BIOS620: Final Project
Leo Ware
5/9/2025

Introduction

Repetitive movement injury is an umbrella term for a large number of muscle or tendon injuries caused by repetitive movement, awkward posture, or other risk factors. Examples include carpal tunnel syndrome. Repetitive movement injury likely affects 7-14% of the population and may be associated with workplace-specific risk factors (Tulder, 2007).

The question of the link between occupational risk factors and repetitive movement injury is an important one. We sought to examine the relationship between poverty status (as defined by the relevant poverty line for household income) and repetitive movement injuries among employed adults in the National Health Interview Survey from 2023.

The National Health Interview Survey is a nationwide survey conducted by the Center for Disease Control to monitor a variety of health outcomes across the US. (CDC, 2023) The survey interviews approximately 27,000 people every year and has been conducted since 1957. In addition to collected data, the CDC publishes survey stratum and weight information for interviewed subjects, making it possible to use complex survey methodologies to more accurately extrapolate survey responses to the US population.

Methods

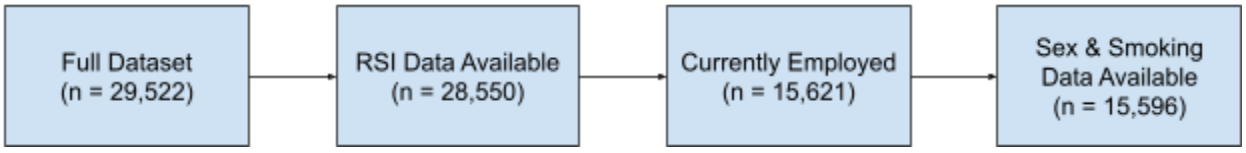
Dataset/Complex Survey Description

Data was taken from the results of the National Health Interview Survey (NHIS) from 2023. The NHIS is a national interview survey overseen by the CDC and conducted by Census Bureau personnel. The survey target population is civilian noninstitutionalized persons residing in the 50 states and DC. The survey uses a stratified cluster sample design designed to represent the entire target population. Sample weights are assigned to each unit that make it possible to adjust for unequal sampling probability and response bias, allowing unbiased estimates of totals in the target population. Standard errors/variance estimation were performed with Taylor series linearization. The survey had a 47.0% response rate and included 29,552 adults in the final sample. (CDC, 2024)

The outcome variable was whether the subject had a repetitive movement injury in the last three months, for example carpal tunnel syndrome. Subjects for which data on repetitive movement injury was unavailable were dropped. The dataset was subset to subjects currently employed, defined as having worked for pay in the last week prior to the study. After subsetting to meet these two criteria, 15,621 subjects remained in the dataset.

A number of covariates from the NHIS dataset were included in the analysis. These included age, sex, race/ethnicity, region (northeast, midwest, south, west), smoking status (whether the subject was ever a smoker), and poverty status (whether the subject's family income was below the poverty line in their area). Age which was binned into 4 brackets (18-24, 25-44, 45-64, 65+). Data on sex and smoking status was unavailable for some subjects. These subjects were dropped, leaving a total of 15,596 in the final analysis.

Figure 1. Study Inclusion Diagram



Statistical Analysis

Analysis was performed using R version 4.4.2. The R package survey was used for statistical analysis (Lumley, 2024). This package supports using complex survey weighting methodologies in the calculation of chi-squared tests and logistic regression, along with other features.

Subjects who had experienced a repetitive strain injury in the last three months were compared to those who had not along each covariate. Raw number, weighted number, and weighted percentage were calculated, with weights accounting for the complex survey design. Comparisons were performed using the chi-squared test implemented in the survey package, which accounts for the complex survey design.

A logistic regression model (quasi-binomial to account for possible overdispersion) was fit to model the relationship between poverty and repetitive movement injury when controlling for age, sec, race/ethnicity, region, smoking status, and poverty status.

Results

Results of the descriptive analysis are displayed in table 1. In the data set, repetitive strain injury is associated with age ($p < 0.001$), race/ethnicity ($p = 0.002$), and smoking status ($p < 0.001$). It was not found to be associated with sex, region, or poverty status.

Table 1: Descriptive Analysis

	Overall		Repetitive Strain Injury			No Repetitive Strain Injury			
value	N (raw)	N (w)	N (raw)	N (w)	w%	N (raw)	N (w)	w%	p-value
Age									
25-44	6712	66118058	686	6449067	9.75%	6026	59668991	90.25%	<0.001*

18-24	1218	19161554	79	1204475	6.29%	1139	17957079	93.71%	
45-64	6041	54442519	711	6122200	11.25%	5330	48320319	88.75%	
65+	1625	10197224	135	802832	7.87%	1490	9394392	92.13%	
Sex									
Female	7745	69460559	823	7062455	10.17%	6922	62398104	89.83%	0.121
Male	7851	80408263	788	7516120	9.35%	7063	72892143	90.65%	
Race/ Ethnicity									
Non-Hispanic White	9850	90568054	1088	9635016	10.64%	8762	80933038	89.36%	0.002*
Hispanic	2665	27957539	227	2162114	7.73%	2438	25795425	92.27%	
Non-Hispanic Asian	1039	10339886	101	951249	9.20%	938	9388637	90.80%	
Non-Hispanic Black	1638	17083030	153	1497226	8.76%	1485	15585804	91.24%	
Non-Hispanic Other	404	3970846	42	332969	8.39%	362	3637877	91.61%	
Region									
Northeast	2403	25777125	240	2420765	9.39%	2163	23356360	90.61%	0.514
Midwest	3449	31538761	339	2967102	9.41%	3110	28571659	90.59%	
South	5730	56751328	588	5432581	9.57%	5142	51318747	90.43%	
West	4014	35852141	444	3758126	10.48%	3570	32094015	89.52%	
Smoking Status									
Never	10699	105045278	998	9207528	8.77%	9701	95837750	91.23%	<0.001*
Former/ Current	4897	44673201	613	5366163	12.01%	4284	39307038	87.99%	
Income to Poverty Ratio									
Above 1	14776	141643920	1539	13889634	9.81%	13237	127754286	90.19%	0.228
Below 1	820	8275435	72	688941	8.33%	748	7586494	91.67%	

* statistically significant at the alpha = 0.05 level
All p-values represent Pearson's X^2 tests with Rao & Scott adjustment. They were calculated using the svychisq function from survey, which accounts for complex survey design.

Table 2 shows the results of modeling. Repetitive strain injury was modeled as a function of age, sex, race/ethnicity, region, poverty status, and smoking status.

Compared to the reference group of subjects aged 25-44, subjects aged 18-24 showed decreased risk (OR = 0.668, p = 0.007) and subjects aged 65 showed decreased risk (OR = 0.737, p = 0.006), when controlling for age, sex, race/ethnicity, region, the poverty status, and smoking status. Compared to the reference group of female subjects, male subjects showed decreased risk (OR = 0.882, p = 0.039) when controlling for age, race/ethnicity, region, poverty status, and smoking status. When compared to the reference group of non-hispanic white subjects, hispanic subjects showed decreased risk (OR = 0.716, p < 0.001) when controlling for age, sex, region, poverty status, and smoking status. When compared to the reference group of never smokers, current/former smokers showed an increased risk (OR = 1.334, p < 0.001) when controlling for age, sex, race/ethnicity, region, and poverty status.

Subjects aged 45-64, non-hispanic asian subjects, non-hispanic black subjects, and non-hispanic other race subjects showed no statistically significant change in risk of repetitive strain injury in the model when compared to the respective reference groups. Region had no statistically significant association with risk in the model when controlling for age, sex, race/ethnicity, poverty status, and smoking status.

Table 2: Model Results

Term	Group	Odds Ratio	95% Confidence Interval	p value
(Intercept)		0.11	(0.091, 0.133)	<0.001*
Age	18-24	0.668	(0.499, 0.894)	0.007*
	25-44	(reference)		
	45-64	1.13	(0.997, 1.281)	0.056
	65+	0.737	(0.594, 0.914)	0.006*
Sex	Female	(reference)		
	Male	0.882	(0.783, 0.993)	0.039*
Race/Ethnicity	Non-Hispanic White	(reference)		
	Hispanic	0.716	(0.587, 0.872)	<0.001*
	Non-Hispanic Asian	0.855	(0.656, 1.114)	0.247
	Non-Hispanic Black	0.838	(0.679, 1.034)	0.099
	Non-Hispanic Other	0.784	(0.497, 1.237)	0.296
Region	Northeast	(reference)		
	Midwest	0.966	(0.782, 1.194)	0.75
	South	1.06	(0.883, 1.273)	0.532
	West	1.216	(0.992, 1.491)	0.061

Income to Poverty Ratio	Above 1	(reference)		
	Below or equal to 1	0.949	(0.703, 1.281)	0.733
Smoking Status	Never	(reference)		
	Current/Former	1.334	(1.173, 1.517)	<0.001*

Discussion

Poverty status showed no statistically significant relationship to repetitive strain injury. This result held in the data overall and was robust to controlling for age, sex, race/ethnicity, region, and smoking status. Because of the complex survey design, we can conclude that, based on this dataset, poverty status does not appear to be a significant predictor of repetitive strain injury among US civilians in the 50 states and DC.

Although no relationship was detected, the modeling approach has significant limitations which should be kept in mind as important context to the conclusion of no association. The model assumed independent effects from each covariate, neglecting the possibility of interaction effects. Additionally, no causal conclusion can be drawn from this model.

AI Statement

No AI was used in the generation of this report.

Data Availability

Data used in this study, as well as tables cited in the report, are available here.
https://docs.google.com/spreadsheets/d/1AKzp-NCeseX9DLGZ-v1VjUoHCTnISdUgKn1-nE5Axss/edit?usp=sharing

Complete analysis code is available here.
https://github.com/leo-ware/bios-620-final

References

2023 NHIS Questionnaires, Datasets, and Documentation | National Health Interview Survey | CDC. (2023). Retrieved May 7, 2025, from https://www.cdc.gov/nchs/nhis/documentation/2023-nhis.html

Lumley, T., Gao, P., & Schneider, B. (2024). *survey: Analysis of Complex Survey Samples* (Version 4.4-2) [Computer software]. https://cran.r-project.org/web/packages/survey/index.html

National Center for Health Statistics. National Health Interview Survey, 2023 survey description. (2024). Center for Disease Control. https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHIS/2023/srvydesc-508.pdf

Seidenberg, A. B., Moser, R. P., & West, B. T. (2023). Preferred Reporting Items for Complex Sample Survey Analysis (PRICSSA). *Journal of Survey Statistics and Methodology*, *11*(4), 743–757. https://doi.org/10.1093/jssam/smac040

Tulder, M. van, Malmivaara, A., & Koes, B. (2007). Repetitive strain injury. *The Lancet*, *369*(9575), 1815–1822. https://doi.org/10.1016/S0140-6736(07)60820-4