

# Occupational vs. Industry Sector Classification of the US Workforce: Which Approach is More Strongly Associated With Worker Health Outcomes?

Kristopher L. Arheart, EdD,<sup>1</sup> Lora E. Fleming, MD, PhD,<sup>1\*</sup> David J. Lee, PhD,<sup>1</sup> William G. LeBlanc, PhD,<sup>1</sup> Alberto J. Caban-Martinez, DO, PhD, MPH,<sup>1</sup> Manuel A. Ocasio,<sup>1</sup> Kathryn E. McCollister, PhD,<sup>1</sup> Sharon L. Christ, PhD, MS, MA,<sup>2,3</sup> Tainya Clarke, MS, MPH,<sup>1</sup> Diana Kachan,<sup>1</sup> Evelyn P. Davila, PhD,<sup>1</sup> and Cristina A. Fernandez, MSc<sup>1</sup>

**Objectives** Through use of a nationally representative database, we examined the variability in both self-rated health and overall mortality risk within occupations across the National Occupational Research Agenda (NORA) Industry Sectors, as well as between the occupations within the NORA Industry sectors.

**Methods** Using multiple waves of the National Health Interview Survey (NHIS) representing an estimated 119,343,749 US workers per year from 1986 to 2004, age-adjusted self-rated health and overall mortality rates were examined by occupation and by NORA Industry Sector.

**Results** There was considerable variability in the prevalence rate of age-adjusted self-rated poor/fair health and overall mortality rates for all US workers. The variability was greatest when examining these data by the Industry Sectors. In addition, we identified an overall pattern of increased poor/fair self-reported health and increased mortality rates concentrated among particular occupations and particular Industry Sectors.

**Conclusions** This study suggests that using occupational categories within and across Industry Sectors would improve the characterization of the health status and health disparities of many subpopulations of workers within these Industry Sectors. *Am. J. Ind. Med.* 54:748–757, 2011. © 2011 Wiley-Liss, Inc.

**KEY WORDS:** National Occupational Research Agenda (NORA); National Institute for Occupational Safety and Health (NIOSH)

<sup>1</sup>Department of Epidemiology and Public Health, Miller School of Medicine, University of Miami, Miami, Florida

<sup>2</sup>Department of Child Development and Family Studies, Purdue University, West Lafayette, Indiana

<sup>3</sup>Department of Statistics, Purdue University, West Lafayette, Indiana

Contract grant sponsor: National Institute for Occupational Safety and Health; Contract grant number: R01OH003915;

Contract grant sponsor: National Institute of Arthritis and Musculoskeletal and Skin Diseases; Contract grant number: F31 AR057687;

Contract grant sponsor: National Cancer Institute; Contract grant number: F31 CA153937.

\*Correspondence to: Lora E. Fleming, MD, PhD, Department of Epidemiology & Public Health, Clinical Research Building, Miller School of Medicine, University of Miami, Room 1049, 10th Floor (R-669), 1120 N.W. 14th Street, Miami, Florida 33136.  
E-mail: lfleming@med.miami.edu

Accepted 27 April 2011

DOI 10.1002/ajim.20973. Published online 13 June 2011 in Wiley Online Library (wileyonlinelibrary.com).

## INTRODUCTION

Occupation and industry classifications have been used for many years as surrogates for occupational exposure to identify and explore issues of worker morbidity and mortality [Monson, 1990; Checkoway, 2004]. This has ranged from the application of occupational/industry classification to national or regional data (particularly in Europe) as part of occupational health surveillance to identify worker subgroups at risk, to the use of specific job titles in longitudinal epidemiologic studies of a particular industry, union, or otherwise identified working population [Guralnick, 1962; Kaminski and Spirtas, 1980; Milham, 1983; Drever, 1995; NIOSH, 1997; Boffetta and Kogevinas, 1999; Fleming et al., 1999; Ma et al., 2006].

Although not without methodological issues (in particular, the misclassification of exposures), analyses using occupational/industry classification can be important for hypothesis generation (e.g., associations between a particular occupational/industry group with a specific health outcome, possibly related to a specific occupational exposure, which need further investigation); the identification of possible high risk worker groups; and with longitudinal data, the ability to observe changes over time in these high risk groups [Fleming et al., 1991; Carpenter, 1999; Bergdahl et al., 2004].

In 1996, the National Institute for Occupational Safety and Health (NIOSH) created the National Occupational Research Agenda (NORA) as a partnership program to stimulate occupational research and improve occupational safety and health in US workplaces [Anon, 2010]. Participation in the creation of the NORA was reportedly broad, including universities, large and small businesses, professional societies, government agencies, and worker organizations. The NORA priorities were based on: The number of workers at risk for a particular injury or illness, the seriousness of the hazard or issue, and the probability that new information and approaches would make a difference. The early NORA priorities focused on specific occupational diseases (e.g., occupational asthma) and specific worker groups (e.g., minority workers). Historically, research in the early NORA priority areas focused on occupational exposures and health effects.

In 2006, NIOSH refocused NORA to a new Industry Sector-based structure currently being developed and implemented through the NORA Industry Sector Councils [Anon, 2010]. With support from NIOSH, these Councils included participants from academia, industry, labor, non-governmental organizations, and government. Each Council created research agendas to provide guidance to the entire occupational safety and health community for moving research to practice in workplaces. In addition, there is a Cross-Sector Research Council to identify opportunities for common research across the Industry Sector Councils (e.g., surveillance, musculoskeletal disorders, health disparities, and safety culture).

The NORA Industry Sectors have created a new perspective for surveillance related to occupational exposures, health disparities, morbidity, and mortality among all US workers. Nevertheless, the use of industry rather than occupation as a surrogate for exposure may lead to even greater misclassification with respect to health outcomes. For example, it is possible that workers in the same occupation (e.g., cleaners) may experience different exposures and health effects across the different NORA Industry Sectors. On the other hand, within each NORA Industry Sector, there may be significant variability both in terms of exposures and health effects since each sector contains

a broad array of occupations (e.g., cleaners, mechanics, clerical, etc).

Self-rated health and overall mortality rates are common measures of morbidity and mortality in occupational and other populations since these data are often relatively easy to obtain, stable, and in the case of self-rated health, highly predictive of both individual morbidity and mortality [Idler and Benyamini, 1997; Singh-Manoux et al., 2007; Bidassie et al., 2010; Dowd and Zajacova, 2010]. In the present study, a nationally representative database was used to examine the variability in both self-rated health and overall mortality risk within occupations across the NORA Industry Sectors, as well as between the occupations within the NORA Industry sectors.

## METHODS

Since 1957, the National Health Interview Survey (NHIS) has been conducted annually by the National Center for Health Statistics (NCHS); the NHIS utilizes a complex sample survey design to obtain population-based strata that are representative of the US civilian non-institutionalized population [Fowler, 1996; Botman et al., 2000]. A mortality linkage was performed by the NCHS for the 1986–2004 NHIS survey participants with the National Death Index (NDI), with follow up through 2006. On average, the linkage was 97% complete; only those under 18 years and those with insufficient information to permit linkage to the NDI were considered ineligible and excluded from the analyses. For the purposes of this study, the public use NHIS mortality file was used, which has been found to provide very similar results when compared to the restricted NHIS-NDI linked data files [NCHS, 2005a; Lochner et al., 2008].

Households were selected by a multi-stage probability sampling strategy involving both clustering and stratification, in order to provide a representative sample of US adults [Massey, 1989]. During the 1986–1996 period, each week a probability sample of households was interviewed by trained personnel to obtain information about the characteristics of each member of the household [Liao et al., 1998]. In most cases, the participants themselves answered interview questions, and for the remaining participants, the responses were obtained from an adult responder in the household. Following incremental modifications in the 1995 and 1996 surveys, the NHIS was completely redesigned in 1997. The redesigned NHIS collects key health information from a single randomly selected adult household member. The response rate ranged between 95–98% during the early NHIS years [Massey, 1989]. The annual response rates to the 1997–2004 adult core have ranged from 72% in survey year 2004 to 80% in survey year 1997 [NCHS, 2005a, 2005b; Atrostic et al., 1999].

Detailed employment information coded by occupation and industry was collected by the NHIS on all subjects  $\geq 18$  years who reported working (paid and unpaid) during the two weeks before the 1986–1996 NHIS surveys and the week prior to the 1997–2004 NHIS surveys [NCHS, 1989; NCHS, 1992; NCHS, 2001; Caban et al., 2005]. Occupation was derived from detailed US Census occupation codes, and collapsed into four occupational status groups used by NCHS: White-collar, service, farm, and blue-collar workers [Krieger et al., 2005]. In addition, we also used 13 somewhat more detailed occupational groups using a NCHS NHIS recode based on the same US Census Occupational Codes (Table I). The eight NORA Industry Sectors, defined using the North American Industry Classification System (NAICS) which has replaced the U.S. Standard Industrial Classification (SIC) system (also based on US Census coding) include: (1) Agriculture, Forestry, Fishing (NAICS code: 11); (2) Construction (23); (3) Healthcare and Social Assistance (62); (4) Manufacturing (31–33); (5) Mining (21); (6) Services (51–56, 61, 71–72, 81, 92); (7) Transportation, Warehousing, Utilities (48–49, 22); and (8) Wholesale, Retail Trade (42, 44–45) [Anon, 2010].

## Statistical Analyses

We used SAS 9.2 (SAS Institute, Inc., Cary, NC) for all analyses in order to take into account sample weights and design effects. Sample weights were adjusted to account for the aggregation of data over the survey years 1986–2004, including the pooled prevalence rate data

**TABLE I.** National Center for Health Statistics (NCHS) Occupational Group Coding

<b>NCHS 4 Occupational Groups</b>	<b>NCHS 13 Occupational Groups</b>
White Collar	Executive, administrative managerial
	Professional specialty
	Technicians/related support
	Sales
	Administrative support occupations, including clerical
Service	Private household
	Protective service
	Service occupations, except protective/household
Farm	Farming, forestry, fishing
Blue Collar	Precision production, craft, repair
	Machine operators, assemblers, inspectors
	Transportation/material moving
	Handlers, equipment cleaners, helpers, laborers

Source: National Center for Health Statistics, Industry and Occupation Coding for Death Certificates 1993.

[Botman and Jack, 1995]. Prevalence rates of poor self-rated health (defined as a report of either fair or poor), both unadjusted and age-adjusted, were considered significantly higher than the overall sample prevalence rate if the subgroup rate was above the upper bound of the 95% confidence interval for the entire sample. This is a variation on the method of testing a one-sample difference in proportions considering the overall sample as the population proportion [Rosner, 2005]. Age-adjusted rates and the corresponding 95% confidence intervals presented are expressed in terms of deaths per 100,000 person years, and were computed by the “direct method” using the year 2000 US population (ages 18 and older) as the standard population. The latter was obtained from the master list of age-adjustment weights from the CDC’s Healthy People 2010 Statistical Notes 20 [Klein and Schoenborn, 2001].

To quantify the relative explanatory power of the eight NORA Industry Sectors, the four NCHS occupational groups, and the 13 NCHS recoded occupational groups, we reported the  $R^2$  attributable to each factor and the amount in common to both factors in the various analysis models. The  $R^2$  is interpreted as the percentage of variance in the dependent variable explained by the independent variables in the model. In this study, the  $R^2$  is interpreted as the strength of the association between the exposures defined by the NORA, NCHS, and occupational categories and the health outcomes; larger  $R^2$  values indicate higher explanatory power.

We used SAS PROC SURVEYLOGISTIC to analyze self-reported health. Five separate models were run: (1) The 4 NCHS occupational groups alone; (2) the 13 occupational groups alone; (3) the 8 NORA Industry Sectors alone; (4) the 4 NCHS occupational groups and the eight NORA industry sectors; and (5) the 13 occupational groups and the 8 NORA industry sectors. The  $R^2$  reported for these models were then arithmetically manipulated to obtain the  $R^2$  attributable to the separate factors. The common  $R^2$  for the NCHS occupational groups and the NORA industry sectors is the sum of the  $R^2$  for models one and three minus the  $R^2$  for model four. The  $R^2$  attributable to the NCHS occupational groups is the  $R^2$  for model one minus the common  $R^2$ , and the  $R^2$  for the NORA industry sectors is the  $R^2$  for model three minus the common  $R^2$ . The same procedure was used for the 13 occupation groups and the NORA industry sectors. For the mortality rates, SAS PROC GLM was used to compute the individual  $R^2$ . An analysis of variance weighted by the inverse of the variance for each group’s annual mortality rate was run. Two models were evaluated: (1) The NCHS occupational groups and the NORA industry sectors; and, (2) the 13 occupational groups and the NORA Industry Sectors. The sums of squares (SS) were used to calculate the common  $R^2$  [(Model SS–NCHS Type III SS–NORA

Type III SS)/Total SS], the NCHS  $R^2$  [NCHS Type III SS/Total SS], and the NORA  $R^2$  [Type III SS/Total SS]. The TYPE III SS for each of the factors was adjusted (independent) of any other factor in the model. The same procedure was used for the 13 occupational groups and the NORA industry sectors model.

## RESULTS

For the pooled analyses from 1986–2004, there were a total of 673,888 adult worker participants in the NHIS, representing an estimated 119,343,749 US workers per year. Among all workers, 5.5% (unadjusted) and 6.6% (age-adjusted) reported poor or fair self-rated health, and the overall mortality rate was 598.0/100,000 person years [py] (with 95% confidence interval: 584.2–601.8/100,000 py).

When reviewing the pooled prevalence rate of reported poor/fair self-rated health, the age adjustment generally increased these estimates. Therefore, in the discussions below, only the age-adjusted prevalence rates of poor/fair self-rated health are described, although both age-adjusted and unadjusted prevalence rates are presented (Table II). The age-adjusted mortality rates are also presented (Table III).

### Within Occupations Across the NORA Industry Sectors

#### *Self-rated health (Table II)*

For the four NCHS occupational groups, the pooled age-adjusted prevalence rates of poor/fair self-rated health were highest and varied across the NORA Industry Sectors most dramatically for the Service workers (4.7–12.4%) and Farm workers (3.5–10.2%), with the least variability and lowest prevalence rate seen among the White collar workers (3.4–6.5%). Somewhat similar patterns were seen among the 13 occupational groups, particularly for those workers reporting low prevalence rates of poor/fair self-rated health. Machine operators, assemblers, inspectors (7.6–15.7%) and handlers, equipment cleaners, laborers (5.6–15.2%) had the highest prevalence rates, which also varied the most dramatically across the NORA Industry Sectors. Private household workers were also high in terms of the prevalence rate of reported poor/fair self-rated health (9.1–14.4%), but not in terms of variability due to the lack of this occupation across all the NORA Industry Sectors. The more professional white-collar occupations had a lower prevalence rate and low variability across the NORA Industry Sectors (e.g., professional specialty workers [3.0–6.2%]).

### *Overall mortality rates (Table III)*

For the four NCHS occupational groups, the age-adjusted overall mortality rates were highest and varied across the NORA industry sectors, most dramatically for the service workers (526.9–967.6/100,000 py) and farm workers (440.1–1007.3/100,000 py), with the least variability and lowest rates found among the white-collar workers (441.5–599.6/100,000 py). Similar patterns were seen among the 13 occupational groups. In addition to the expected high mortality rates and variability across NORA Industry Sectors seen among Protective services workers (666.9–1557.0/100,000 py), high rates and variability were also found among farming, forestry, fishing workers (394.4–1007.0/100,000 py) and handlers, equipment cleaners, and laborers (659.2–1161.0/100,000 py). Private household workers were also high in terms of overall mortality rates (500.3–1315.0/100,000 py), but not in terms of variability due to the lack of this occupation across all the NORA industry sectors. Again, the more professional white-collar occupations had low age-adjusted overall mortality rates and low variability across the NORA Industry Sectors (e.g., professional specialty workers [408.8–690.1/100,000 py]).

### Between the Occupations Within the NORA Industry Sectors

#### *Self-rated health (Table II)*

For the four NCHS occupational groups, the pooled age-adjusted prevalence rates of poor/fair self-rated health were highest and most varied between the occupations within each NORA Industry Sector for the Manufacturing Sector (4.3–12.4%), the Construction Sector (5.1–13.6%), and the Healthcare/Social Assistance Sector (3.5–11.1%). Although not as high or as variable, there was at least a 4% difference in the prevalence rate of poor/fair self-rated health among the four NCHS occupational groups within the other NORA Industry Sectors, except for Wholesale/Retail trade sector (6.5–9.9%). Within each NORA Industry Sector, in general the lowest prevalence rates were concentrated in the white-collar workers and the highest prevalence rates among the Service and Farmers Industry Sectors. Using the 13 occupational groups, the Healthcare/Social Assistance Sector (3.3–15.7%), the Mining Sector (3.0–14.4%), and the Construction Sector (3.0–13.6%) had the highest report of poor/fair self-rated health and the greatest variability between the occupations within each NORA Industry Sector. There was at least 7% difference in the prevalence rate of reported poor/fair self-rated health among the 13 occupational groups within the other NORA Industry Sectors, except for Wholesale/Retail Trade Sector (6.1–9.9%). In general, within each NORA

**TABLE II.** Unadjusted and Age-Adjusted Pooled Prevalence Rates of Poor/Fair Self-Rated Health by NORA Industry Sector and Occupational Group Among US Workers in the 1986–2004 National Health Interview Survey

Occupational Categories	NORA Industry Sector																													
	Total	Agriculture, Forestry, Fishing				Healthcare/Social Assistance				Manufacturing				Mining				Services				Transportation, Warehousing, Utilities				Wholesale/Retail Trade				
		Self-rated Health	N	Fair <sup>a</sup>	% Poor/ Fair <sup>a</sup>	Self-rated Health	N	Fair <sup>a</sup>	% Poor/ Fair <sup>a</sup>	Self-rated Health	N	Fair <sup>a</sup>	% Poor/ Fair <sup>a</sup>	Self-rated Health	N	Fair <sup>a</sup>	% Poor/ Fair <sup>a</sup>	Self-rated Health	N	Fair <sup>a</sup>	% Poor/ Fair <sup>a</sup>	Self-rated Health	N	Fair <sup>a</sup>	% Poor/ Fair <sup>a</sup>	Self-rated Health	N	Fair <sup>a</sup>	% Poor/ Fair <sup>a</sup>	
Overall Total	673,888	5.5	5,45	5.63	19,115	7.5	43,203	5.6	80,869	5.9	116,980	6.0	3,872	6.2	235,481	5.0	48,196	5.1	126,172	5.7	7.3									
NCHS Occupational Categories																														
White Collar	394,138	4.1	4.06	4.23	2,484	4.2	10,639	4.3	56,344	4.1	46,942	3.3	1,676	3.9	169,205	3.8	25,464	4.2	81,384	5.3	6.5									
Service	90,673	8.6	8.34	8.84	207	8.1	214	11.2	21,856	10.2	19,49	11.6	59	5.6	42,455	8.7	1,552	6.4	22,381	6.9	9.9									
Farm Worker	18,594	8.1	7.50	8.64	15,589	8.1	125	7.1	152	4.7	619	9.6	3	0.0	1,732	8.2	53	5.4	321	6.5	6.6									
Blue Collar	170,483	7.0	6.82	7.16	835	8.3	32,225	5.9	2,517	9.6	67,470	7.7	2,134	8.0	22,089	8.0	21,127	6.0	22,086	6.3	8.4									
13 NHIS Occupational Categories																														
Executive, administrative managerial	93,384	3.7	3.59	3.89	635	4.6	6,489	4.3	7,813	4.1	15,230	2.9	622	3.3	42,668	3.6	6,103	3.8	13,824	4.6	6.1									
Professional specialty	99,968	3.1	3.01	3.28	788	3.8	954	3.5	27,574	3.7	10,175	2.4	400	3.1	54,406	2.9	3,001	2.4	2,670	4.8	6.2									
Technicians/related support	23,772	3.9	3.62	4.24	251	4.5	378	3.1	8,104	4.4	4,251	3.5	185	3.9	7,642	3.6	1,966	3.7	995	4.7	6.6									
Sales	75,457	5.0	4.81	5.23	138	3.2	453	5.5	4,666	2.5	4,837	2.8	67	5.5	16,675	4.5	1,643	3.3	51,178	5.5	6.7									
Administrative support occs, including clerical	101,557	5.0	4.84	5.17	672	4.4	2,365	4.9	12,387	5.1	12,449	4.9	402	5.2	47,814	4.9	12,751	5.1	12,717	5.4	5.4									
Private household	4,827	14.4	13.13	15.67	0	.	4	0.0	103	15.7	5	0.0	0	.	4,695	14.4	0	.	20	6.8	9.1									
	14.4	13.21	15.76				0.0		14.4						14.5															

(Continued)



**TABLE III.** Age-Adjusted Overall Mortality Rates/100,000 Person Years by NORA Industry Sector and Occupational Group Among US Workers of the 1986–2004 National Health Interview Survey

Occupational Categories	NORA Industry Sector														
	Total	Agriculture, Forestry, Fishing		Healthcare/Social Assistance		Manufacturing		Mining		Services		Transportation Warehousing, Utilities		Wholesale/Retail Trade	
		Mortality Rate <sup>a</sup>	Mortality Rate <sup>a</sup>	Mortality Rate <sup>a</sup>	Mortality Rate <sup>a</sup>	Mortality Rate <sup>a</sup>	Mortality Rate <sup>a</sup>	Mortality Rate <sup>a</sup>	Mortality Rate <sup>a</sup>	Mortality Rate <sup>a</sup>	Mortality Rate <sup>a</sup>	Mortality Rate <sup>a</sup>	Mortality Rate <sup>a</sup>	Mortality Rate <sup>a</sup>	Mortality Rate <sup>a</sup>
Overall Total	673,888	593.0	584.2	601.8	797.4	714.4	489.9	577.9	636.3	558.7	619.3	639.2			
NCHS Occupational Categories															
White Collar	394,138	509.3	500.0	518.8	492.3	571.4	441.5	500.2	528.5	485.4	515.4	599.6			
Service	90,673	683.5	662.0	705.7	526.9	967.6	591.0	866.2	649.4	702.7	799.4	720.6			
Farm Worker	18,594	846.7	804.9	890.6	841.1	1007.3	652.6	920.1	0.0	926.6	440.1	394.4			
Blue Collar	170,483	708.8	691.9	726.1	808.3	773.7	672.7	622.5	749.2	793.4	725.9	747.5			
13 NHIS Occupational Categories															
Executive, administrative managerial	93,384	528.0	511.0	545.6	594.0	595.8	439.0	480.4	372.2	536.3	476.6	591.7			
Professional specialty	99,968	434.0	418.8	449.7	408.8	473.4	448.5	456.9	690.1	415.2	525.5	487.2			
Technicians/related support	23,772	446.8	412.8	483.6	677.1	707.2	395.3	403.9	1032.0	484.9	505.1	551.6			
Sales	75,457	603.9	583.6	635.0	554.4	742.3	644.9	572.3	747.4	570.9	579.5	617.9			
Administrative support occs, including clerical	101,557	490.9	475.0	507.4	463.5	496.7	427.2	542.8	464.5	466.8	527.5	556.3			
Private household	4,827	652.1	596.6	712.6		0.0	500.3	1315.0		654.4		669.9			
Protective service	11,734	910.5	836.6	990.8	666.9	1253.0	1557.0	930.0	0.0	884.7	1216.0	942.9			
Service occs, except protective/Household	74,112	651.5	628.6	675.2	553.5	952.4	575.7	852.5	908.6	662.1	713.1	710.7			
Farming, forestry, fishing	18,594	846.7	804.9	890.6	841.1	1007.0	652.6	920.1	0.0	926.6	440.1	394.4			
Precision production, craft, repair	73,107	685.8	660.3	712.3	775.6	753.1	598.6	610.5	764.4	744.4	572.2	675.2			
Machine operators, assemblers, inspectors	44,584	616.0	587.0	646.5	526.6	595.4	649.0	597.5	560.3	744.3	586.3	610.1			
Transportation/material moving	28,115	826.1	783.3	871.2	814.2	862.9	810.1	807.4	787.3	823.4	824.0	828.1			
Handlers, equipment cleaners, helpers, laborers	24,677	820.9	772.0	873.0	1093.0	838.5	797.5	666.8	659.2	1161.0	773.2	826.9			

<sup>a</sup>Overall mortality rates/100,000 person year (py).

Industry Sector, the lowest prevalence rates were concentrated in the white-collar and professional workers, and the highest prevalence rates among the machine operators, assemblers, inspector workers and the transportation, material moving workers.

### **Overall mortality rates (Table III)**

For the four NCHS occupational groups, the highest age-adjusted overall mortality rates and greatest variability by occupation were seen in the Construction sector (571.4–1007.3/100,000 py) and the Services sector (485.4–926.6/100,000 py). In general, there was at least a 300/100,000 py difference between the four NCHS occupational categories within each NORA industry sector. In general, the low mortality rates were concentrated among the white-collar workers and the high mortality rates among the farm workers and blue-collar workers. Using the 13 occupational groups, the highest within the NORA Industry Sector mortality rates and greatest variability across occupational group were seen for the Healthcare/Social Assistance Sector (395.3–1557.0/100,000 py) and the Manufacturing sector (403.9–1315.0/100,000 py). With the exception of the Wholesale/retail trade (487.2–942.9/100,000 py), there was at least a 700/100,000 py difference in the mortality rates between the occupations within the NORA Industry Sectors. The lowest mortality rates were concentrated in the white-collar and professional workers, and the highest among the protective service workers and the handler, equipment cleaner, helper, laborer workers.

### **Explanatory power**

The overall  $R^2$  for the NORA-NCHS analysis of self-rated health was 0.64, suggesting a significant amount of variance was explained by both constructs. The individual  $R^2$  values were 0.01 for NORA, 0.56 for NCHS, and 0.07 in common; the four NCHS occupational categories accounted for the major portion of the explained variance, while the amount explained by NORA industry categories was negligible. The overall  $R^2$  for the NORA-NCHS analyses of the mortality rates was 0.92, indicating a strong association between both factors and mortality rates. The individual  $R^2$  for NORA was 0.19, for NCHS 0.57, and in common 0.16. The four NCHS occupational categories explained more of the variation than did the NORA industry categories.

The overall  $R^2$  for the NORA-13 occupations analysis of self-rated health was 0.74, suggesting a significant amount of variance was explained by both constructs. The individual  $R^2$  values were 0.01 for NORA, 0.66 for the 13 occupations, and 0.07 in common. The 13 occupational groups accounted for the major portion of explained

variance, while the amount explained by NORA industry categories was negligible. The overall  $R^2$  for the NORA-13 occupations analyses of the mortality rates was 0.89, indicating a strong association between both factors and mortality rates. The individual  $R^2$  for NORA was 0.05, for the 13 occupations 0.59, and in common 0.25; again, the 13 occupational groups explained more of the variation than did the NORA industry categories.

## **DISCUSSION**

As others have found in the US and elsewhere, we found considerable variability in the prevalence rates of age-adjusted self-rated poor/fair health and in the age-adjusted overall mortality rates in a representative sample of all US workers from 1986–2004 using a nationally representative database [Guralnick, 1962; Kaminski and Spirtas, 1980; Milham, 1983; Drever, 1995; NIOSH, 1997; Boffetta and Kogevinas, 1999]. The variability was significant and greatest when examining these data within the NIOSH NORA Industry Sectors by occupation. In addition, we identified an overall pattern of increased poor/fair self-reported health and increased mortality rates concentrated in particular occupations and in particular NORA Industry Sectors.

There are several implications of these findings with respect to the variability in health outcomes by occupation and by NORA Industry Sector. Currently, the NORA Industry Sector Councils are gathering and analyzing sector-specific data on the exposures and health outcomes of workers between and within each of the eight NORA Industry Sectors. This study suggests that to decrease possible exposure misclassification and inadequate characterization of the health status of many subpopulations of workers within these industry sectors, it would be important to further investigate health outcomes by occupation within each of the NORA Industry Sectors.

Only relatively recently has a health disparities focus been specifically applied to occupational health in the US [Murray, 2003; Caban et al., 2005; Krieger et al., 2005]. This study identified significant health disparities when examined by both occupation and the NORA Industry Sectors with respect to health outcomes. By occupation, the general pattern is that white collar/professional workers have substantially better self-reported health and lower mortality rates compared to other occupational groups, particularly compared to farming, forestry, fishing workers, handlers, equipment cleaners, and laborers, protective service workers, private household workers, and machine operators, assemblers, inspectors. Within the NORA industry sectors, the Healthcare/Social Assistance sector, the Construction sector, and the Manufacturing sector were more likely to have the highest and greatest variability of poor/fair self-reported health and overall mortality rates.

Therefore, these findings substantiate that workers of lower socio-economic class are more likely to suffer from poor health.

## Limitations

Although the death linkage data used in this study to create the mortality rates provide a longitudinal view of all US workers who participated in the NHIS between 1986 and 2004 with follow up through 2006, the NHIS data are collected cross-sectionally. Therefore, it is not possible to explore with these data whether working in certain occupations leads to poorer health, whether those with poorer health are more likely to be employed in certain occupations, or whether (as is most likely) there is some combination of factors. This analysis was not intended to demonstrate causation, but rather to explore the relative explanatory power of using occupational and industry coding as a surrogate for exposure, and the possibility of misclassification bias when using surrogate measures with low explanatory power. Finally, the data used for the poor/fair health assessment are self-reported; however, as discussed above, self-report of health has been shown in multiple studies to be a robust predictor not only of future health, but also of future mortality [Idler and Benyamini, 1997; Singh-Manoux et al., 2007; Bidassie et al., 2010; Dowd and Zajacova, 2010].

## CONCLUSIONS

The NHIS data provide a unique opportunity to examine various aspects of US worker health, including morbidity and mortality, both cross-sectionally and longitudinally, using a representative sample of the entire US population. This study provides evidence of significant health disparities with respect to self-rated health and mortality rates in certain occupations and NORA Industry Sectors. However, these analyses also support the need to examine US worker health not only by NORA Industry Sector, but by occupation both within and across these Industry Sectors to truly understand these health disparities in the US workplace.

## REFERENCES

- About NORA. Partnerships, Research and Practice 2010. <http://www.cdc.gov/niosh/nora/about.html>. Accessed September 2010.
- Atrostic B, Bates N, Burt G, Silverstein A, Winters F. Nonresponse in federal household surveys: New measures and new insights. 1999. [http://www.fcsn.gov/committees/ihsng/portland\\_3\\_\\_120299.pdf](http://www.fcsn.gov/committees/ihsng/portland_3__120299.pdf). Accessed 7-30-04, 2004.
- Bergdahl IATK, Eriksson K, Hedlund U, Nilsson T, Flodin R, Jarvholm B. 2004. Increased mortality in COPD among construction workers exposed to inorganic dust. *Eur Resp J* 23(3):402–406.
- Bidassie B, McGlothlin JD, Mena I, Duffy VG, Barany JW. 2010. Evaluation of lifestyle risk factors and job status associated with back injuries among employees at a mid-western university. *Appl Ergon* 41(1):106–114.
- Boffetta P, Kogevinas M. 1999. Occupational cancer in Europe. *Environ Health Perspect* 107 (Suppl 2):227.
- Botman SL, Jack SS. 1995. Combining National Health Interview Survey Datasets: Issues and approaches. *Stat Med* 14(5–7):669–677.
- Botman SL, Moore TF, Moriarty CL, Parsons VL. 2000. Design and estimation for the National Health Interview Survey, 1995–2004. *Vital Health Stat* 2(130):1–31.
- Caban AJ, Lee DJ, Fleming LE, Gomez-Marin O, LeBlanc W, Pitman T. 2005. Obesity in US workers: The National Health Interview Survey, 1986–2002. *Am J Public Health* 95(9):1614–1622.
- Carpenter LRE. 1999. Cancer and occupation in women: Identifying associations using routinely collected national data. *Environ Health Perspect* 107(Supplement 2):299–303.
- Checkoway H. 2004. *Research methods in occupational epidemiology*. 2nd edition. New York: Oxford University Press.
- Dowd JB, Zajacova A. 2010. Does self-rated health mean the same thing across socioeconomic groups? Evidence from biomarker data. *Ann Epidemiol* 20(10):743–749.
- Drever F. 1995. The registrar general's decennial supplement for England & Wales. Vol. 10. London: Office of Population Censuses and Surveys Health and Safety Executive, Series DS.
- Fleming LE, Ducatman AM, Shalat SL. 1991. Disease Clusters: a Central and Ongoing Role in Occupational Medicine. *J Occ Med* 33:818–825.
- Fleming LE, Bean JA, Rudolph M, Hamilton K. 1999. Mortality in a cohort of licensed pesticide applicators in Florida. *Occup Environ Med* 56(1):14–21.
- Fowler FJ Jr. 1996. The redesign of the National Health Interview Survey. *Public Health Rep* 111(6):508–511.
- Guralnick L. 1962. Mortality by occupation and industry among men 20–64 years of Age: US, 1950. Vol. 532. Washington, DC: NCHS.
- Idler EL, Benyamini Y. 1997. Self-rated health and mortality: A review of twenty-seven community studies. *J Health Soc Behav* 38(1):21–37.
- Kaminski R, Spirtas R. 1980. Industrial characteristics of persons reporting morbidity during health interview surveys conducted in 1969–1974. Cincinnati, Ohio: National Institute of Occupational Safety and Health.
- Klein RJ, Schoenborn CA. 2001. Age adjustment using the 2000 projected US population. *Healthy People 2010 Statistical Notes*, Number 20. <http://www.cdc.gov/nchs/data/stant/statnt/statnt10.pdf>. Accessed September 3, 2004.
- Krieger N, Barbeau EM, Soobader MJ. 2005. Class matters: U.S. versus U.K. measures of occupational disparities in access to health services and health status in the 2000. U. S. National Health Interview Survey. *Int J Health Serv* 35(2):213–236.
- Liao Y, Cooper RS, Cao G, Durazo-Arvizu R, Kaufman JS, Luke A, McGee DL. 1998. Mortality patterns among adult hispanics: Findings from the NHIS, 1986–1990. *Am J Public Health* 88(2):227–232.
- Lochner K, Hummer RA, Bartee S, Wheatcroft G, Cox C. 2008. The public-use National Health Interview Survey linked mortality files: Methods of reidentification risk avoidance and comparative analysis. *Am J Epidemiol* 168(3):336–344.

- Ma F, Fleming LE, Lee DJ, Trapido E, Gerace TA. 2006. Cancer Incidence in Florida Professional Firefighters, 1981 to 1999. *JOEM* 48:883–888.
- Massey JT. 1989. National Center for Health Statistics (U.S.). *Design and estimation for the National Health Interview Survey, 1985-94*. Hyattsville, Md.: U.S. Dept. of Health and Human Services, Public Health Service, Centers for Disease Control, National Center for Health Statistics.
- Milham S. 1983. Occupational mortality in Washington state 1950–79. Cincinnati, OH: NIOSH, Dept HEW #83-116.
- Monson R. 1990. Occupational epidemiology. 2nd edition. Boca Raton: CRC Press.
- Murray LR. 2003. Sick and tired of being sick and tired: Scientific evidence, methods, and research implications for racial and ethnic disparities in occupational health. *Am J Public Health* 93(2):221–226.
- NCHS. 1989. Classified index of industries and occupations, 1989. Washington, DC: US Department of Health and Human Services.
- NCHS. 1992. Industry and occupation coding for death certificates 1993. Hyattsville, MD: Public Health Service.
- NCHS. 2001. Documentations for the NHIS Surveys 2001. In: US Dept of Health and Human Services N. ed. Washington, DC: National Center for Health Statistics (NCHS).
- NCHS. 2005a. The 1986–2000 National Health Interview Survey linked mortality files: Matching methodology, October 2005. National Center for Health Statistics. Office of Analysis and Epidemiology [http://www.cdc.gov/nchs/data/datalinkage/matching\\_methodology\\_nhis\\_final.pdf](http://www.cdc.gov/nchs/data/datalinkage/matching_methodology_nhis_final.pdf) (accessed 11-15-07).
- NCHS. 2005b. Summary health statistics for the U. S. adults: National Health Interview Survey, 2004. *Vital Health Stat* 10:228.
- NIOSH. 1997. Mortality by occupation, industry and cause of death: 24 reporting states (1984–88). Cincinnati, OH: NIOSH #1997-114.
- Rosner BA. 2005. Fundamentals of biostatistics, 6th edition. Pacific Grove, CA: Brooks.
- SAS Institute I. 2008. SAS 9.2 Software. *Products and Solutions* <http://www.sas.com/software/sas9/-section=1>.
- Singh-Manoux A, Gueguen A, Martikainen P, Ferrie J, Marmot M, Shipley M. 2007. Self-rated health and mortality: Short- and long-term associations in the Whitehall II study. *Psychosom Med* 69(2): 138–143.