded from this analysis. All datasets were first merged by cycle year, then all yearly data were merged into one. Only individuals who were 18 years and older with complete or partial audiometry (i.e. "AUAEXSTS = Complete" or "AUAEXSTS = Partial") exam were included in the datasets.[19]

2.2.2 Sample Weight Adjustment and Survey Design

The survey weight for the working dataset for this study was calculated based on NHANES Tutorial Module 3: Weighting.[20] The calculation used the codes provided by NHANES with correct adjustments to the number of years (18) in the working dataset. Since the method provided by NHANES does not specify whether the cycle years should be continuous or not while combining datasets for multiple cycles, two different weight calculations were performed. The first method was to calculate combined weight for all 18 years, whereas the second method was to calculate one weight for the first 14 continuous years (1999-2012) and another for the last 4 years (2015-2018), then combine both together. The survey design was then computed based on the 18-year sample weight using "dplyr" (2/7/2022)[21] and "srvyr" (Ellis et al., 2/20/2022)[22] packages with race/ethnicity as the strata.

2.2.3 Occupational Noise Exposure Estimates (<u>Analysis completed by Dr. Roberts</u>)

The O*NET database contains survey results that reflect occupational noise exposure in the workplace.[11] One of the O*NET survey question is "Sounds, Noise Levels are Distracting, etc. (element ID IV.C.2.b.1.a)", which uses an ordinal measure from 1 to 5 to record the frequency of occupational noise exposure. The 5 responses are 1) "never", 2) "once a year or more but not every month", 3) "once a month or more but not every week", 4) "once a week or more but not every day", and 5) "every day".[23] The job titles of participants were grouped by the Standard Occupational Classification (SOC) System major (i.e. most broad) codes.[24] The O*NET score was the mean survey question score for each SOC category, which reflected only the frequency, but not the intensity, of the exposure. Each SOC category was then cross listed with the occupation categories from NHANES and assigned to each individual in the working dataset of this study. Lastly, the O*NET score of each SOC was merged into the dataset as well. The computation of this section was performed in STATA (StataCorp, College Station, TX) by Dr. Roberts.

2.2.4 Demographic and Socioeconomic Characteristics

Chi-square tests were used to gather descriptive statistics of demographic and SES variables between people with and without hearing loss in the sample population. To account for the non-random design of the HNANES, the chi-square test was corrected with survey-specific weights for the survey design. Hearing loss was defined based on hearing threshold, where individuals with pure tone average (PTA) greater than 25 dB were categorized as having hearing loss.[13] Individuals with family income less than $20,000 were defined as low income based on NHANES categorization.[9] And individuals who worked less than 35 hours per week were defined as unemployment/underemployment based on the Bureau of Labor Statistics.[25] Lastly, for people with occupations that are in the fourth or fifth quintile, their occupations were considered "noisy". All analyses were performed through RStudio with the "survey" package (Lumley 7/19/2021).[26]

2.2.5 Linear Regression Models

Liner regression models were used to predict hearing thresholds. O*NET scores were used as independent variable in the first linear regression model, whereas the average test result for speech frequencies (i.e. hearing threshold) was used as the continuous dependent variable. In the second model, all occupation categories were divided into 5 quintiles of O*NET scores as the independent variable for the regression model, whereas the dependent variable was the same as the first model. Additionally, to ease interpretation, the intercepts were centered at 18 years old (the youngest age of the dataset used for this paper). Lastly, to account for the non-linear effect of noise exposure and age, (centered-age)^2 was modeled into the regression as well. The equation for the linear regression model is

$$\mathrm{y} = a + b_1 x_1 + b_2 x_2 + b_3 x_3 \ (1)$$

where y was hearing threshold, x_1 was O*NET score (first linear regression model) or O*NET score quintile (second linear regression model), x_2 was age (centered at 18), and x_2 was (centered-age)^2. The coefficients (b_1, b_2, and b_3) were indicators of change in hearing thresholds as the values of independent variables change. All models were performed through RStudio with the "survey" package (Lumley 7/19/2021).[26]

2.2.6 Multiple Logistic Regression Models

Multiple logistic regression models were used to assess the association between various demographic and SES factors and hearing loss. The demographic variables were sex (male/female), age (<45/>=45), and race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, and Other). The SES variables used were education (<High School/>= High School), income (family income <$20,000/>=$20,000), and working hours (<35 hours per week/>=35 hours per week) as the marker for unemployment/underemployment. All demographic and SES variables were categorized based on the study done by Emmett and Francis.[9] Lastly, occupational noise exposure (O*NET score quintile <4/>=4) was used as the occupational exposure variable to indicate whether the job was noisy based on the significant cutoff from the Choi et al. study.[11] A null model was first fitted with hearing loss, the dependent variable, only. Then demographic and SES factors were added in one by one. At each step of the modeling process, an Akaike information criterion (AIC) value was generated, and a total of 7 models were generated. The regression equation for the final logistic regression model is

$$\ln \frac{y}{1-y} = a + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + b_5 x_5 + b_6 x_6 + b_7 x_7 \quad (2)$$

where y was hearing threshold, x_1 was noisy job, x_2 sex, x_3 age, x_4 race, x_5 education, x_6 income, and x_7 employment. Similar to Equation 1, The coefficients (b_1, b_2, …, b_7) were indicators of change in hearing thresholds as the values of independent variables change. The adjusted odds ratios were calculated based on these coefficients. All models and odds ratio calculations were performed through RStudio with the “survey” package (Lumley 7/19/2021).[26]

3. Results

3.1 Data Collection from NHANES Database

In total, 16,078 data points for analysis. All 16,078 individuals had results for hearing test and basic demographic information. However, not every individual had a complete record of SES variables. Therefore, the number of data points for each variable varies due to data collection differences. For O*NET score, n=4,349, for education level, n=10,746, for income, n=9,095, and for weekly working hours, n=7,845.

3.2 Sample Weight Adjustment

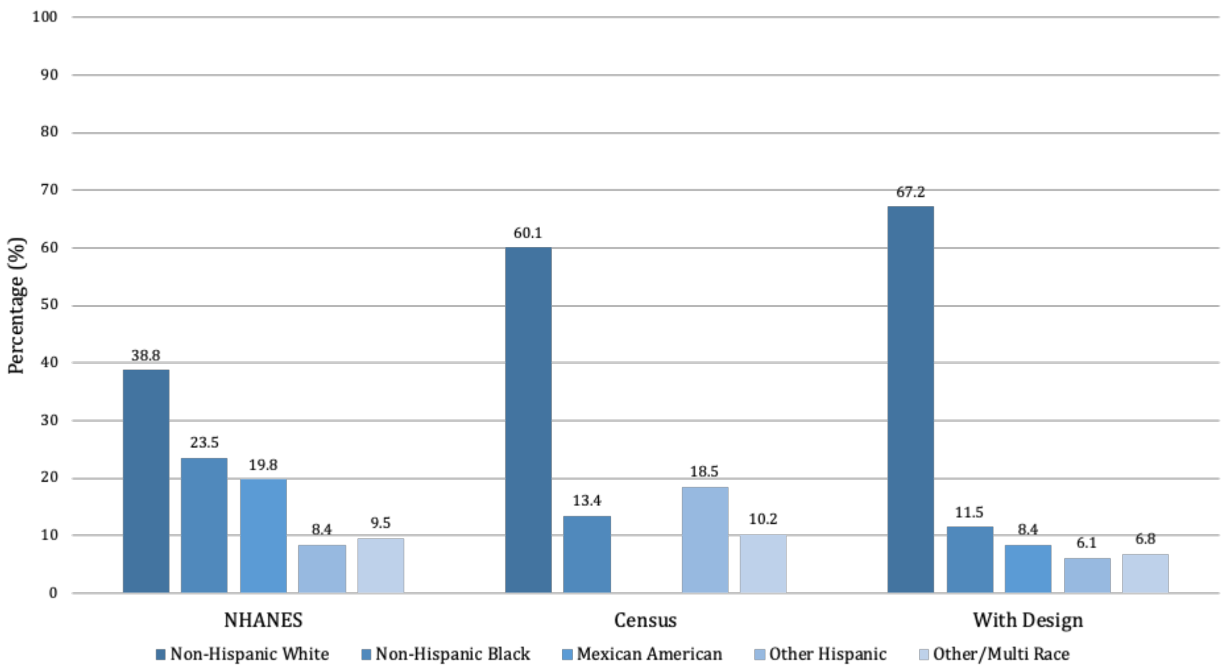

Figure 1. Race/Ethnicity distribution in NHANES original data, Census data, and NHANES data adjusted with proper sample design.

The two methods of sample weight calculation rendered similar results. The corrected proportions of each race/ethnicity group from these two methods are similar to each other, and similar to the 2021 Census data[27]. To simplify the rest of the analysis, the 18-year combined method was used through the analysis to generate the backbone survey design for regression models. Figure 1 shows the race/ethnicity distribution of the original NHANES dataset, the Census data, which represents the actual distribution of the population, and the NHANES dataset distribution after the weighted design was implemented. Before the survey weight adjustment, NHANES data contained 61.2% people of color, where the actual population proportion was 29.9% based on the Census data. After adjustment, this proportion in the NHANES dataset was lowered to 22.8%.

3.3 O*NET Scores

Table 1. SOC job titles, corresponding O*NET scores and quintile, and number of people in each quintile for the dataset used.

Job Title	O*NET Score	Quintile	Number of People
Computer and mathematical occupations	1.755	1	991
Business and financial operations occupations	2.372		
Life, physical, and social science occupations	2.456		
Sales and related occupations	2.508		
Food preparation and serving related occupations	2.536	2	726
Education, training, and library occupations	2.542		
Arts, design, entertainment, sports, and media occupations	2.562		
Community and social service occupations	2.593		
Healthcare support occupations	2.704	3	1058
Management occupations	2.719		
Architecture and engineering occupations	2.733		
Personal care and service occupations	2.768		
Healthcare practitioners and technical occupations	2.776		
Office and administrative support occupations	2.979	4	920
Building and grounds cleaning and maintenance	3.051		
Protective service occupations	3.253		
Farming, fishing, and forestry occupations	3.280		
Production occupations	3.816	5	654
Installation, maintenance, and repair occupations	3.828		
Construction and extraction occupations	4.153		
Transportation and material moving occupations	4.189		

Table 1 displays the results for O*NET scores for each occupation categories. There are 21 total categories with O*NET scores range from 1.755 to 4.189. The first quintile has scores range from 1.755 to 2.508, the second quintile from 2.536 to 2.704, the third quintile from 2.719 to 2.776, the fourth quintile from 2.979 to 3.280, and the fifth quintile ranges from 3.816 to 4.189. There were also different numbers of individuals in each quintile, where the third quintile had the most individuals (1058) and the fifth quintile had the least (654).

3.4 Demographic and Socioeconomic Characteristics

Table 2. Demographic characteristics

Characteristics	Normal Hearing n (%)	Hearing Loss n (%)	P-value
Demographics (n=16,078)			
Male	6294 (46.8)	1567 (59.4)	<0.001*
Age ≥ 45	5317 (39.6)	2410 (91.4)	<0.001*
Race/Ethnicity (n=16,078)			<0.001*
Non-Hispanic White	5113 (38.0)	1469 (55.7)	
Non-Hispanic Black	3242 (24.1)	368 (13.9)	
Hispanic	3599 (26.8)	620 (23.5)	
Other	1486 (11.1)	181 (6.9)	
Education (n=10,746)			
Education < High School	3933 (45.6)	1351 (63.4)	<0.001*
Income (n=9,095)			
Income < $20,000	2327 (30.4)	513 (35.4)	0.755
Working Hours (n=7,845)			
Weekly Working Hours < 35	2038 (28.5)	250 (35.8)	0.325
Occupation with O*NET Scores (n=4,349)			
Noisy Occupation	1623 (39.4)	171 (40.5)	0.413

The prevalence of hearing loss stratified by various SES factors are presented in Table 2. People with hearing loss were more likely to be male (59.4% vs. 46.8%, p<0.001), 45 years or older (91.4% vs. 39.6%, p<0.001), and non-Hispanic white (55.7% vs. 38.0%, <0.001), than people with normal hearing. In terms of SES factors, individuals with hearing loss were more likely to not have completed high school (63.4% vs. 45.6%, p<0.001). Low income, underemployment/unemployment, and having a noisy job were not significantly associated with hearing loss.

3.5 Linear Regression Models

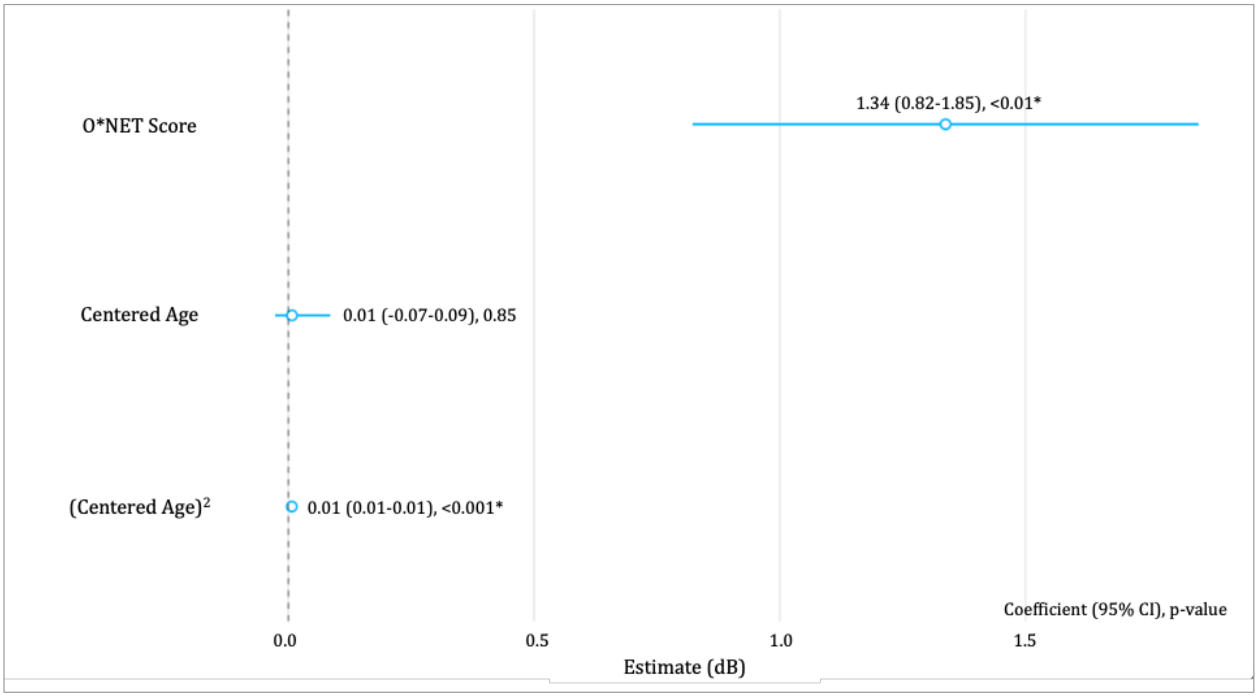

Figure 2a. Results from the first linear regression model with O*NET score being the main regressor

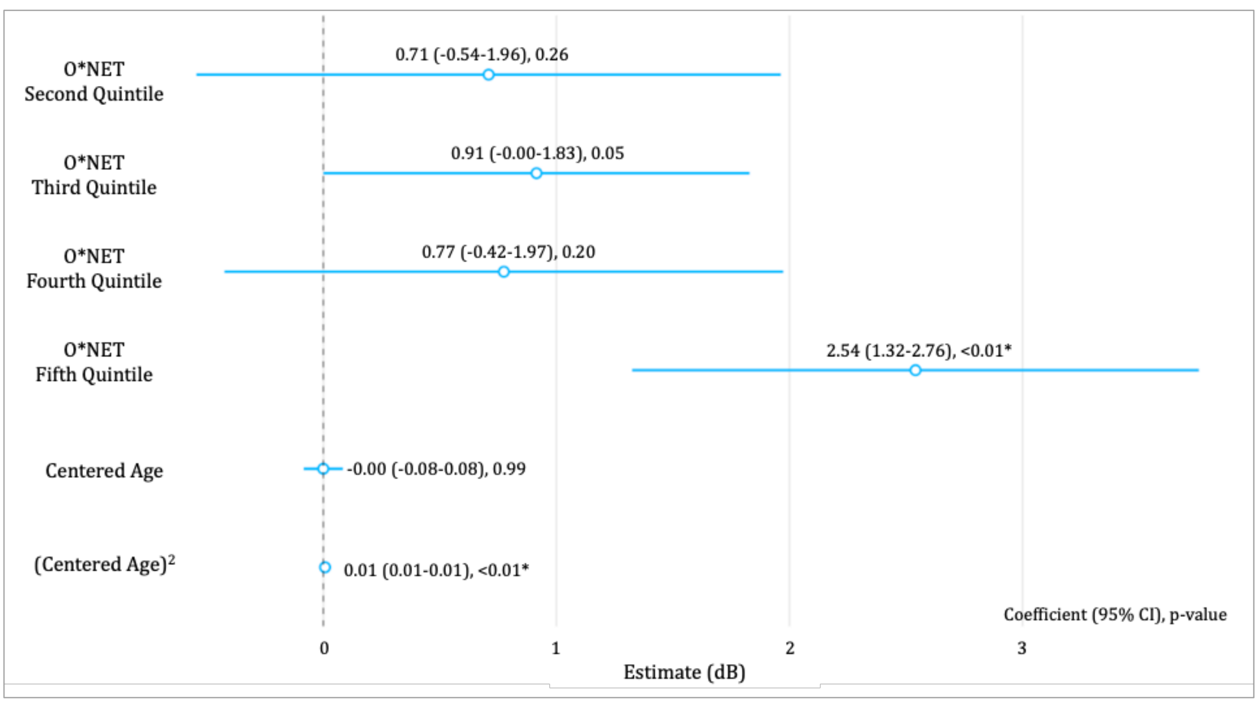

Figure 2b. Results from the second linear regression model with O*NET score quintile being the main regressor

The coefficients and p-values from linear regression models are presented in Figures 2a and 2b. Figure 2a shows that for every unit increase in O*NET score, there was 1.34 dB significant increase in hearing threshold (i.e. decline in hearing ability). Figure 2b shows that compared to the first quintile of O*NET score, there were significant increase in hearing threshold in the fifth quintile. Comparing to individuals with jobs that had O*NET scores in the first quintile, individuals with jobs that had O*NET scores in the fifth quintile had worse hearing test threshold, where their test results were 2.54 dB higher ($p<0.01$). Additionally, for both models, the regression model results with independent variable $(\text{centered-age})^2$ were also significant. For every 1 unit increase in $(\text{centered-age})^2$, starting at 18 years of age, hearing test average increased 0.01 dB ($p<0.01$).

3.6 Multiple Logistic Regression Models

Table 3. AIC results with hearing loss as the dependent variable

Model	Model Independent Variables	AIC	Differences
0	Null	12802.908	
1	Loud Job	2852.921	9949.987
2	Loud Job + Sex	2835.299	17.622
3	Loud Job + Sex + Age	2419.370	415.929
4	Loud Job + Sex + Age + Race	2418.350	1.02
5	Loud Job + Sex + Age + Race + Edu	2361.449	56.901
6	Loud Job + Sex + Age + Race + Edu + Income	1387.951	973.498
7	Loud Job + Sex + Age + Race + Edu + Income + Employment	1332.182	55.769

The step-by-step modelling process is presented in Table 3. AIC values decreased as more independent variables were added. Large decreases in AIC values were seen when age (415.929 points lower than the previous step) and income (973.498 points lower than the previous step) were added. Model 7 with all seven independent variables was chosen for further analysis because it returned the lowest AIC value.

Table 4. Unadjusted and adjusted odds ratio from logistical regression models

	Unadjusted		Adjusted	
Characteristics	**OR (95% CI)**	**P-value**	**OR (95% CI)**	**P-value**
Sex				
Female	Reference		Reference	
Male	1.44 (1.28-1.63)	<0.001*	1.74 (1.13-2.68)	0.015*
Age				
< 45 years	Reference		Reference	
>= 45 years	12.53 (10.17-15.44)	<0.001*	8.67 (3.87-19.43)	<0.001*
Race/Ethnicity				
Non-Hispanic White	Reference		Reference	
Non-Hispanic Black	0.46 (0.40-0.54)	<0.001*	0.67 (0.49-0.92)	0.018*
Hispanic	0.61 (0.52-0.73)	<0.001*	0.52 (0.31-0.88)	0.017*
Other	0.72 (0.54-0.94)	0.018*	1.20 (0.46-3.13)	0.700
Education				
>= High School	Reference		Reference	
< High School	2.09 (1.84-2.37)	<0.001*	2.04 (1.25-3.31)	0.006*
Income				
>= $20,000	Reference		Reference	
< $20,000	1.02 (0.88-1.20)	0.755	0.61 (0.31-1.18)	0.134
Employment				
>= 35 hours/week	Reference		Reference	
< 35 hours/week	1.14 (0.88-1.48)	0.325	1.38 (0.81-2.34)	0.219
Noisy Job				
Normal Job	Reference		Reference	
Noisy Job	1.16 (0.82-1.64)	0.413	0.96 (0.51-1.83)	0.900

The unadjusted and adjusted odds ratios (ORs) generated from model 7 are shown in Table 4. The significance of each variable was mostly consistent between unadjusted and adjusted ORs except for Other race/ethnicity, which had a significant unadjusted OR but an insignificant adjusted one. After adjustment, there were significant associations to hearing loss present in sex, age, non-Hispanic Black race category, Hispanic race category, and education. Male had 1.74 times higher odds (OR=1.74, CI=1.13-2.68, p=0.015) of hearing loss than female. People who were 45 years and older had more than 8 times higher odds (OR=8.67, CI=87-19.43, p<0.001) of hearing loss than those who were younger. Compared to Non-Hispanic White individuals, Non-Hispanic Black individuals were 0.67 times as likely (OR=0.67, CI=0.49-0.92 p=0.018) to have hearing loss, where the likelihood was 0.52

(OR=0.52, CI=0.31-0.88, p=0.017) in Hispanic individuals. Lastly, individuals with low education attainment had about 2 times higher odds (OR=2.04, CI=1.25-3.31, p=0.006) of hearing loss than those had more education. Low income, underemployment/unemployment, and noisy job were not found to have significant associations with hearing loss.

4. Discussion

In this study, we found that occupational noise exposure leads to higher hearing thresholds (i.e. worse hearing) at speech frequencies in the general U.S. population. The associations were established through building a novel occupational exposure score and a dataset that was properly weighted based 1999-2018 NHANES cycles. Additionally, when modeled together with demographic, SES, and occupational factors, low education attainment was found to be a significant risk factor for hearing loss in the general population.

In agreement with previous studies, our results show significant differences in prevalence and odds of hearing loss with demographic strata. Hearing loss prevalence was higher in male, in those 45 years of age or older, and in non-Hispanic White. Similarly, the odds of hearing loss were higher in male and in people who are 45 years of age or older, but less in non-Hispanic Black and Hispanic population. Male sex and age are known non-modifiable risk factors of hearing loss that have been discussed in multiple previous studies.[28] Male tends to have worse hearing than female due to their occupations and recreational activities.[29,30] For example, the occupations with O*NET scores in the fifth quintile are production; installation, maintenance, and repair; construction and extraction;

and transportation, where all of them are male-dominant occupations.[29] Age in NIHL is corresponded to the cumulative effect of noise exposure, which was also modeled in the regression models of this study and will be touched on with the model discussion. As for the race/ethnicity factor, there has not been a consensus on the reason behind non-Hispanic Black and Hispanic races being protective factors.

The significance of occupational noise exposure in the general population was mostly seen in the linear regression results. The significant change in using O*NET score as a regressor showed that occupational noise exposure leads to worse hearing. Every 1 unit increase in the O*NET score indicates that the frequency of noise exposure in the workplace goes up a level. For example, if the score increases from 1 to 2, the frequency would increase from "never" to "once a year or more but not every month". Thus, the result indicated that for every level increase in noise exposure frequency, people needed the volume to be 1.34 dB louder at hearing frequencies than before. Since the decibel is on a logarithmic scale, where a 3 dB increase means the power of the sound doubles,[3] a 1.34 dB increase gives more than 30% increase in sound power. Similarly, from the results using the O*NET quintile as a regressor, the 2.54 dB increase in the fifth quintile means people essentially need a doubled sound power to hear at speech frequencies comparing to those in the first quintile. In addition to the main regressors of these two models, covariate $(\text{centered-age})^2$ was also found to be significant in both linear regression models. This indicated that $(\text{centered-age})^2$ accurately represented the non-linear effect of age and exposure accumulation. These results were somewhat consistent with the Choi et al. study, where the trend of higher hearing threshold as a result of increase in O*NET score stands, yet the quintiles that showed significance were different from the Choi et al. study. Among all three different models Choi and colleagues used, the fourth and the fifth quintiles were always the ones showing significant increase in hearing thresholds, rather the fifth quintile only in our study. However, since we used the broadest SOC occupation codes rather than the more detailed ones in the Choi et al. study on top of a completely different dataset, it is very likely that misclassification has happened during the process, which can potentially explain the differences.

Though higher frequencies of occupational noise exposure were shown to be associated with higher hearing threshold (i.e. wore hearing), there was no significant odds between O*NET scores and hearing loss from the logistic regression model results. There are several possible explanations to this result. First, since the NHANES is a cross-sectional study, longitudinal effects can hardly be inferred from analyses based on NHANES data. While occupational noise exposure clearly led to worse hearing on a population level, the time needed for the effects to accumulate to the extent of hearing loss cannot be factored into the model with the cross-sectional dataset. Secondly, using the fourth quintile as the cutoff for whether an occupation was noisy was a choice based on the Choi et al. study, where they found fourth and fifth quintiles significant in their regression models. However, since the characteristics distribution was different in our study from theirs, this categorization may not be the most accurate, and there can be other ways to define the cutoff based on the O*NET score. Additionally, O*NET data collections target workers who were actively on the job during the survey period.[12] Therefore, there could be bias in sampling due to healthy worker effects, where the people who had hearing loss that prevented them to be working were not sampled in the dataset. Lastly, since the multiple logistic regression models have not been done before with O*NET score, it is difficult to compare these results to previous studies. Nevertheless, although occupational noise exposure cannot be definitively concluded a risk factor for hearing loss with the current data, it was indeed associated with decline in hearing ability at speech frequencies.

As for SES stratifications, the only significant factor shown in the prevalence and odds ratios was low education attainment. The prevalence of hearing loss was higher in individuals with low education attainment was in agreement with the Emmett and Francis study. However, Emmett and Francis also found income and unemployment/underemployment to be significant strata in prevalence of hearing loss, which were not seen with this study population. The differences are most likely due to the size of the sample population and the differences in sample weights because of the larger range of NHANES cycles.

In terms of the odds ratios, Emmett and Francis measured the risk of having different SES in individuals of hearing loss, whereas in our study, the risk of hearing loss was measured in individuals with different SES. Our analysis results showed that low education attainment increased the odds of having hearing loss, and the Emmett and Francis study showed that having hearing loss increased the odds of having low education attainment. These two results made low education attainment a very interesting factor when studying hearing loss in the general population. Studies have shown that hearing loss during school year decrease students' abilities to pursue more education, yet low education often led to noisy occupations, which further makes hearing threshold higher at speech frequencies.[31] Additionally, Emmett and Francis also found that hearing loss was the risk factor for low income and unemployment/underemployment, yet these two factors were not found to significantly increase the odds of hearing loss in our study population.

There are several limitations in the study. The biggest limitation came from the NHANES's study design itself. With the cross-sectional design, temporality cannot be established, yet it is an extremely important factor that plays into the cumulative effect of noise exposure and NIHL. Within the variables of NHANES, the factor of unemployment/underemployment could not be defined with more details, because NHANES did not collect whether the person was unable to maintain more than 35 hours of work per week, or chose not to work more than 35 hours per week. This also provided another possible explanation for the insignificant association we saw between this SES factor and hearing loss in the study. Due to NHANES's data collection methods, each individual participant had different measurements, which lead to the different numbers of datapoints throughout the analyses of this study with the occupational subset being the smallest (4,349 out of 16,078). To ensure the representativeness of the models, a sensitivity analysis was done by running all regression models with only people who were 45 years older or above. The results varied slightly in numbers, but the general trends were exactly the same as analyses with the full dataset. This indicated that although with some missing data, our analysis still reflects the true trends in the sample population.

Additionally, since the SOC system was designed for track economic trends instead of actual exposure, the O*NET score, measuring the frequency of noise exposure in the workplace, can only serve as a proxy for the level of exposure. Although it is sufficient to represent occupational noise exposure on a low to high scale, the extent of the exposure and whether the exposure exceeds safe limits cannot be inferred from this measure. Furthermore, only the least detailed SOC codes were used for this analysis, and it is possible that lack of the clear dose-response relationship in the logistics model is due to the lack of specificity in the codes. Nevertheless, without a large occupational cohort or detailed employment information, the O*NET score is still a useful measure for occupational noise exposure in the general population. Overall, another detail that might be considered as a limitation is that neither NHANES nor O*NET collected information on hearing protection use, which can potentially skew results in those occupations that require to have hearing protection.

Lastly, there are some possible future directions to be considered based on results and limitations. In terms of the methods themselves, SES factors can possibly be refined more, and other covariates related to possible noise exposure, such as housing status, can be added to the model. Recently, there have been discussions on methods for generalized noise exposure study in general population using data collected through Apple Watch[32], which can open doors for data varieties, awareness, and most importantly, temporality. This may lead to possibilities in the future for studying total noise exposure. In addition to pulling data from NHANES, a large occupational cohort that infer temporality is more than needed to study chronic occupational exposures like noise. Since the association between noise exposure in the workplace and reduction in hearing ability at speech frequencies is seen with the general population, studies of hazard reduction and protection implementation can be done for the general population as well. Several studies have been focused on methods to reduce disability-adjusted life years for noised-exposed workers in certain industries[8], which can hopefully be expanded to larger population.

5. Conclusion

With the high prevalence of hearing loss in the U.S. population, quantifying occupational and socioeconomical risk factors for noise-induced hearing loss is crucial. In this study, we found that occupational noise exposure was associated with higher hearing thresholds at speech frequencies in the U.S. population and low education attainment was associated with increased odds of hearing loss. To better refine these relationships and establish prediction models, a large occupational cohort for longitudinal study is very much needed in the field. Overall, the awareness of occupational noise exposure in all occupations should be heightened and more focus should be given to people who started work earlier in life without obtaining higher education. These implications can help not only researchers, but public health practitioners, educators, and leaders to help reduce exposure and preserve hearing on a population level.

References

1. Masterson, E. A., Bushnell, P. T., Themann, C. L. & Morata, T. C. Hearing Impairment Among Noise-Exposed Workers — United States, 2003–2012. *MMWR Morb. Mortal. Wkly. Rep.* **65**, 389–394 (2016).
2. Eric Bieledeld. Anatomy and Physiology of the Ear: Normal Function and the Damage Underlying Hearing Loss. in *The Noise Manual* (eds. Deanna K. Meinke, Elliott H. Berger, Dennis P. Driscoll, Richard L. Neitzel, & Kathryn Bright) (AIHA, 2022).
3. U.S. Department of Health and Human Services. *Criteria for a recommended standard... occupational noise exposure, revised criteria 1998.* https://www.cdc.gov/niosh/docs/98-126/ (1998) doi:10.26616/NIOSHPUB98126.
4. Tak, S., Davis, R. R. & Calvert, G. M. Exposure to hazardous workplace noise and use of hearing protection devices among US workers-NHANES, 1999-2004. *Am. J. Ind. Med.* **52**, 358–371 (2009).
5. Shkembi, A., Smith, L., Roberts, B. & Neitzel, R. Fraction of acute work-related injuries attributable to hazardous occupational noise across the USA in 2019. *Occup. Environ. Med.* **79**, 304–307 (2022).
6. Themann, C., Suter, A. & Stephenson, M. National Research Agenda for the Prevention of Occupational Hearing Loss—Part 1. *Semin. Hear.* **34**, 145–207 (2013).
7. Neitzel, R. L., Swinburn, T. K., Hammer, M. S. & Eisenberg, D. Economic Impact of Hearing Loss and Reduction of Noise-Induced Hearing Loss in the United States. *J. Speech Lang. Hear. Res.* **60**, 182–189 (2017).
8. Themann, C. L. & Masterson, E. A. Occupational noise exposure: A review of its effects, epidemiology, and impact with recommendations for reducing its burden. *J. Acoust. Soc. Am.* **146**, 3879–3905 (2019).
9. Emmett, S. D. & Francis, H. W. The Socioeconomic Impact of Hearing Loss in U.S. Adults. *Otol. Neurotol.* **36**, 545–550 (2015).
10. National Center for Health Statistics. National Health and Nutrition Examination Survey Overview.
11. Choi, Y.-H., Hu, H., Tak, S., Mukherjee, B. & Park, S. K. Occupational noise exposure assessment using O*NET and its application to a study of hearing loss in the US general population. *Occup. Environ. Med.* **69**, 176–183 (2012).
12. U.S. Department of Labor. O*NET Data Collection Program. https://www.dol.gov/agencies/eta/onet/data-collection.
13. National Institute for Occupational Safety and Health. Inquiring Ears Want to Know - A Fact Sheet about Your Hearing Test.
14. National Center for Health Statistics. National Health and Nutrition Examination Survey: Plan and Operations, 1999–2010. (2013).
15. National Center for Health Statistics. National Health and Nutrition Examination Survey: 1999–2020 Survey Content Brochure.
16. National Center for Health Statistics. National Health and Nutrition Examination Survey, 2015−2018: Sample Design and Estimation Procedures. (2020).
17. Center for Disease Control and Prevention. National Health and Nutrition Examnination Survey Audiometry Procedures. (2018).
18. Christopher J. Endres. Package 'nhanesA'. (2021).
19. National Center for Health Statistics. 2017-2018 Data Documentation, Codebook, and Frequencies: Audiometry.

20. National Center for Health Statistics. NHANES Tutorial Module 3: Weighting.
21. Package ‘dplyr’.
22. Greg Freedman Ellis, Thomas Lumley, Tomasz ˙Zółtak, Ben Schneider, & Pavel N. Krivitsky. Package ‘srvyr’.
23. O*NET Resource Center. The O*NET® Content Model. https://www.onetcenter.org/content.html.
24. Standard Occupational Classification (SOC) System. https://www.bls.gov/soc/.
25. U.S. Bureau of Labor Statistics. Labor Force Statistics from the Current Population Survey: Concepts and Definitions. https://www.bls.gov/cps/definitions.htm.
26. Thomas Lumley. Package ‘survey’. (2021).
27. U.S. Census Bureau. U.S. Census Bureau QuickFacts: United States. https://www.census.gov/quickfacts/fact/table/US/PST045221.
28. Nelson, D. I., Nelson, R. Y., Concha-Barrientos, M. & Fingerhut, M. The global burden of occupational noise-induced hearing loss. *Am. J. Ind. Med.* **48**, 446–458 (2005).
29. U.S. Bureau of Labor Statistics. Labor Force Statistics from the Current Population Survey: Household Data.
30. Feder, K., Marro, L., McNamee, J. & Michaud, D. Prevalence of loud leisure noise activities among a representative sample of Canadians aged 6–79 years. *J. Acoust. Soc. Am.* **146**, 3934–3946 (2019).
31. Roberts, B. & Neitzel, R. L. Noise exposure limit for children in recreational settings: Review of available evidence. *J. Acoust. Soc. Am.* **146**, 3922–3933 (2019).
32. Neitzel, R. L. *et al.* Toward a better understanding of nonoccupational sound exposures and associated health impacts: Methods of the Apple Hearing Study. *J. Acoust. Soc. Am.* **151**, 1476–1489 (2022).

Appendix

Supplementary Table 1. NHANES survey contents for datasets used in this study

Component	Sample Description	NHANES Cycle Years										
		99-00	01-02	03-04	05-06	07-08	09-10	11-12	13-14	15-16	17-18	19-20
Questionnaire												
Demographics	All ages (12-85)											
Hearing	1 year and over											
Occupation	16 years and over											
Income	All ages (12-85)											
Examination												
Audiometry	½ sample (20-69 years)*											

Blue: Component or lab test conducted on original sample description
Green: Change from original sample description
White: Component or lab test not conducted
*20-69 years were tested for most of the cycles except for 05-06 (12-19 years and 70 years and above), 07-08 (12-19 years), and 09-10 (12-19 years and 70 years and above).