LONG WORK HOURS AND ADVERSE HEALTH OUTCOMES:

DEFINING A HEALTH RISK-BASED THRESHOLD

FOR WORK HOUR EXPOSURE

by

SARAI CAROLYN HARRIS CONWAY, MA

APPROVED:

LISA ANNEMARIE POMPEII, PHD

DAVID GIMENO RUIZ DE PORRAS, PHD

ROBERT EDMUND ROBERTS, PHD

DEAN, THE UNIVERSITY OF TEXAS
SCHOOL OF PUBLIC HEALTH

DEDICATION

To Craig and Harrison

LONG WORK HOURS AND ADVERSE HEALTH OUTCOMES:

DEFINING A HEALTH RISK-BASED THRESHOLD

FOR WORK HOUR EXPOSURE

by

SARAI CAROLYN HARRIS CONWAY
BA, Vanderbilt University, 1996
MA, Syracuse University, 2002

Presented to the Faculty of The University of Texas

School of Public Health

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS
SCHOOL OF PUBLIC HEALTH
Houston, Texas
May, 2015

ACKNOWLEDGEMENTS

This project could not have been completed without the guidance of Lisa Pompeii, David Gimeno, and Robert Roberts; the input of Jack Follis; and the intellectual and moral support of Craig Conway, Harrison Conway, Erin Morrison, Renata Benjamin and Laura Lee.

LONG WORK HOURS AND ADVERSE HEALTH OUTCOMES:

DEFINING A HEALTH RISK-BASED THRESHOLD

FOR WORK HOUR EXPOSURE

Sarai Carolyn Harris Conway, MA, PhD
The University of Texas
School of Public Health, 2015

Dissertation Chair: Lisa Pompeii, PhD

Long work hours (LWH) have been studied as an independent risk factor for adverse health conditions for more than 50 years. However, no previous study has sought to determine the threshold of work hours beyond which there was evidence of increasing risk of poor health or to describe the functional form of work hours and the risk of cardiovascular disease.

Data from the Panel Study of Income Dynamics were used to: 1) identify the LWH cut points that best predicted increased risk of poor self-reported general health (SRGH), incident cardiovascular disease (CVD), and incident cancer; and 2) evaluate the presence of a dose-response relationship between work hour duration and the risk of CVD.

Regarding the selection of the statistically optimized cut points, this study suggests that the work hour threshold of 52 hours of work per week or more, on average, over a minimum of 10 years best predicted increased risk of all three outcomes of interest: poor SRGH (RR: 1.28; 95% CI: 1.06−1.53), incident CVD (RR: 1.42; 95% CI: 1.24−1.63), and incident cancer (RR: 1.62; 95% CI: 1.22−2.17). On the question of a dose-response relationship between work hours and CVD, increasing work hour durations were associated

with increased risk of incident CVD across a set of nested, hierarchical logistic regression models. The restricted cubic spline model was determined to best fit the data, the results of which suggested a 33% increase in risk of CVD among individuals working 55 hours per week, on average, for a minimum of 10 years compared to those working 45 hours per week, on average, for the same duration.

This project is not only the first to identify statistically optimized exposure cut points for LWH that are based on the evidence of health risk, but also offers the first characterization of the dose-response relationship between work hour duration and CVD. These results indicate that LWH could be recognized as a potential risk indicator for poor SRGH, CVD, and cancer, thereby contributing to the growing body of evidence regarding LWH and adverse health outcomes.

TABLE OF CONTENTS

List of Tables i

List of Figures iii

List of Appendices iv

Chapter I: Background 1
- Long Work Hours: Definitions and National Regulations 1
- Time Trends in Hours Worked 4
- Long Work Hours and Adverse Health Outcomes 8
- Long Work Hours and Self-Reported General Health 13
- Potential Confounding and Moderating Factors 29
- Proposed Mechanisms of Effect for Long Work Hours 31
- Conceptual Framework for the Association between Long Work Hours and Adverse Health Effects 32
- Work Hours as an Occupational Risk Factor 36
- Long Work Hours Exposure Misclassification 38
- Public Health Significance 42
- Specific Aims 43

Chapter II: Methods 45
- The Panel Study of Income Dynamics − Parent Study 45
  - Overview 45
  - Sampling, Tracking, and Response 45
  - Informed Consent 48
  - Data Collection 48
- The Current Study 49
  - Overview 49
  - Study Population and Inclusion/ Exclusion Criteria 49
  - Study Variables 52
- Data Analysis 59
  - Hypothesis 1 59
  - Hypothesis 2 65
- Study Power 69
- Strengths and Limitations 71
- Human Subjects Considerations 72

Chapter III: The Identification of a Long Work Hours Cut Point for Predicting Elevated Risk of Poor Self-Reported General Health 74
- Name of Journal Proposed for Article Submission: Epidemiology 74
- Abstract 74
- Background 75
- Methods 76

Study Population....76
Data Collection....77
Statistical Analysis....78
Results....82
Discussion....83
Acknowledgements....87
References....89

Chapter IV: Application of a Long Work Hours Cut Point for Predicting Elevated Risks of Cardiovascular Disease and Cancer....97
Name of Journal Proposed for Article Submission: Epidemiology....97
Abstract....97
Background....98
Methods....100
Study Population....100
Data Collection....101
Statistical Analysis....103
Results....105
Discussion....107
Acknowledgements....109
References....113

Chapter V: Dose-Response Relation between Work Hours and Caridovascular Disease: Findings from the Panel Study of Income Dynamics....119
Name of Journal Proposed for Article Submission: Journal of Occupational and Environmental Medicine....119
Abstract....119
Background....120
Methods....121
Study Population....121
Data Collection....122
Statistical Analysis....123
Results....126
Discussion....128
Acknowledgements....131
References....137

Chapter VI: Conclusions....145

Appendices....147

References....150

LIST OF TABLES

Chapter I: Background

Table 1. Work hours in a sample of European Union countries, 2010 3

Table 2: Increases in U.S. work hours, 1979-2007 6

Table 3. Distribution of sociodemographic and occupational characteristics by weekly work hours in adult, working participants in the Panel Study of Income Dynamics, 2011 wave (N=10,194) 9

Table 4. Evidence of the strength of association between long work hours and several health-related outcomes in 53 selected studies, 1996-2013..................................11

Table 5. Studies of LWH and SRGH...16

Chapter II: Methods

Table 1. PSID working status and work hour measures, 1986-2011*....................................53

Table 2. PSID self-reported general health measure, 1986-2011* 55

Table 3. Construction of primary outcome variable 56

Table 4. PSID specific physical health measures, 1986-2011* 57

Table 5. Statistical power of measures of association between LWH and the outcomes of interest 70

Chapter III: The Identification of a Long Work Hours Cut Point for Predicting Elevated Risk of Poor Self-Reported Health

Table 1. Measures[a] of model fit, calibration and model discrimination for long work hours and self-reported general health: Panel Study of Income Dynamics, 1986-2011. 88

Chapter IV: Application of a Long Work Hours Cut Point for Predicting Elevated Risks of Cardiovascular Disease and Cancer

Table 1. Measures[a] of model fit, calibration and model discrimination for long work hours and incident coronary heart disease: Panel Study of Income Dynamics, 1986-2011. 110

Table 2. Measures[a] of model fit, calibration and model discrimination for long work hours and incident cancer: Panel Study of Income Dynamics, 1986-2011. 111

Table 3. General description of the combined study populations by selected characteristics at study baseline (1986) by average work hours per week: Panel Study of Income Dynamics, 1986-2011. 112

Chapter V: Dose-Response Relation between Work Hours and Cardiovascular Disease: Findings from the Panel Study of Income Dynamics

Table 1. General description of the study population by selected characteristics at study baseline (1986) by outcome category: Panel Study of Income Dynamics, 1986-2011. 133

Table 2. Association between work hour duration and incident cardiovascular disease: Panel Study of Income Dynamics, 1986-2011. 134

Table 3. Association between work hour duration and incident cardiovascular disease in part-time workers (<35 work hours per week, on average) and full-time (≥35 work hours per week, on average): Panel Study of Income Dynamics, 1986-2011. 135

LIST OF FIGURES

Chapter I: Background

Figure 1. Historical trends in total annual work hours per worker in selected countries, 1870-2000....5

Figure 2. Historical trends in average weekly work hours per U.S. worker, 1830-1980 5

Figure 3: Typical number of weekly work hours reported by U.S. workers, 2001-2014....7

Figure 4. The individual health evaluation process to determine SRGH status 14

Figure 5. Conceptual framework for studying LWH....34

Chapter II: Methods

Figure 1. Diagram of the PSID sampling frame 46

Figure 2. Construction of CVD incidence multi-year summary variable....58

Figure 3. Construction of incident cancer multi-year summary variable 60

Figure 4: ROC curve illustrating Youden Index J for cut point c....64

Figure 5. Proposed directed acyclic graph for exploring the relationship between LWH and illness 67

Chapter V: Dose-Response Relation between Work Hours and Cardiovascular Disease: Findings from the Panel Study of Income Dynamics

Figure 1. Restricted cubic spline model for the relationship between work hour duration and incident cardiovascular disease: Panel Study of Income Dynamics, 1986-2011. 136

LIST OF APPENDICES

Appendix A: University of Texas Health Science Center Institutional Review Board Outcome Notices..147

CHAPTER I: BACKGROUND

Long Work Hours: Definitions and National Regulations

Long work hours (LWH) refers to daily or weekly durations of work exceeding a threshold of work hours generally defined in terms of the standard work hours of a full-time employee.[1] Assuming the threshold of standard work hours is exceeded, LWH encompasses all productive activities and all jobs in paid employment (including self-employment) and, generally, is restricted to activities in the formal economy. As such, it is a measure of time actually worked by an employee and is, by definition, demarcated by the fact that its duration exceeds the collective prevailing work hours.[2]

A distinction is made between LWH, overtime work, and shift work, as the latter two are specific to contractual or scheduled hours of work, respectively.[2] *Overtime work* (or "overtime") is characterized as work hours that exceed the contracted amount of work time per week defined in a formal agreement and applies to part-time or full-time employees designated as overtime eligible by their employers, local guidelines, or national mandates.[3,4] By definition, overtime equates to LWH, but LWH is not necessarily overtime. This is because a worker must exceed a pre-determined number of work hours in a week to reach overtime, whereas patterns of daily overwork that sum to fewer hours than the overtime threshold could be considered LWH (e.g., if LWH is defined as more than 8 hours of work per day, and overtime is defined as exceeding 40 work hours in a week, a four-day workweek of 10 work hours per day would be considered LWH but not overtime).[1] *Shift work* is a work schedule pattern that diverges from standard daytime hours by involving unusual or irregular hours, such as night work or rotating work periods.[5] Although a lengthy shift could be classified as LWH, shift work is characterized by scheduled periods of work that occur earlier or later in the day than the standard work hours of a full-time employee, which delineates shift work from LWH. [1]

Most industrialized countries limit the number of work hours by establishing a maximum number of total hours to be worked during a specified period of time (e.g., 40 hours per week).[6] Although national regulations regarding work time maximums vary across industrialized nations − as does the enforcement of such standards and the consequences for employers that surpass the maximums − only the U.S., Japan, and New Zealand permit workweeks of more than 60 hours.[7] Further, among developed nations, the U.S. is alone in having no nationally-imposed limit on the number of hours an individual over the age of 16 is allowed to work per week in general industry.[8] Even so, the 40-hour workweek is generally considered the standard for full-time employment in the U.S. In a nation-wide survey of U.S. employers, approximately three-quarters of respondents characterized their organizational norm for a full-time workweek as 40 hours, with an overall study range of 32 to 48 hours per week, not including overtime.[9]

With the exception of transportation workers, all existing work hour restrictions for specific industries and occupations have been established by a combination of professional association guidelines and state laws rather than federal mandate.[10] Federal measures to limit extreme work schedules have been proposed in the interest of public safety for a few occupations such as health care professionals due to the possibility of fatigue-induced errors, injuries, or accidents. However, no federal legislation has been passed to limit the maximum number of hours worked per day or per week for workers outside the transportation industry.

Currently, significant variation in allowable work hour durations exists within and across professions, despite the fact that the physiological responses to fatigue resulting from long work hours are equivalent.[11]

Variability among the legislated maximum workweek, the generally accepted standard workweek, and the actual hours worked exists not only in measures of work hours in the U.S., but also in measures of time at work around the world. In the European Union (E.U.), for example, regulations limit work time to a maximum of 48 hours per week, including overtime.[12,13] Even so, the statutory maximum workday and workweek vary across the E.U. as do the self-reported hours worked by full-time employees, as shown in Table 1.[14] In Asia, Korean workers are limited to 40 work hours per week with an additional 12 hours of overtime per week according to the Labor Standards Act of 1997.[15] However, despite such regulations, the OECD reported that Koreans worked longer hours than laborers in any other surveyed country in 2011.[16] The evidence of illegal practices and exemptions for certain

Table 1. Work hours in a sample of European Union countries, 2010

Country	Statutory Maximum Workday (hours)	Statutory Maximum Workweek (hours)	Reported Average Workweek, 2010 (hours)
Belgium	8	38	38.6
Finland	8	40	37.8
France	10	48	38.0
Germany	8*	48	40.5
Greece	12	48	39.7
Spain	9^	40	38.1
United Kingdom	13	48	40.5

* Daily work hours may be up to 10 hours as long as an 8-hour per day average is maintained over a six-month period
^ Daily work hours may be up to 12 hours as long as a 9-hour per day average is maintained over a collectively agreed upon period of time

Source: European Foundation, 2011 [14]

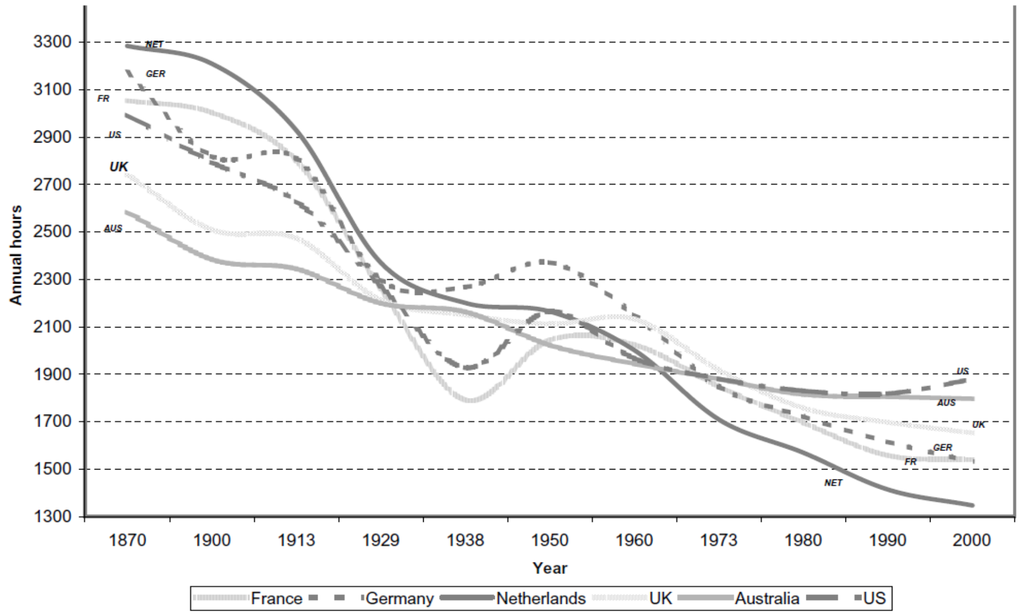
industrial sectors and occupational categories that undermine statutory work hours and overtime limits also varies by country.[17]

Time Trends in Hours Worked

U.S. laborers began calling for reduced work hours as early as 1791, when Philadelphia carpenters unsuccessfully went on strike to demand a 10-hour workday limit. Although a small proportion of workers lobbied successfully for work hour reductions in the 19th Century, 16-hour workdays and 6-day workweeks were common during that time in the U.S.[18] The campaign for 8-hour work days, which picked up momentum in the 1870s, was one of the first organized efforts to alter the conditions of work in the U.S.[18,19] Although Henry Ford is widely attributed with introducing the 8-hour workday as an industrial standard in 1914,[20] it was not a widespread practice in the U.S. until the passage of the Fair Labor Standards Act in 1937, which applied to approximately 20 percent of the workforce.[21] Since that time, U.S. work hours declined gradually across most occupations (with the exception of a spike during World War II and the years immediately thereafter) until the last two decades of the 20th Century,[22] at which point work hours began a slow increase, both in terms of average annual work hours (Figure 1)[7,23] and average number of work hours per week (Figure 2).[24-27]

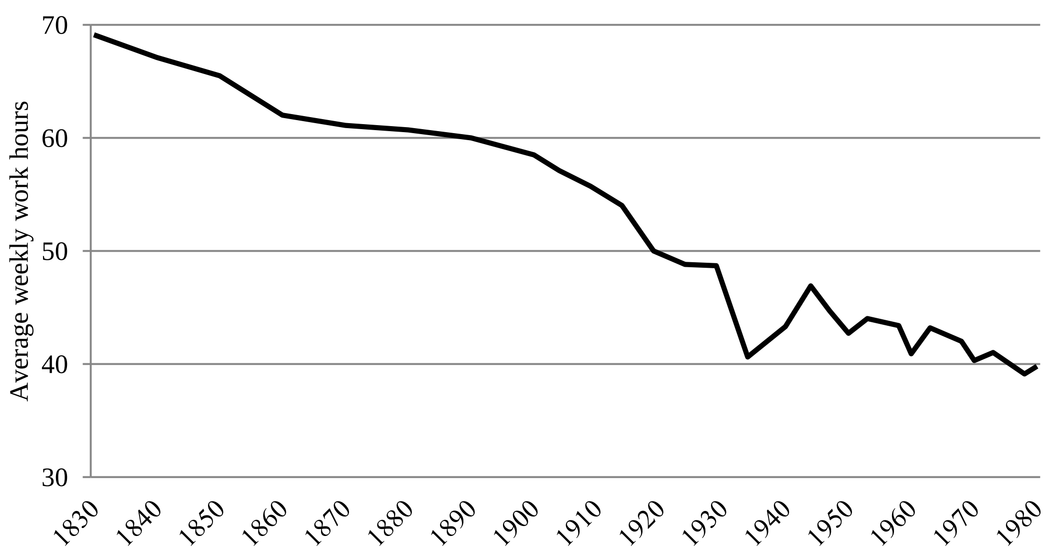
The increase in U.S. work hours since the end of the 1970s has been seen in terms of average annual hours worked, the number of weeks worked annually, and average weekly work hours, with the largest increases seen among female workers (Table 2).[23] Since 2000,

Figure 1. Historical trends in total annual work hours per worker in selected countries, 1870-2000



Source: Huberman, 2002[28] in Lee *et al.*., 2007[7]

Figure 2. Historical trends in average weekly work hours per U.S. worker, 1830-1980



Adapted from: U.S. Department of Interior, 1883[24]; U.S. Senate, 1893[25]; Owen, 1988[26;] Whaples, 1990[27]

U.S. work hours have stabilized as some of the longest among developed countries. In terms of annual work hours, the U.S. ranked second in average annual work hours in 2011 among wealthy industrialized member nations according to data from the Organisation for Economic Co-operation and Development (OECD); the U.S. ranked first among the same population in 2012 (although it should be noted that the highest ranking country in 2011 did not report data in 2012).[29] OECD data from 2012 also indicated that U.S. workers averaged more work hours than those in every western European country except Greece, and performed the equivalent of up to three additional months of work each year compared to workers in, Belgium, Denmark, France, Germany, Ireland, Luxembourg, the Netherlands, Norway, and the United Kingdom.[29] In terms of the average weekly work hours, U.S. full-time workers' hours have held relatively steady since 2000, ranging from 44.9 to 46.9 hours per week (Figure 3). Greater variation in average weekly work hours has been evident in part-time workers during the same period, which is generally attributed to the effects of the economic downturn that began after the turn of the 21st century.[30]

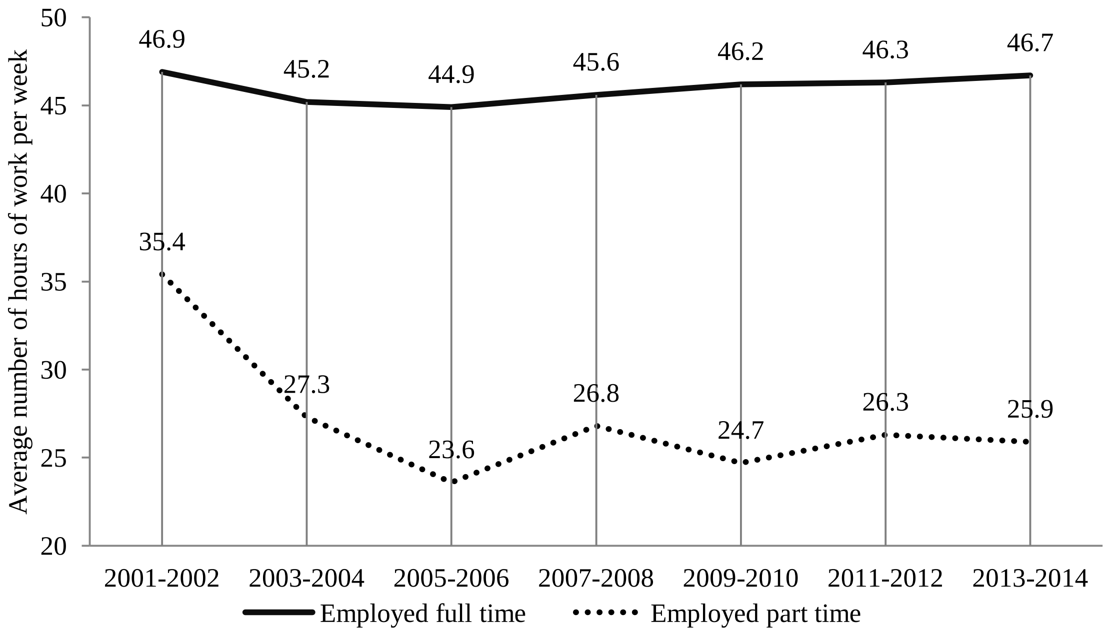
Not all U.S. occupations have experienced a lengthening of hours at work over the last thirty-five years. Instead, work hour expectations are increasingly bifurcated by socio-

Table 2: Increases in U.S. work hours, 1979-2007

Workers	Annual Hours Worked		Change, 1979-2007		
	1979	2007	Annual work hours (%)	Weeks worked (%)	Weekly work hours (%)
All	1,687	1,868	10.7	8.1	2.4
Males	1,915	2,000	4.4	4.0	0.4
Females	1,438	1,729	20.3	13.6	5.8

Adapted from: Mishel, 2013[23]

Figure 3: Typical number of weekly work hours reported by U.S. workers, 2001-2014



Source: Saad, 2014[30]

economic class, with LWH associated with managerial and professional workers while shorter work hours are more likely among low-wage service, manual, and production employees.[11,31] Low-wage employees tend to have less job control, including less flexibility in terms of work hours and fewer options for working additional hours, if desired.[31,32] Thus, those employees reporting increasingly longer hours tend to be male professionals with relatively high levels of job control and job satisfaction.[31]

Nationally-representative data from the 2011 wave of the Panel Study of Income Dynamics (PSID) on work schedule practices, 32.8 percent of U.S. workers over the age of 18 reported working more than 40 hours per week (Table 3). Compared to those working 40 hours per week or less, those working more than 40 hours per week were disproportionately male and older, with higher levels of educational attainment and salaried positions. Across industries, 44% (n=1,062) of non-service workers report working more than 40 hours per week, whereas only 29.4% (n=2,273) of service workers report doing so. In terms of occupational types, 40.7% (n=896) of those in manual occupations report working more than 40 hours per week compared to only 30.6% (n=2,437) of non-manual workers.

Work hours practices are influenced by economic, cultural, legal, and institutional factors.[33] Many incentives exist for LWH, which appeal to both workers and their employers, including higher pay, the prospect of professional advancement, and a demonstration of usefulness to their employer. For employers, the need to meet production demands and to minimize employee benefit and recruitment costs are reasons that organizations allow − and often encourage − their employees to work longer hours instead of hiring additional staff.[10,33]

Long Work Hours and Adverse Health Outcomes

Evidence that exposure to LWH increases the risk of deleterious health conditions has been documented in the scientific literature for more than 50 years.[34] A relationship between LWH and health has been demonstrated for outcomes including poor general health,[35] cardiovascular disease,[36] musculoskeletal disorders,[37] work-related injuries,[38,39] depression, [40] and sleep disruption.[41] Even so, the data on each health outcome remain relatively thin. Table 4 presents a range of health related outcomes related to LWH with their accompanying number of tests of association and a summarized indicator of the evidence of association.[1,42] These outcomes were analyzed in 53 studies published from 1996 to 2013, which were selected for inclusion based on the following criteria: (1) evidence was provided in

Table 3. Distribution of sociodemographic and occupational characteristics by weekly work hours in adult, working participants in the Panel Study of Income Dynamics, 2011 wave (N=10,194)

	Weekly Work Hours		
	≤ 40 hours/week n (%)	>40 hours/week n (%)	*p*-value
Total	6854 (67.2)	3340 (32.8)	
Sex			*p*<0.001
Male	2712 (39.6)	2176 (65.2)	
Female	4142 (60.4)	1164 (34.9)	
Age			*p*<0.001
18-24	653 (9.5)	140 (4.2)	
25-34	2036 (29.7)	985 (29.5)	
35-44	1375 (20.1)	832 (24.9)	
45-54	1398 (20.4)	785 (23.5)	
>54	1392 (20.3)	598 (17.9)	
Educational level (years of education)			*p*<0.001
< 11 years	431 (6.3)	160 (4.8)	
11 - 12 years	2433 (35.5)	898 (26.9)	
>12 years	3990 (58.2)	2282 (68.3)	
Employment status			*p*=0.341
Self-employed	6014 (87.9)	2912 (87.3)	
Employed by others	826 (12.1)	425 (12.7)	
Number of workers at job location			*p*<0.001
≤ 10 workers	2565 (39.7)	1037 (32.5)	
11 - 49 workers	1252 (19.4)	649 (20.4)	
50 - 249 workers	1449 (22.4)	731 (22.9)	
≥ 250 workers	1202 (18.6)	772 (24.2)	
Industry			*p*<0.001
Services	5471 (80.2)	2273 (68.2)	
Non-services	1349 (19.8)	1062 (31.8)	
Occupational type			*p*<0.001
Non-manual	5527 (80.9)	2437 (73.1)	
Manual	1306 (19.1)	896 (26.9)	
Pay status: salaried vs. hourly			*p*<0.001
Salaried	1594 (30.4)	1600 (58.9)	
Hourly	3646 (69.6)	1118 (41.1)	

original, peer-reviewed articles; (2) studies were cohort, case-control, or cross-sectional in nature; (3) studies either excluded shift work from the exposure of LWH or analyzed it separately from LWH; (4) included a reference group that worked approximately 40 hours per week; (5) analyzed work hours and/ or overtime hours at the individual level (as opposed to the group, division, or organizational level); (6) examined health outcomes that could lead to morbidity and/ or mortality; and, (7) provided a quantitative estimate of the relationship between LWH and the outcome measures.

Of the 25 health-related outcomes analyzed, only four were evaluated in 10 or more tests of association: sleep disruption, cardiovascular disease, psychological ill health, and heart rate. The majority of health-related outcomes in this sample (14/25, or 56%) were analyzed in five or fewer studies, which prevents meaningful conclusions about the true nature of the relationship between LWH and those outcomes due to the paucity of evidence. Generally speaking, the more limited the evidence on the relationship between LWH and a specific health outcome, the more likely that results of such investigations are inconsistent or contradictory (Table 4). The absence of consistent findings on the LWH−health relationship is typically attributed to several common methodological weaknesses. The majority of studies on LWH have been undertaken on data collected on specific occupational groups (e.g. nurses, office workers) or workplace settings (e.g. factories, hospitals), often with small sample sizes, both of which limits the generalizability of study findings to other working populations.[44-47] The analysis of LWH and health through the use of representative national databases is rare, and existing studies have typically been conducted on data from countries other than the U.S., with the majority of research occurring on populations in Japan,[48]

Table 4. Evidence of the strength of association between long work hours and several health-related outcomes in 53 selected studies, 1996-2013

	Tests of association (N)	Summarized evidence for association
Mortality		
All-cause	2	±
Morbidity		
Cardiovascular disease	4	+
Hypertension	6	±
Diabetes	4	±
Metabolic syndrome	1	+
Psychiatric morbidity	6	++
Physiology		
Heart rate	10	±
Blood pressure	5	+
Immunity	2	−−
Cholesterol	4	−
Fasting blood sugar	1	+
Noradrenaline	1	−
Cognitive function	6	−−
Work disability		
Disability retirement	1	+
Sickness absence	1	−
Subjective health		
General ill health	6	+
Psychological ill health	10	++
Physical ill health	2	++
Fatigue	6	++
Behavior		
Sleep disruption	27	++
Increased alcohol consumption	8	±
Smoking	6	+
Drug use	1	0
Unhealthy eating habits	2	+
Decreased physical activity	4	0
Body mass index/ obesity	7	±

Sources: Artazcoz *et al.*, 2013[43]; Bannai and Tamakoshi, 2014[1]; van der Hulst, 2003[42]; Virtanen *et al.*, 2012[36]

Key:

++ = positive association considered particularly strong due to number of studies and/ or methodological rigor of studies

+ = positive association found in more than half of studies

± = mixed findings due to approximately equal positive and negative evidence

− = negative association found in more than half of studies

−− = negative association considered particularly strong due to number of studies and/ or methodological rigor of studies

0 = no significant association

the Scandinavian countries,[49] and the United Kingdom.[36] Most studies of the association of LWH with health statuses and behaviors have analyzed samples composed of salaried or self-employed workers, but the working conditions, compensation schemes, and motivation behind working long hours may differ significantly between these two groups as well among hourly-wage laborers.[50] Some studies that have not demonstrated an association between LWH and health have treated work hours as a continuous variable and assumed a linear relationship,[47,51,52] which may underestimate the LWH−health relationship by failing to identify individuals who have chosen reduced work hours due to poor health.[53] Several researchers have pointed to the need for further study of potentially confounding or interacting factors, such as gender, work characteristics, family responsibilities, or domestic work.[12,54]

Several systematic reviews have examined the health effects of LWH[1,4,10,33,34,42] and have demonstrated the lack of a consistent framework for measuring work hours as an occupational risk factor. For example, some studies have defined LWH in terms of overtime[55,56] or extended shift hours[44] while others examined daily[57] or weekly[58] work hours. Subdivisions of exposure categories in terms of hours worked – representing lesser or greater levels of exposure – have varied similarly.[1,36,42] Differentiating LWH from overtime work and shift work has been inconsistent, though the three concepts are recognized as distinct.[1,59,60] These limitations have rendered much of the research on LWH not comparable and, ultimately, inconclusive. Inadequate exposure assessment is consistently identified as one of the main sources of variability in the literature.[1,33]

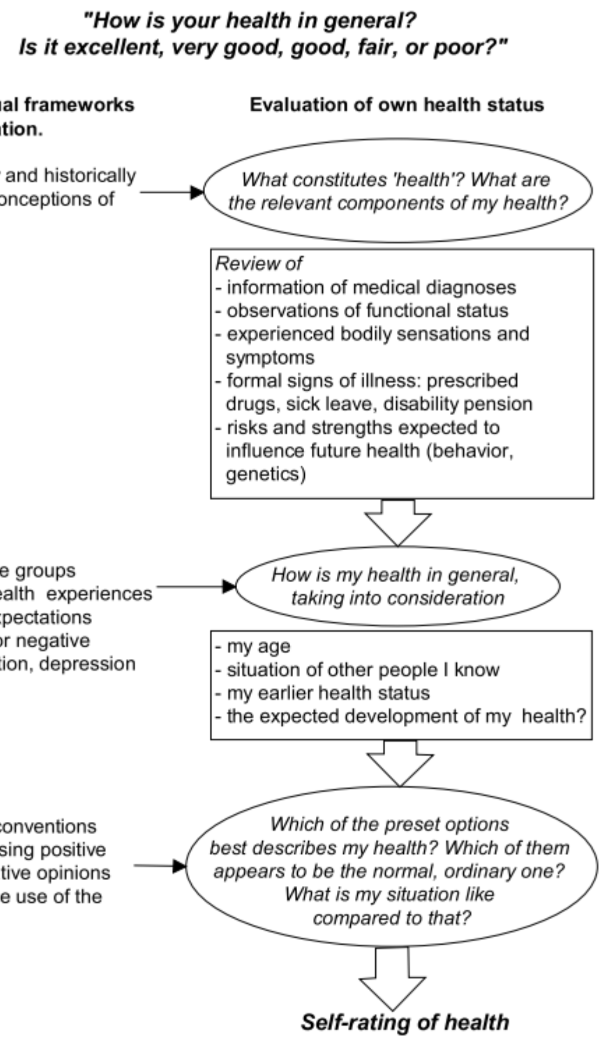
Long Work Hours and Self-Reported General Health

Measures of self-reported general health (SRGH) provide a single, composite health score reflecting an individual's perception of his or her collective well-being and morbidity status.[61] SRGH provides an individual's assessment of his or her health, which is based on a combination of subjective and objective biological, psychological, and social aspects of health.[62] A conceptual framework for the process of individual health evaluation is presented in Figure 4.

SRGH has been shown to be a better predictor of mortality than medical diagnostic criteria[62,63] and are used widely as an index of the overall health of an individual or, when aggregated, of a population.[61] As a summary measure of population health, SRGH allows for comparisons of the health of one population with another or for monitoring changes in the health of a single population over time.[64] Because working long hours may have a greater impact on some aspects of health than others, self-assessed general health is an appropriate measure for evaluating a broad range of physiological outcomes that may be associated with the exposure through the use of a composite measure.

Whereas some studies have demonstrated a positive association between LWH and poor SRGH,[35,43,47,57,65] others have not[53, 66] (Table 5). Studies reporting a positive association between LWH and poor SRGH found a significant LWH−health relationship among individuals working more than 45 hours per week,[35] individuals working more than 60 hours per week who had between one and 10 years of job tenure,[65] junior physicians averaging 10.2 or more work hours per day,[57] soldiers averaging 76 or more work hours per week,[47] men working more than 40 hours per week who were culturally expected to

Figure 4. The individual health evaluation process to determine SRGH status



Source: Jylha, 2009 [62]

provide the majority of the household income,[43] women who worked more than 60 hours per week,[65] women working more than 40 hours per week who were culturally expected to provide a substantial amount of the household income,[43] and men and women working more than 50 hours per week who were culturally expected to provide a relatively equal amount of the household income.[43] These studies suggested that factors associated with LWH differ by sex. More consistent associations between work hours and health outcomes were shown in populations of men than of women, and one study detected a dose-response relationship between increasing work hours and health outcomes in men.[53] Among studies reporting an association between LWH and poor SRGH, researchers found that individuals exposed to LWH were 1.2 to 6.4 times more likely report poor SRGH, compared to individuals in control groups that worked fewer hours.

No significant work hours−health association was reported in two studies for individuals working more than 40 hours per week,[53,66] and one study reported no LWH−health association for men or women who were not culturally expected to be a significant contributor of household income.[43]

Table 5. Studies of LWH and SRGH

First Author, Year	Study Design and Sample	Work Hours Measure	Justification for Work Hour Risk Categories	General Health Measures	Other Outcome Measures	Statistical Analyses	Covariates	Results*
Artazcoz, 200766	Cross-sectional Data source: 2002 Catalonian Health Survey N = 2792 Spain males: 59% A_r: 16-64 years Occup: diverse Organ: diverse	SRWH: WH/w WH categories, h/w: <30, 30-40¶, >40	no explanation provided for specific categories	self-assessed general health status	mental health status self-reported hypertension job satisfaction health-related behaviors	MLogRA interaction term between hours of paid work and of domestic work included in model results stratified by sex	demographic factors socio-economic status type of work contract weekly hours of domestic work marital status children 0-3 years of age in home	Adjusted ORs (95% CIs): Males, WH/w: - <30: 1.2 (0.5, 3.0) - 30-40: ref - >40: 0.9 (0.6, 1.2) Females, WH/w: - <30: 1.2 (0.7, 2.0) - 30-40: ref - >40: 1.3 (0.8, 1.9)
Artazcoz, 200953	Cross-sectional Data source: 2006 Catalonian Health Survey N = 7103 Spain males: 56% A_r: 16-64 years Occup: diverse Organ: diverse	SRWH: WH/w WH categories, h/w: <30, 30-40¶, 41-50, 51-60	no explanation provided for specific categories	self-assessed general health status	mental health status self-reported hypertension job satisfaction health-related behaviors	MLogRA results stratified by sex	employment conditions family and domestic characteristics	Adjusted ORs (95% CIs): Males, WH/w: - <30: 1.3 (0.7, 2.4) - 30-40: ref - 41-50: 1.0 (0.8, 1.2) - 51-60: 1.4 (1.0, 2.1) Females, WH/w: - <30: 1.3 (1.0, 1.7) - 30-40: ref

								- 41-50: 1.0 (0.8, 1.4) - 51-60: 1.4 (0.7, 2.6)
Artazcoz, 201343	Cross-sectional Data source: 4th EWCS N = 15583 25 European countries males: 60% A_m: 39.1 years A_r: 16-64 years Occup: diverse Organ: diverse	SRWH: WH/w main job, WH/w all other jobs WH categories, h/w: 30-40¶, 41-50, 51-60	no explanation provided for specific categories desire to evaluate “moderately long working hours,” defined as 40-60 h/w desire to avoid “extremely long working hours,” defined as >60 h/w	self-assessed impact of job on health	work-related stress work-related psychological distress	MLogRA results stratified by geographic location, sex	socio-demographic factors type of work contract family responsibilities identity of main contributor of household income geographical location	Adjusted ORs (95% CIs): Males, WH/w: Anglo-Saxon: **- 41-50: 4.1 (2.9, 5.8)** **- 51-60: 6.4 (4.2, 9.9)** Continental: **- 41-50: 2.2 (1.8, 2.7)** - 51-60: 1.5 (0.9, 2.7) Eastern European: - 41-50: 1.1 (0.8, 1.4) **- 51-60: 2.4 (1.5, 3.6)** Nordic: - 41-50: 1.2 (0.8, 1.7) - 51-60: 2.4 (1.0, 5.7) Southern European: **- 41-50: 1.7 (1.4, 2.0)** **- 51-60: 2.6 (1.7, 4.0)**

								Females, WH/w: Anglo-Saxon: - 41-50: 1.6 (1.0, 2.6) - 51-60: 0.8 (0.4, 1.9) Continental: - 41-50: 0.8 (0.5, 1.2) - 51-60: 1.1 (0.6, 2.2) Eastern European: - 41-50: 1.0 (0.7, 1.3) - **51-60: 3.4 (1.6, 7.2)** Nordic: **- 41-50: 2.2 (1.2, 3.9)** - 51-60: 2.4 (0.7, 8.5) Southern European: - 41-50: 1.7 (1.2, 2.3) - 51-60: 0.5 (0.2, 1.0)
Baldwin, 199767	Cross-sectional Data source: Scottish university medical student cohort	SRWH: WH/w mean: 76 range: 33-128	no risk categories identified (WH treated as continuous variable)	Total GHQ, Somatic GHQ, self-assessed general health status	mental health status job performance social dysfunction physical ailments in last year	Factor analysis, Corr, MLinRA significant covariate associations	not specified	Correlation coefficients (*p*-value): HW & Total GHQ: ns WH & Somatic

	N = 142 U.K. males: 55% A_m: 25 years Occup: physicians Organ: hospitals				alcohol consumption	controlled for each other		GHQ: ***r* = 0.3 (*p*=0.002)** WH & self-assessed health: ns
Ettner, 200135	Cross-sectional Data source: 1995 MIDUS N = 2048 U.S. males: 50% A_r: 25-74 years Occup: diverse Organ: diverse	SRWH: WH/w WH categories, h/w: <35 (“part time”), 35-45¶ (“full time”), >45 (“more than full time”)	WH categories defined to be consistent with Ross & Mirowsky, 199568	self-assessed impact of job on health	none	MLogRA results stratified by job characteristic category	occupation-based and self-reported job characteristics socio-demographic factors personality characteristics work scheduling wage rate	Adjusted ORs (95% CI): Occupation-based job characteristics, WH/w: **- <35: 1.6 (1.2, 2.0)** - 35-45: ref **- >45: 0.7 (0.6, 0.8)** Self-reported job characteristics, WH/w: - <35: 1.3 (1.0, 1.7) - 35-45: ref **- >45: 0.8 (0.6, 0.9)**
Grosch, 200631	Cross-sectional Data source: 2002 GSS N = 1744	SRWH: WH/w WH categories, h/w: 1-34 (“part time”), 35-40¶ (“full time”), 41-	no explanation provided for specific categories	QWL	mental health status physical ailments in last year occupational	MLogRA	socio-demographic factors	Adjusted ORs (95% CI): Adjusted model, WH/w: - 1-34: 1.0 (0.7,

	U.S. males: 48.7% A_m: 41 years Occup: diverse Organ: diverse	48 (“low overtime”), 49-69 (“medium overtime”), ≥70 (“higher overtime”)			injuries in last year work-related stress job satisfaction job-family interference			1.4) - 41-48: 0.9 (0.6, 1.4) - 49-69: 0.5 (0.3, 0.9) **- ≥70: 1.9 (1.1, 3.5)**
Jex, 199957	Cross-sectional N = 2273 U.S. males: 96% A_m: 25 years Occup: soldiers Organ: U.S. Army	SRWH, WH/d mean: 10.23 SD: 2.85	no risk categories identified (WH treated as continuous variable)	composite physical strain index created for this study (list of stress-related health symptoms, frequency of those symptoms in past month)	psychological strain job satisfaction organizational commitment	MRegA interactions tested between WH and self-efficacy, WH and collective efficacy	work overload task significance self-efficacy collective efficacy	Parameter estimate, main effect model (SE; *p*-value): **WH/w: 0.16 (0.07; *p*=0.019)**
Song, 201465	Cross-sectional Data source: 2008 KLIPS N = 3699 Korea males:65% A_r: 25-64 Occup: diverse Organ: diverse	SRWH, WH/w WH categories, h/w: ≤40¶, 41-52, 53-60, >60	WH categories defined to be consistent with Korea’s Labor Standards Act and Won *et al.* 200869	self-assessed general health status	none	MLogRA results stratified by sex, job tenure - both factors also retained in multivariate models	socio-demographic factors occupational characteristics health-related behaviors	Adjusted ORs (95% CI): Adjusted model, WH/w: - ≤40: ref - 41-52: 1.1 (0.9, 1.3) - 53-60: 1.1 (0.9, 1.3) - >60: 1.4 (1.1, 1.8) Males, WH/w: - ≤40: ref - 41-52: 1.0 (0.8, 1.3) - 53-60: 0.9 (1.8, 1.3)

- >60: 1.1 (0.8, 1.6)

Females, WH/w:
- ≤40: ref
- 41-52: 1.2 (0.9, 1.6)
- 53-60: 1.2 (0.9, 1.7)
- **>60: 2.0 (1.3, 3.0)**

Tenure ≤1 yr, WH/w:
- ≤40: ref
- 41-52: 0.9 (0.6, 1.2)
- 53-60: 0.9 (0.6, 1.4)
- >60: 1.0 (0.6, 1.6)

Tenure, >1-≤10 yrs, WH/w
- ≤40: ref
- 41-52: 1.1 (0.9, 1.5)
- 53-60: 1.2 (0.9, 1.6)
- **>60: 1.8 (1.3, 2.5)**

Tenure, >10 yrs, WH/w:
- ≤40: ref
- 41-52: 1.3 (0.8, 1.9)

								- 53-60: 1.1 (0.6, 1.8) - >60: 1.4 (0.7, 2.7)
Taris, 2011[70]	Prospective Data source: Study on Health at Work N = 649 The Netherlands males: 84% A_m: 41 years Occup: diverse Organ: diverse	SRWH: OT/w mean: 5.1 range: 0-11.4	no risk categories identified (WH treated as continuous variable)	composite subjective health index created for this study (self-assessed health, self-assessed physical health)	health behaviors BMI	SEM	job demands job control socio-demographic factors	Standardized maximum likelihood estimates (p-value): **WH & subjective health: -0.06 (p<0.05)**

Key:

* = bolded values are statistically significant at α=0.05; ¶ = reference category; # = number; /d = per day; /m= per month; Am = mean age; Ar = age range; Corr = correlation; EWCS = European Working Conditions Survey; GHQ = General Health Questionnaire; GSS = General Social Survey; h = hours; KLIPS: Korean Labor and Income Panel Study; MIDUS = Mid-Life in the United States study; MLinRA = multivariate linear regression analysis; MLogRA = multivariate logistic regression analysis; MRegA= multi-level regression analysis; Occup = type of occupation; Organ = type of organization; QWL = Quality of Work Life module; ref = reference group; SE = standard error; SRWH = self-reported work hours; WH = work hours; yr = year; yrs = years

Artazcoz and colleagues (2007) evaluated the effect of gender on associations between LWH, health, and health-related behaviors in a cross-sectional study of 2,792 salaried contract workers aged 16 to 64 years in Catalonia, Spain, participating in the 2002 Catalonian Health Survey. Work hours for the previous week and perceived health status were self-reported during in-person interviews. SRGH responses were transformed into a dichotomized variable representing good vs. poor general health. No statistically significant associations between number of work hours and SRGH were seen for men or women in any of the hours of paid work per week variables. An interaction term for LWH and hours of domestic work was included in the models for health-related outcomes, which did not produce a significant result. Consistent differences were seen by gender for the other health-related outcome variables associated with LWH. Statistically significant associations were found for six of the seven health-related outcomes among women (poor mental health status, hypertension, job dissatisfaction, smoking, no leisure time physical activity, daily number of hours slept ≤6), but only one of the seven outcomes among men (daily number of hours slept ≤6).[66]

Artazcoz and colleagues (2009) cross-sectionally analyzed the relationship of LWH with health status (including SRGH) and health-related behaviors in the 7,103 salaried contract workers participating in the 2006 Catalonian (Spain) Health Survey. Participants were surveyed at home in face-to-face interviews about their work hours and health status. Work hours data were based on self-reported hours worked during the previous week. SRGH was assessed by asking participants to characterize their health according to standard categories, which were later transformed into a binary variable representing good vs. poor health. Adjusted odds ratios were stratified by sex and reported for each of the categories of work hours, but no statistically significant relationship were found for poor SRGH among any of the work hours categories. However, consistent differences were seen between the sexes in the relationship of LWH with the other health outcomes that were measured. LWH was significantly associated with three of seven outcomes among women (job dissatisfaction, smoking, daily number of hours slept ≤6) whereas it was associated with six of seven outcomes among men (poor mental health status, self-reported hypertension, job dissatisfaction, smoking, no leisure-time physical activity, daily number of hours slept ≤6), and a gradient of effect was seen among men.[53]

Artazcoz and colleagues (2013) evaluated the association between LWH and three health outcomes (including work-related general health) in a cross-sectional study of 15,583 workers participating in the 2005 European Working Conditions Survey. Work schedules and health statuses were determined during face-to-face interviews in the participants' homes. The impact of work on general health was determined with a binary-response question ("*Does your work affect your health, or not?*") followed with a checklist of 16 health issues for those who responded affirmatively. Those who reported working more than 60 hours per week were excluded from the analysis. Results were stratified by geographic location and sex. LWH was associated with health problems among workers in all five geographic categories, although not consistently across all exposure categories. Only Nordic males and Continental females did not have a statistically significant association between any of the work hour categories and work-related poor health. A gradient of effect was suggested by the analyses, with a positive trend demonstrated in most groups between work hour durations and poor health status, although this was not uniformly the case. The number of significant associations was higher among men than women (18 out of 30 vs. 10 out of 30, respectively), which the authors attributed to gender-specific social norms.[43]

Baldwin and Wrate (1997) examined the associations of work characteristics with attitudes, health outcomes, and job performance among 142 junior doctors in the U.K. using a cross-sectional study design. Participants on the U.K. mainland were interviewed in-person at baseline during the winter of 1993/1994; those elsewhere received a mailed questionnaire. The entire cohort was surveyed by mail one year later, which was the basis for the majority of the data used in this analysis. Hours of work were measured by asking respondents to provide the number of hours worked in the last week. Physical health was assessed using the somatic symptoms subscale of the 28-item General Health Questionnaire. Participants who were interviewed were also asked to identify how often they had experienced any of 25 common health problems in the last year, and these baseline data were also part of the analysis, despite the fact that they did not necessarily capture the health of the participants at follow-up. Hours worked was not significantly correlated to the total score on the GHQ or to the number of health problems in the last year, but it was significantly related to scores on the somatic symptom subscale indicating general poor health (r=0.26, p=0.002).[47]

Ettner and Gryzwacz (2001) investigated the association between work characteristics and physical and mental health in 2,048 participants in a cross-sectional analysis of the 1995 Mid-Life in the United States (MIDUS) study. Job- and health-related information was collected during a telephone interview and two subsequent mailed questionnaires; this analysis included only those participants who completed the interview and submitted both questionnaires. Occupation, industry, and job characteristic data were self-reported with additional details provided by appending information from the *Dictionary of Occupational Titles* to the occupational codes assigned to each participant. The impact of work on health was evaluated with a categorical response question ("*Overall, what kind of effect does your job have on your physical health?...on your mental health?*"). Results were stratified by self-reported job characteristics and by occupation-based job characteristics provided by the appended information. In the analysis of occupation-based measures, compared to those individuals working full time (35-45 hours per week), workers reporting working more than full time (>45 hours per week) were 31 percent less likely to report a positive impact of work on health (OR: 0.69; 95% CI: 0.6, 0.8) whereas part time workers (<35 hours per week) were 57 percent more likely to report a positive impact (OR: 1.6; 95% CI: 1.2, 2.0). A similar trend was seen in the analysis of self-reported job characteristics data, although the result for part time workers was not statistically significant.[35]

Grosch and colleagues (2006) conducted a cross-sectional study of 1,744 General Social Survey respondents from the 2002 wave. Work hour durations for the previous week and general health status were self-reported during face-to-face interviews using the 76-item Quality of Work Life (QWL) module developed by the National Institute for Occupational Safety and Health (NIOSH). Five work hour categories were delineated for analytical purposes, which captured work hours from 1 hour per week to more than 70 hours per week. Participants working 70 hours or more per week were almost twice as likely (OR = 1.9; 95% CI: 1.1, 3.5) to report low levels of general health when compared to participants working 35 to 40 hours per week. Though not statistically significant, the results for individuals working between 41 and 69 hours per week suggested that they experienced better general health than those working 35 to 40 hours per week, which may indicate that moderate amounts of overtime attract workers who feel well enough to work longer hours.[31]

Jex and Bliese (1999) explored the potential moderating effect of efficacy beliefs on work-related stressors among a primarily young, male population of 2,273 U.S. Army soldiers in 36 companies in a cross-sectional study. Work hours measures were self-reported by the soldiers and were dependent upon company membership. General physical health status was determined using a composite physical strain index, derived from a 24-item stress-related health problems scale and the frequency of those health problems in the last month, both of which were self-reported. Work hours was significantly associated with the physical health score (β=0.157; SE=0.067; p=0.019). Interaction terms for work hours and efficacy and work hours and collective efficacy were introduced into the model but were not statistically significant.[57]

Song and colleagues (2014) conducted a cross-sectional study of 3,699 participants of the Korean Labor and Income Panel study (2008 wave) who were employed full-time. Measures of work hours and health status were self-reported by participants. Weekly work hours were not significantly associated with SRGH in the adjusted multivariate model. However, when stratified by sex, women working more than 60 hours per week were twice as likely (OR: 2.0; 95% CI: 1.3, 3.0) to report poor health than those working 40 or fewer hours. The authors attributed the fact that women reported lower levels of SRGH than men to the disproportionate amount of time women spend, on average, on household duties as well as gender disparities of the labor market, which may contribute to women's job discontinuity and necessitate longer work hours to close gaps in experience. In a model stratified by job tenure, individuals with between one and 10 years on the job who worked more than 60 hours per week were 1.8 (95% CI: 1.3, 2.5) times more likely to report poor health when compared to those working 40 hours or less. The authors suggested that this demonstrated the possibility of a dose-response pattern over years of working extended hours. However, they did not indicate if they had further subdivided and analyzed the job tenure variable to support that hypothesis.[65]

Taris and colleagues (2011) examined the influence of the behavioral lifestyle mechanism on the work hours–health relationship in a prospective study of 649 Dutch full-time workers from the two-year Study on Health at Work. Information on work hours and health statuses was collected via an online survey at baseline and at follow up. The responses to two measures of self-assessed health ("*How would you evaluate your health? How would you evaluate your physical health in general?*") were averaged to form a composite general health indicator. The difference between actual work hours and contracted work hours was calculated, and the resulting overtime value was treated as a continuous variable. Structural equation modeling yielded a statistically significant correlation coefficient for overtime hours and SRGH (standardized effect=-0.06; $p<0.05$). This result should be interpreted with caution, however, as the authors indicated that individuals with lower SRGH scores at baseline were more likely to drop out of the study than those with higher SRGH scores.[70]

The disparate results of the studies under discussion have been attributed to two factors: (1) the true health effects of work, and (2) individual attributes resulting in differential reporting behaviors. The nature of the true health effects of work depends on the particular characteristics of the respondent's job, in conjunction with individual factors that moderate the effects of the working conditions. Thus, the true health effects of work are believed to vary both within and across occupations and industries, which may result in contradictory findings depending on the study sample composition.[71] The individual traits that result in differential reporting behaviors depend on the respondent's tendency to emphasize primarily positive or negative health states, understanding of the measure being surveyed, and likelihood of revealing undesirable health conditions.[35] It has been suggested that self-reports of health may be influenced by personality, self-perception, or cultural heritage in a way that clinical measures are not.[71] Thus, the factors influencing reporting behavior tend to vary at the individual-level, which reduces the possibility that they will bias the reporting of job effects on health if they are evenly distributed throughout the study population.[35]

Potential Confounding and Moderating Factors

Numerous individual and work characteristics have been identified as potential confounders and controlled for in multivariate models, including demographic factors, socio-economic status, employment conditions, work contract types, and levels of self-efficacy.[43,57,66] However, the same worker, job, and organizational characteristics have been analyzed as moderating factors in other multivariate analyses.[33,35,43,53,66] Researchers have not commented on whether these variables are potentially both confounders and modifiers,[72] although one recent study adjusted for and stratified by sex and job tenure, which suggests that they believed the two variables were acting as both.[65]

The potential for self-selection bias resulting from a portion of the exposed choosing to engage in long work hours has been acknowledged, but less attention has been focused on the influence of choosing or being required to work LWH for lengthy periods of time, such as over the course of several years.[4,12] It has been suggested that those who work long hours for an extended period of time should be analyzed as a survivor population.[12] A form of the healthy worker effect, survivorship bias is the consequence of the selection of relatively healthy individuals into the workplace and the subsequent early exiting of workers who develop poor health. The resulting population has superior health compared to that of the initial population, which may confound measures of association[73,74] and complicate the identification of an appropriate comparison population.[12] In this situation, work status is related to future exposure status, and so a reciprocal relationship exists between exposure and work status.[75] Special analytical techniques have been proposed to address this specific type of survivorship bias, known as time-varying confounding affected by prior exposure.[74]

Work conditions may differ between normal work hours and the additional work time that constitutes LWH, and ascertaining those differences has not received much scholarly attention. Among professionals, overtime work hours − which are, by definition, LWH − have been shown to be less stressful than regular work hours due to the perception of overtime as uninterrupted, productive work time[76]; among manual laborers, however, extended shifts have been associated with reduced productivity, low preference, and high stress.[77,78] Depending on the population under study, the contradictory conditions experienced as a result of LWH have the potential to attenuate or inflate the measures of association if not correctly assessed in terms of job control, job satisfaction, or preference for LWH.[4]

Proposed Mechanisms of Effect for Long Work Hours

Two primary mechanisms of effect have been hypothesized to explain the impact of LWH on health: (1) the physiological recovery mechanism, and (2) the behavioral lifestyle mechanism. Both center on the impact of physiological changes resulting from the exposure to LWH, but the manner in which those changes are triggered differs between the two.[42,65,79,80] It should be noted that the proposed mechanisms are not believed to be mutually exclusive and may occur at the same time. It is also possible that each one may influence a specific set of physiological effects resulting in distinct health complaints, or they may reinforce the effects of the other, but the evidence on these hypotheses is limited.[42]

LWH is believed to interfere with physiological recovery mechanisms because the exposure requires longer time at work, thereby reducing the amount of rest and recovery time, particularly in terms of sleep hours.[81,82] It is unclear whether short sleep durations result from stress-related sleep disruptions, less time available for sleep, or both.[42] It is generally accepted that the physiological demands made by the effort required by work would be reversed if given adequate recovery time. In this situation, however, the body's physiobiological systems experience incomplete recovery and are not allowed to return to their baseline levels.[70] The remaining negative load is compounded by the effort required to maintain the same working energy the next day, and so on.[83] This leads to a greater recovery requirement,[70] which, if unmet, may result in disruption of the central nervous, endocrine, and immune systems and, eventually, physiological and psychological health complaints.[84,85]

Another hypothesis is that LWH are associated with deleterious lifestyle choices (e.g. smoking, unhealthy diet, lack of exercise), which lead to physiological changes and increased risk for adverse health outcomes.[86-88] LWH have been cited as a barrier to physical activity due to the reduction in discretionary time.[70] Long work hours have also been associated with increased odds for unhealthy weight gain,[89,90] smoking,[65] increased alcohol use,[65,89,91] and decreased physical activity.[65,92]

Whether due to insufficient recovery, to lifestyle factors, or both, the relationship between LWH and ill health is believed to be directly and indirectly mediated by stress. The stress of LWH acts directly by intensifying the demands on an increasingly fatigued worker, and it acts indirectly by extending the worker's period of exposure to other sources of workplace stress.[4] Disturbances in stress hormone excretion have been found in situations of high workloads, including LWH.[84,93,94] High levels of stress have been shown to be a contributing factor in psychiatric problems,[95] cardiovascular disease,[96] musculoskeletal disorders,[97] and gastrointestinal distress.[98]

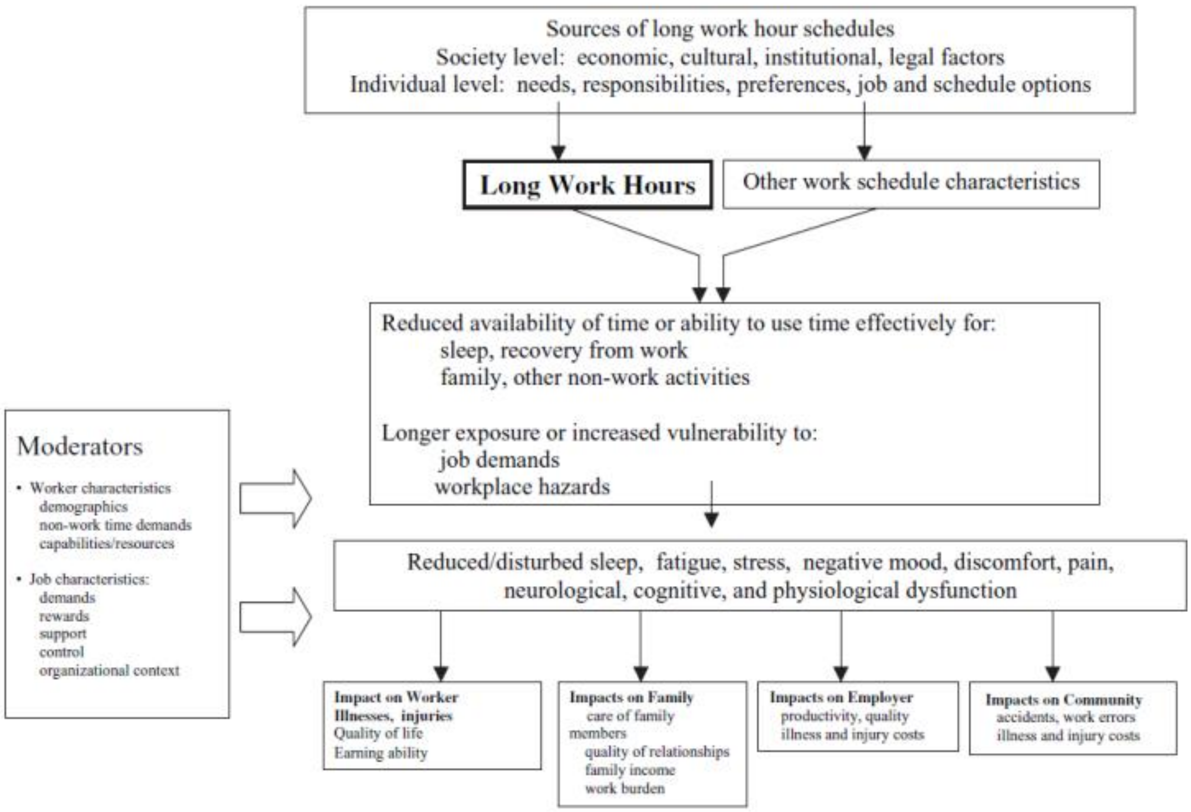
Conceptual Framework for the Association between Long Work Hours and Adverse Health Effects

In the mid-1990s, NIOSH turned its attention to issues of work organization through the establishment of the National Occupational Research Agenda (NORA), which was intended to spur a critical evaluation of occupational health in an attempt to identify areas warranting future research.[99] The subsequent report produced by the NIOSH Organization of Work (OOW) Work Group in 2002 enumerated the systemic changes that had occurred in the process and nature of work organization during the 20th Century. The OOW Work Group indicated that "researchers need to develop improved methods for measurement of work hours, give more attention to safety outcomes, and focus on populations most likely to work long hours".[11(p923)] To that end, NIOSH organized a Long Working Hours Team (LWH Team) in 2003 to clarify and address the issues raised in the OOW 2002 report.[11]

A conceptual model developed by the NORA LWH Team to promote a more thorough approach to assessing the exposure to LWH is shown in Figure 5.[10] This model advances frameworks previously introduced by Barton *et al.* (1995)[100] and the NORA OOW Work Group[101] by attempting to catalogue factors that might contribute to the detrimental effects of LWH (although the authors indicate that LWH may be beneficial, as well, in terms of potentially reducing the psychological distress associated with family roles and conflicts[102]).[33] Beginning with the society- and individual-level determinants of work schedule patterns, the framework highlights the role of increased workplace exposures and reduced recovery time in contributing to physiological changes that may result in short- and long-term health effects. Potential worker- and job-specific moderators are also listed. Most significantly, the complexity of this schematic emphasizes that simply counting the total number of hours worked is not sufficient for characterizing the work schedule exposure or evaluating the relationship between work scheduling and health and safety risks.[33]

Determining the effect of LWH on short- and long-term health requires quantitatively estimating the exposure duration, exposure intensity, and demographic and occupational

Figure 5. Conceptual framework for studying LWH



Source: Caruso *et al.*, 2006[33]

characteristics of workers at risk as accurately as possible. Although the number of publications investigating the health effects of LWH is increasing, the majority of studies have evaluated the association between hours worked and specific health outcomes without defining the relationship between LWH and work status (e.g., part-time vs. full-time, salaried vs. hourly wage), shift work, overtime, number of jobs, industry, occupation, workplace psychosocial stressors, or worker characteristics. Further, few studies have considered the influence of LWH on health- and non-health-related outcomes simultaneously, though research has demonstrated that working long hours has a negative effect on worker family dynamics[10,33] and workplace safety.[4,39,103]

Understanding and measuring the health effects of LWH is a complex task that goes beyond simply quantifying the duration of time at work. Instead, worker health is influenced by a combination of work schedule patterns, work intensity, rest time, and job control in addition to demographic variables, socio-economic factors, and family dynamics.[11] It is difficult to isolate the effect of LWH from the effect of occupational stress, in general, because LWH often occur in situations of work overload.[4] Further, time at work cannot be separated from the content and conditions of the work performed. Time spent performing repetitive, physically-demanding, poorly paid tasks is believed to affect the health of workers differently than the same amount of time spent in rewarding, well paid professional activities.[11]

Despite the contribution of conceptual models such as the one developed by the NORA LWH Team, existing frameworks for LWH have not been operationalized into exposure definitions, exposure assessment models, or analytical models to be applied to real-life data. Current approaches to LWH are one-dimensional in their focus on arbitrarily defined measures of what constitutes “long” work hours and simplistic analytical models. The disadvantages of existing work hour measures and definitions contribute to acknowledged limitations in exposure assessment measures in areas of occupational research that depend on accurate measures of time at work and, ultimately, time-at-risk.

Work Hours as an Occupational Risk Factor

Time at work is a foundational measure of occupational exposure. The duration and structure of work hours is one of the most fundamental and most easily quantifiable factors in assessing occupational exposures associated with work processes.[11] Work hours are typically utilized in occupational epidemiology as either a surrogate measure for time-at-risk for occupational exposures or are coupled with other workplace factors to develop composite measures of exposure (e.g., as a component in an interaction term or an index). As a result, the accurate measurement of work hours is a critical element when evaluating workplace exposures.[22]

The use of work hours as a measurement of occupational exposure time-at-risk is standard among federal regulatory bodies as well as industry-specific professional associations. The U.S. Occupational Safety and Health Administration's (OHSA) Permissible Exposure Limits (PELs) define the legal maximum levels of exposures to hazardous substances per eight-hour work-day, with actual exposures calculated using individuals' work hours.[104] NIOSH and the American Conference of Governmental Industrial Hygienists publish recommended exposure limits (RELs)[105] and threshold limit values (TLVs),[106] respectively, that complement OSHA's PELs and are calculated similarly. Professional associations, such as the American Industrial Hygiene Association and the International Council on Mining & Metals, instruct their member organizations to characterize potentially hazardous occupational processes and tasks by hours of work and to define periods of exposure in terms of work hours.[107]

Further, work hours function on the micro- and macro-levels as both a temporal structuring mechanism and a contributor to the larger work experience. At the individual level, work scheduling structures virtually all other aspects of day-to-day life,[108] thereby influencing social dynamics as well as individual behaviors. At the organizational level, work hours are often part of a multi-variable factor referred to as “work organization” that includes management practices, wage systems, task-level work content, and psychosocial job attributes, such as effort-reward imbalance, social support, job demands, and work control. Although previously considered negligible, these work characteristics have been seen as increasingly important to the field of occupational health over the last ten years or so.[109] Because it is recognized that the organization of work impacts all workers in every industry and that the nature of work has changed – and continues to change – dramatically over time, it has been argued that these variables must be re-evaluated to better understand their impact on workers as the standard assumptions made about working life are no longer valid.[11,109]

As the definition of standard work hours influences the definition of LWH, differences in standard work hours may account for differences in the results of some studies on the health effects of LWH. To account for this, several scholars have proposed a single definition of LWH to be applied internationally based on the convention of 40 hours of work per week and 8 hours of work per day.[1] However, the issue with such suggestions is that the recommended cut-points are not selected to accurately distinguish risk categories for the effects of LWH. Such cut-points are defined based on convention or subjective selection criteria rather than by applying statistical methods to determine the LWH value that minimizes the misclassification of risk groups through the use of large, representative data sets. Inaccurately assessed work hours may be considered a misclassification of the exposure measure, and selecting cut-points in this manner may result in misclassified risk groups.

Long Work Hours Exposure Misclassification

Scientific rationale to support the use of a specific quantity of work hours as an indicator of health risk is lacking. Studies of LWH overwhelmingly define work hours in terms of 40 hours per week, either as a dichotomized cut-point or as a category end-point in a multilevel exposure categorization. This is most likely due to the fact that 40 work hours per week is the generally accepted labor standard for full-time work in the U.S. and several other industrialized nations. Other cut points used to define LWH have tended to add hours in five- or ten-hour increments (e.g. Artazcoz *et al.* 2007 exposure categories of <30, 30-40, >40 work hours per week) or in additional 8-hour durations (e.g. Ribet and Derriennic 1999 exposure categories of ≤48 and >48 hours per week). Such selections appear to be arbitrary, as researchers very rarely explain the rationale for their categories. For example, only two of the nine studies listed in Table 5 provided any justification for their application of any particular work hours ranges. Ettner and Grzywacz 2001 stated that they utilized the exposure categories of a previous study (Ross and Mirowsky 1995) to maintain consistency, but the earlier study provided no explanation regarding its definition of work hour categories. Song and colleagues (2014) employed Korea's legal limits to define long work hours for their sample of Korean employees. The authors followed the convention set in a report from the Korean Ministry of Employment and Labor (Won *et al.* 2008) to characterize extended overtime hours.

Given that the primary goal of the studies under discussion here was to evaluate the association between LWH and health outcomes and behaviors, defining risk categories in terms of a standard labor practice and, typically, one or more accompanying arbitrarily defined categories has not provided a reasonable theoretical construct for assessing the exposure to date. The use of an industry standard would be appropriate if the research questions focused solely on assessing the frequency with which individuals worked LWH. However, the use of an industry standard is of questionable validity for evaluating the health risks associated with LWH, as there is no evidence that this method of attributing work hours risk accurately measures that risk. A single cut-point dichotomizing work hours into standard hours vs. extended hours suggests two risk categories: low risk and high risk. Similarly, multilevel exposures suggest risk categories relative to the referent category.

Research on the LWH−health relationship frequently employs the duration of 40 work hours per week as a touchstone in measuring occupational exposure and characterizing health risks due to LWH. However, the reliance on a general industry standard − e.g., the 40 hour workweek − to delimit health risk is inherently flawed because it has been neither created nor validated for such a purpose. This proposed research project is premised on the hypothesis that the inconsistent results and generally weak measures of effect seen in the LWH−health relationship may result from exposure misclassification due to the unfounded use of a generally accepted workplace practice as a measure of health risk.

Exposure misclassification is a form of information bias that can affect the assessment and interpretation of the exposure-outcome relationship.[110] Non-differential exposure misclassification is present when errors in exposure classification are unrelated to an individual's outcome status, and, thus, all individuals have the same probability of being misclassified relative to their exposure status. Previous studies of the health effects of LWH have employed arbitrary and unfounded work hours categories to assign health risk status, but those studies have done so consistently across their unique study populations. Because the exposure misclassification is independent from outcome status in these studies, it is non-differential. Non-differential misclassification of binary exposures attenuates the relationship between the exposure and the outcome.[111] Many studies of LWH have applied a single cut-point to define standard versus long work hours (e.g., Tarumi *et al.* 2003 work hour categories of <45 and ≥45 hours per week; Shields 1999 work hour categories of ≥35 to <41 and ≥41 hours per week), although multiple exposure categories are more common. Similarly, non-differential misclassification tends to reduce the measures of effect and bias results towards the null when the exposure is assessed using multiple categories (e.g., Ettner and Grzywacz 2001 work hour categories of <35, 35-45, and >45 hours per week; Virtanen *et al*. 2012 work hour categories of 7-8, 9, 10, 11-12 hours per day) or as a continuous variable (e.g., Jex and Bliese 1999).[112] Exposure misclassification errors as small as 10 to 20 percent have been shown to have a large impact on risk estimates.[113] The effect of non-differential exposure misclassification on measures of association has been estimated to be more than 30 percent under most circumstances, and up to 70 percent.[110] It is suggested, then, that the misclassification of exposure status may have diluted the effect of LWH, resulting in an underestimation of the strength of the work hours−health association in the existing literature.

Thus, a validated threshold of effect for LWH is needed and has been called for by numerous researchers.[4,33,34,114,115] Based on his review of the literature, Harrington (1994) concluded that "working weeks in excess of perhaps 48-56 hours a week are 'harmful'."[114(p703)] In response to Harrington's statement, Folkard (1994) argued that the existing evidence was not adequate to establish specific recommendations on work hour risk thresholds,[115] a position supported by Sparks and colleagues (1997), who contended that conclusions such as Harrington's (1994) were "based…on a purely qualitative assessment of the studies reviewed" rather than on statistically-derived rationale.[34(p393)] Instead, Sparks *et al.* (1997) performed a meta-analysis of the existing data based on the hypothesis that "health symptoms may only arise if the hours worked exceed a particular number," which appeared to be 48 hours per week given their results.[34(p401)] Later that year, Spurgeon and colleagues (1997) posed the question "at what level of overtime do effects begin to occur?" and responded that "hours which exceed 50 are associated with increased occupational stress."[4(p371)] Given the limitations of the literature, however, they stipulated that the "adoption of 50 hours as a threshold…might be viewed as somewhat arbitrary or perhaps cautious"[4(p371)], and they concluded that "currently available data are insufficient to determine exactly how many hours people should be required to work if they are to remain safe and healthy."[4(p374)] More recently, Caruso and colleagues (2006) argued that flawed methodological approaches had precluded the identification of a valid threshold for safe work hours and indicated that "a wide array of exacerbating and mitigating factors often makes it difficult to establish safe limits for working hours."[33(p931)]

Health-oriented exposure assessments and risk analyses based on the presumption of a standard 40-hour workweek are disconnected from the health outcomes they are intended to investigate. An improved measure of LWH would advance our understanding of the profile of risk associated with prolonged work hours, particularly when considered with industry- or occupation-specific exposures. This research proposes developing the first statistically optimized exposure cut-point for LWH based on a large, population-based sample of U.S. workers. The resulting analytical model represents an unusually comprehensive method of assessing the occupational exposure of LWH, the health effects of which are widely studied, but poorly understood due to inconsistent, over-simplified, and systematically erroneous methodological approaches.

Public Health Significance

The proportion of U.S. workers working extended hours has increased substantially over the last 30 years.[116] It is estimated that almost one-third of the U.S. workforce regularly works more than a 40-hour workweek and one-fifth works more than 50 hours.[117] Combined with the increasing reliance of employers on extended schedules, this trend has raised concerns about the risks to worker health and safety if work hours become excessive. Incomplete knowledge about the risk categories of LWH and insufficient exposure assessment models for evaluating the effect of LWH on health may have contributed to a substantial human and economic burden from exposure-related decreased quality of life, declining life expectancies, increased healthcare costs, job-related injuries, lost productivity, and greater employee turnover.[77,78] This research took a unique approach to studying the work hours–health relationship, and the results make a significant contribution to improving work policy, employment practice, and future research on workplace exposures.

Specific Aims

This study focused on the duration and health outcomes of hours worked by U.S. workers to assess and define work hours-related health risks. The purpose of the study was to identify a threshold value for exposure to LWH beyond which adverse health outcomes were detected and that could be used to characterize the relationship between LWH and SRGH, which represented a much more comprehensive exposure assessment methodology for LWH than previously existed in the literature. This work was conducted using de-identified data from the U.S. Panel Study of Income Dynamics, a nationally representative household panel survey with more than 40 years of longitudinal data and a current enrollment of more than 22,000 individuals.

The specific aims were used to address the overall research hypothesis that LWH increases the risk of adverse health conditions. This study had the following specific aims and corresponding hypotheses:

Aim 1. Identify the statistically optimized cut point of LWH that most accurately predicts increased risk of poor SRGH status after long-term LWH exposure.

Aim 2. Determine the applicability of a general health based cut point on cardiovascular disease (CVD) risk and cancer risk by (1) identifying statistically optimized cut points for CVD and cancer and (2) assessing the validity of these cut points against that of the cut point derived in Aim 1.

<u>Hypothesis corresponding to Aims 1 and 2</u>: An identifiable threshold exists beyond which long-term exposure to LWH is detrimental to workers' physical health.

<u>Aim 3</u>. Assess the presence of a dose-response relationship between working hours and CVD.

<u>Hypothesis corresponding to Aim 3</u>: A dose-response relationship exists between LWH and CVD in which the risk of CVD follows a monotonic increasing dose response as work hours per week increases.

CHAPTER II: METHODS

The Panel Study of Income Dynamics − Parent Study

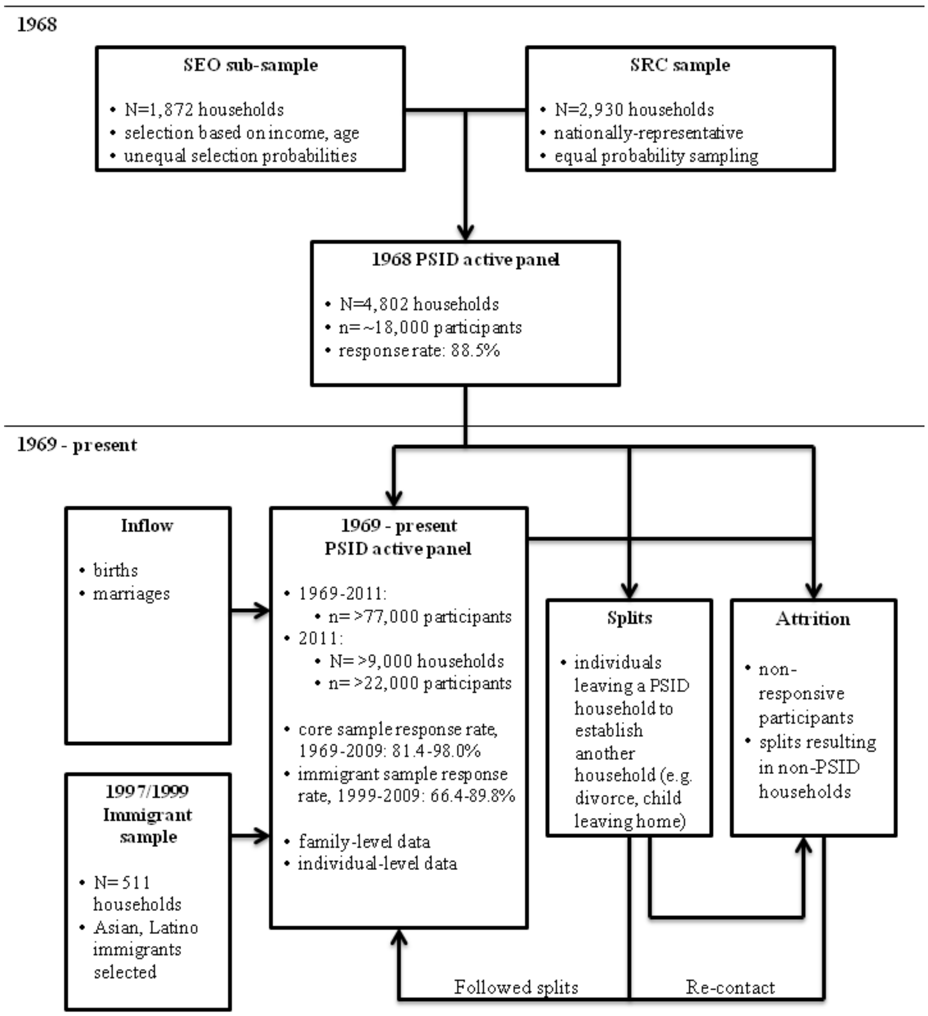
Overview

The PSID is an on-going longitudinal survey of a representative sample of U.S. residents and the family members with whom they reside.[118] Initiated by the U.S. Census Bureau in 1966 as the Survey of Economic Opportunity (SEO), the PSID was acquired and expanded by the Survey Research Center (SRC) at the University of Michigan in 1968 and has been ongoing to current date.[119] The initial sample contained over 18,000 individuals living in 5,000 families in the U.S. As children in PSID families mature and establish independent households, they and their families are enrolled into the PSID, as well. The sample size is currently more than 22,000 individuals in 9,000 families.[120]

Sampling, Tracking, and Response

The 1968 PSID sample consisted of two independent samples (Figure 1). The first sample, which was a sub-sample of the 1966 SEO cohort, was comprised of 1,872 low-income families with heads-of-household under the age of 60 years. This sample was confined to Standard Metropolitan Statistical Areas (SMSAs) and to non-SMSAs in the Southern U.S., and, as such, it involved unequal selection probabilities. The second sample was a cross-sectional, national sample drawn by the SRC that was an equal probability sample of households in the 48 contiguous states, which yielded 2,930 completed interviews.[119] To maintain the representative nature of the PSID population, a sample of 511 immigrant families was added in 1997/1999.[118] This sample was comprised of Latino and Asian immigrants, which were the two largest subsets of immigrants to the U.S. since

Figure 1. Diagram of the PSID sampling frame



Source: Adapted from Ko, 2013[121]

1968.[120] The combined PSID sample − which contains the SRC, SEO, and immigrant samples − is a sample with unequal selection probabilities and differential non-response over the survey waves. Compensatory weighting is needed to adjust for these factors in statistical analyses, and PSID data files contain longitudinal weights for this purpose.[119]

The PSID's tracking rules stipulate that all members of the original 1968 family units and their descendants (by birth or adoption) are to be followed to any new, non-institutional living situations. Institutionalized participants are re-contacted once they leave institutional housing based on information provided by their families. Sample members over the age of18 years are eligible for inclusion as separate family units if they establish economically independent households, such as in the case of children leaving home. Family members who are not sample members (e.g. the spouse of a sample member) will not be tracked if they establish an independent household, such as in the case of divorce.[119]

The response rate was 88.5 percent in the first wave of the PSID (1968),[119] 81.4 percent in the second wave (1969), and has ranged from 93.0 to 98.0 percent for the core sample since that time.[118] Prior to 1993, non-responsive participants were considered lost-to-follow-up[119]; since 1993, non-responders have been re-contacted in subsequent waves in an effort to increase retention rates.[118] The cumulative rate of attrition is estimated to be approximately half of the original 1968 sample.[122] Despite evidence of higher cumulative attrition in lower-income families,[122,123] recent assessments of the representativeness of PSID data have shown that PSID parameter estimates are largely unbiased by the effects of non-response, attrition, and non-coverage (e.g., Fitzgerald 2011; Gouskova *et al.* 2010; Bosworth and Anders 2008; Li *et al.* 2010) and that weighted PSID data achieves a high level of national-representativeness.[123]

Informed Consent

Informed consent is obtained prior to beginning each interview in each study wave. Respondents are informed that participation is voluntary and that they may refuse to answer questions or decline having their interview recorded.[124] In addition, the Health Sciences and Behavioral Sciences Institutional Review Board of the University of Michigan reviews the PSID data collection and distribution protocols and survey instruments annually to ensure the ethical treatment and protection of research participants.[118]

Data Collection

The PSID main interview was performed annually between 1968 and 1997 and has been conducted biennially since 1997. Information is collected on each family member, with additional details gathered on the heads of household and their spouses or long-term cohabitators, if present.[118] Although initially gathered via face-to-face interview, data have been collected by telephone since 1973.[119] PSID was historically an economically-based survey; however, a health survey component was added in 1999.[125] Most data from 1968 to 2011 are de-identified and publically available. Survey question topics include socio-demographic characteristics, employment status and industry/ occupation information, health status, income and finances.[119]

The Current Study

Overview

The purpose of this study was to examine the effects of LWH on health outcomes among a large cohort of U.S. workers. LWH has been positively associated with various adverse health outcomes,[35,37-39,40,41] but there is a lack of a standard, accepted framework for measuring work hours as an occupational risk factor.[1,33] This study aimed to address some of the methodological issues, such as inconsistent exposure definitions, in order to better demarcate the potential risks of exposure to LWH on worker health. More specifically, this research was intended to identify a threshold value for exposure to LWH beyond which adverse health effects are detected as measured by self-reported general health. Next, the identified cut-point of work hours (that infers LWH) associated with SRGH was tested with other health outcomes — specifically, CVD and cancer — in this cohort. This assessment allowed us to determine if the definition of LWH was unique to SRGH or if it was applicable across outcomes and, thus, broadly generalizable to multiple health-related outcomes. Finally, work hour data was utilized to examine the form or distribution of its relationship with CVD to assess if a dose-response relationship exists. These efforts characterized the exposure much more comprehensively than had previously existed in the literature.

Study Population and Inclusion/ Exclusion Criteria

The current study was a population-based observational study that employed a retrospective cohort design. A longitudinal sample of U.S. workers from the PSID was used to develop a more consistent method of characterizing work hours exposures. Four subcohorts of the total sample were defined based on the inclusion criteria listed below. Because the number of participants meeting each set of inclusion criteria varied, the size of each sub-cohort varied, as well.

Inclusion Criteria

PSID respondents from 1986 to 2011 were included in this study if they (1) were 18 years of age or older in 1986 and (2) reported non-zero work hour data for at least 10 years between 1986 and 2011. A total of 3,596 of 60,214 PSID participants met our initial study inclusion criteria. Because little is known about the effect of LWH over specific exposure durations (e.g., one year vs. five years vs. 15 years of exposure), the study duration of 25 years was selected to ensure that the study data captured a long induction period for the development of any potential adverse health outcomes correlated with LWH exposure. Including only those participants reporting at least 10 years of work hour data was intended to provide sufficient information on each individual's work hour trends over the study period.

For the outcome of SRGH, from the 3,596 participants meeting the initial inclusion criteria, twenty-seven were excluded for lacking work hour data for the years prior to a report of poor SRGH. We further excluded 22 participants who were missing responses to the SRGH question in more than five waves and 300 participants reporting poor SRGH at baseline (i.e., prevalent cases). Additionally, the study was restricted to individuals whose average work hours over the study duration were equivalent to full-time employment (defined as an average of at least 35 hours of work per week),[126,127] resulting in the exclusion of 1,041participants. Limiting the analysis to those individuals who averaged at least 35

WH/w over the study duration seemed appropriate given that our exposure of interest is LWH, which, by definition, refers to daily or weekly durations of work exceeding a threshold of work hours generally defined in terms of the standard work hours of a full-time employee.[1] These exclusions resulted in a cohort of 2,206 for the Aim 1 analysis.

For the outcome of CVD, participants also were required to have either a complete set of responses to the CVD incidence question for each wave of data collected during the study period or a complete set of responses to those questions prior to reporting an affirmative response to CVD incidence. Participants reporting CVD status at the baseline health study (1999) were excluded from this analysis. The Aim 2 CVD cohort was comprised of 1,806 participants, with 777 cases of incident CVD (or, 43.0%), after excluding 276 participants due to missing CVD data, 658 prevalent cases of CVD, and 856 participants reporting average weekly work hours of less than 35 hours from the 3,596 participants meeting the initial inclusion criteria. The Aims 2 and 3 CVD cohorts were constructed independently and were of different sizes. This was because the Aim 3 cohort was not restricted to full-time workers, as was the case with the Aim 2 cohort, but it did exclude individuals reporting other chronic prevalent conditions at the baseline health study (1999), which was not the case with the Aim 2 cohort. Thus, the Aim 3 CVD cohort was comprised of 1,926 participants, with 822 cases of incident CVD (or, 42.7%), after excluding those missing CVD data (n = 364), those reporting CVD at the baseline health study in 1999 (n = 658), and those reporting other prevalent chronic health conditions or disabilities at baseline (n = 529) or missing those data (n = 119) from the 3,596 participants meeting the initial inclusion criteria.

For the assessment of incident cancer status, participants were required to have either a complete set of responses to the cancer question for each wave of data collected during the study period or a complete set of responses to that question prior to reporting an affirmative response to a cancer diagnosis. Participants reporting cancer diagnoses (excluding skin cancers) at the baseline health study (1999) will be excluded from this analysis. For the Aim 2 cancer cohort, 2,196 participants met our criteria, of whom 263 (12.0%) reported incident cancer, after excluding 116 participants for missing information on cancer status, 205 prevalent cases of cancer, and 1,079 participants reporting average weekly work hours of less than 35 hours from the 3,596 participants meeting the initial inclusion criteria.

Study Variables

Long Work Hours

Measurement and Definition of Work Hours in the PSID: The PSID captures a comprehensive picture of time worked by participants, including questions on working status (PSID waves: 1968-2011), annual hours spent working (PSID waves: 1968-2011), annual overtime hours (PSID waves: 1985-2011), and number of jobs and hours worked by job (PSID waves: 1983-2011) (Table 1). Information is captured on each employed participant's industry and occupation and is coded according to 1970 Census Bureau codes (PSID waves: 1986-2001) or 2000 Census Bureau codes (PSID waves: 2003-2011). The industry and occupation variables will be transformed to create a consistent set of values by using cross-walks provided by the U.S. Census Bureau. The start and stop dates of each job that participants were working at the time of the interview or had worked since the previous

Table 1. PSID working status and work hour measures, 1986-2011*

PSID survey questions
• We would like to know about what you do--are you working now, looking for work, retired, keeping house, a student, or what?
• In your work for [EMPLOYER NAME], what (is/was) your occupation? What sort of work (do/did) you do? What (are/were) your most important activities or duties?
• What kind of business or industry was that in?
• When did you start and when did you stop working for this employer? Please give me all of the start and stop dates if you have gone to work for [THIS EMPLOYER] more than once. (*captured by month and by job for up to four jobs*)
• How many weeks out of the year did you actually work on this job in [PREVIOUS YEAR], not including any time off that you told me about earlier?
• On average, how many hours a week did you work on all of your jobs during [PREVIOUS YEAR]?
• On the average, how many hours a week did you work on this job? (*captured by job for up to five jobs*)
• How many hours a week did that overtime amount to in [PREVIOUS YEAR]?

* = PSID waves from 1986 to 1997 occurred annually; PSID waves from 1999 to 2011 occurred every other year

interview were ascertained, including the month and year of employment for each job (PSID waves: 1984-2011). Validation studies have demonstrated a strong positive correlation (Pearson's $r = 0.64$) between cross-sectional PSID self-reported total hours worked in the previous year and weekly work hours reported by employers.[128]

Definition and Construction of Current Study LWH Exposure Variable: The average number of hours worked per week on all jobs in the previous year was calculated by dividing the total number of hours worked per year by the total number of weeks worked in that year. The total number of hours worked in a given year included hours of work on participants' main jobs, hours of overtime, and hours on extra jobs.

The arithmetic mean of participants' self-reported weekly work hours for each year in which they report non-zero work hours was calculated to create a summary average weekly work hours variable. This summary average variable was considered an index of the participants' exposure to LWH over a minimum of 10 years' time. Using the arithmetic mean of past LWH exposure as a summary measure of exposure assumes slowly or partially reversible physiological processes.[129] This was appropriate given the mechanisms of effect hypothesized for LWH, which center on the body's inability to repair LWH-related physiological damage due to the lack of sufficient recovery time, resulting in the gradual accumulation of effect. The use of an arithmetic mean to represent exposure over time assumed homogeneity among the years of exposure, however, which precluded more detailed examinations of exposure patterns, such as short term, high exposure periods.[129]

For Aims 1 and 2, whether an individual was categorized as exposed or unexposed to LWH depended on the specifics of the statistical test being performed. A series of tests were conducted to determine the statistically optimized cut point for LWH, and participants were considered exposed to LWH if their average weekly work hours summary variable was greater than the dichotomized cut point being tested.[130] The testing was executed against a comprehensive set of potential cut points from 36 to 65 work hours per week. The definition of LWH differed in each testing sequence, as a different dichotomous value for LWH was defined for each test (i.e. ≥45 hours per week, ≥46 hours per week, ≥47 hours per week, and so forth). Each participant was categorized as exposed or non-exposed based on the threshold value of LWH being analyzed.

For Aim 3, work hours was treated as a continuous variable in the analysis of the dose-response relationship between LWH and CVD. Thus, the summary average weekly work hours variable was used as the exposure variable, and the shape of the functional form was described using a series of nested logistic regression models.

Self-Reported General Health

Measurement and Definition of SRGH in the PSID: General health was defined as a self-reported, categorical variable in which a participant summarized his or her overall health status. Since 1984, PSID participants have been asked if their health, in general, is "excellent, very good, good, fair, or poor" (PSID waves: 1986-2011) (Table 2). In an assessment of the quality of the health data captured in the PSID, Andreski *et al.* (2007) concluded that measures of self-reported general health status align fairly closely between the PSID and the U.S. Health and Retirement Study, which is another widely used longitudinal panel survey focused on employment and health. Andreski and colleagues also determined that multivariate models applied to data from the PSID and the U.S. National Health Interview Survey, which is the most widely used nationally representative health survey, produced similar parameter estimates for the explanatory factors associated with SRGH, including age, race, sex, marital status, and education.[125]

Table 2. PSID self-reported general health measure, 1986-2011*

PSID survey question
• Would you say your health in general is excellent, very good, good, fair, or poor?

* = PSID waves from 1986 to 1997 occurred annually; PSID waves from 1999 to 2011 occurred every other year

Definition and Construction of Current Study SRGH Variable: A dichotomous variable was constructed from the PSID responses to the question about general health (as detailed above), with good SRGH defined as participant responses of “excellent”, “very good”, and “good”, and poor SRGH defined as responses of “fair” or “poor”, as detailed in Table 3.

Table 3. Construction of primary outcome variable

Variable	PSID Coding	Proposed Study Coding
self-reported general health	1. Excellent 2. Very Good 3. Good	0. Good
	4. Fair 5. Poor	1. Poor

Self-Reported Physical Health Conditions

Measurement and Definition of Self-Reported Physical Health Conditions in the PSID: Beginning in 1999, PSID participants were surveyed to determine whether they been diagnosed by a physician with specific physical health conditions (PSID waves: 1999-2011) (Table 4).[120] Studies on the reliability and validity of interview data on chronic physical health conditions have demonstrated varied agreement between self-reports of physician diagnoses and diagnoses extracted from medical records, insurance records, or disease registries, or made during health examinations. In general, the overall agreement between self-responses and health records has been reported as moderate,[131] with excellent agreement for cardiovascular disease (including hypertension)[131-133] and stroke,[133] moderate to excellent agreement for diabetes[132-135] and depression,[136] moderate agreement for cancer,[137] and relatively low agreement for hypercholesterolemia,[132] osteoarthritis/ rheumatoid arthritis,[138] respiratory disease,[131,135] and musculoskeletal disease.[131]

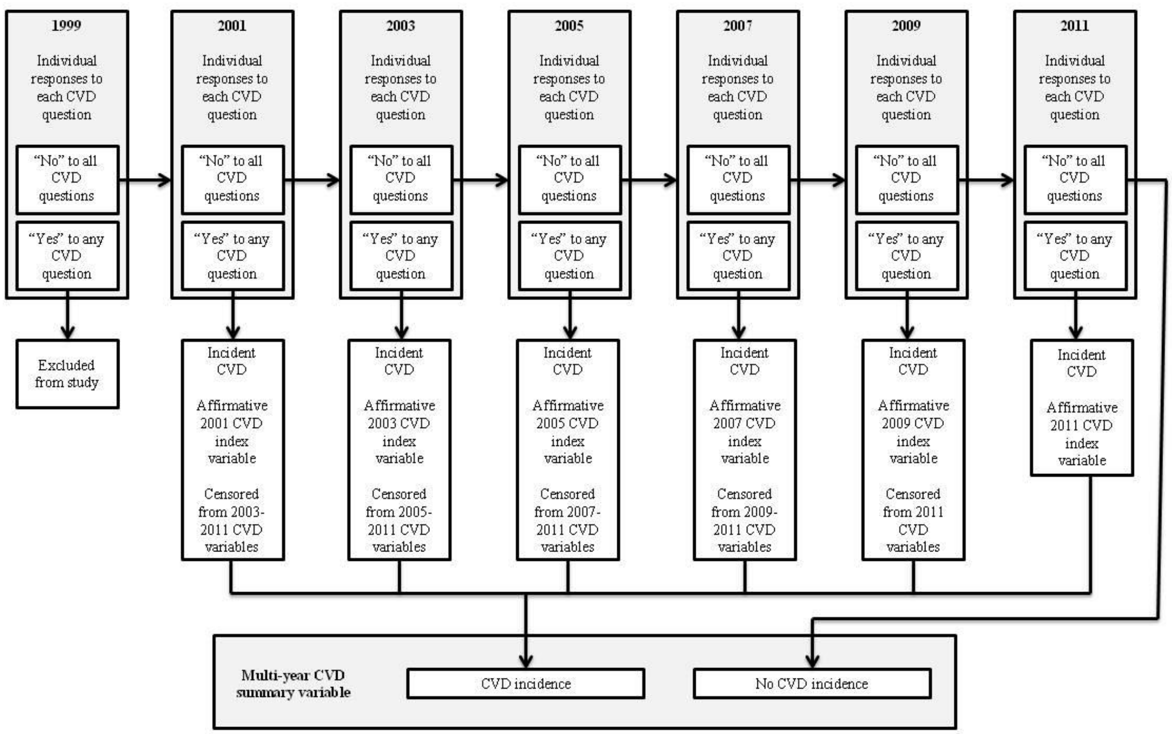
Table 4. PSID specific physical health measures, 1986-2011*

PSID survey questions	Waves available
Has a doctor ever told you that you have or had any of the following— • Arthritis or rheumatism • Asthma • Cancer or a malignant tumor • Chronic lung disease • Coronary heart disease, angina, or congestive heart failure • Diabetes or high blood sugar • Heart attack • Hypertension or high blood pressure • Other serious chronic conditions • Stroke	1986-2011

* = PSID waves from 1986 to 1997 occurred annually; PSID waves from 1999 to 2011 occurred every other year

Definition and Construction of Current Study CVD Incidence Multi-Year Summary Variable: Incident CVD during the study period was defined by an affirmative multi-year CVD summary variable. This multi-year summary variable was generated from each year's CVD index variable beginning in 1999. The yearly CVD index variables were created from individual responses to four PSID questions on the following CVD-related outcomes: heart attack occurrence; coronary heart disease, angina, or congestive heart failure diagnoses; hypertensive status; and stroke occurrence. The PSID questions on CVD-related outcomes are all phrased as "Has a doctor ever told you that you have or had any of the following?" and require dichotomous responses. A dichotomous index variable for CVD was generated for each year based on participant responses, with affirmative answers to any CVD-related question defined as a positive CVD outcome in a given year. From these index variables, a dichotomous CVD multi-year summary variable was constructed from participant responses between 2001 and 2011 (per the study's inclusion criteria, participants reporting CVD at baseline in 1999 were excluded) to represent CVD incidence status over the study period. A schematic outlining the construction of the CVD outcome variables is provided in Figure 2.

Figure 2. Construction of CVD incidence multi-year summary variable



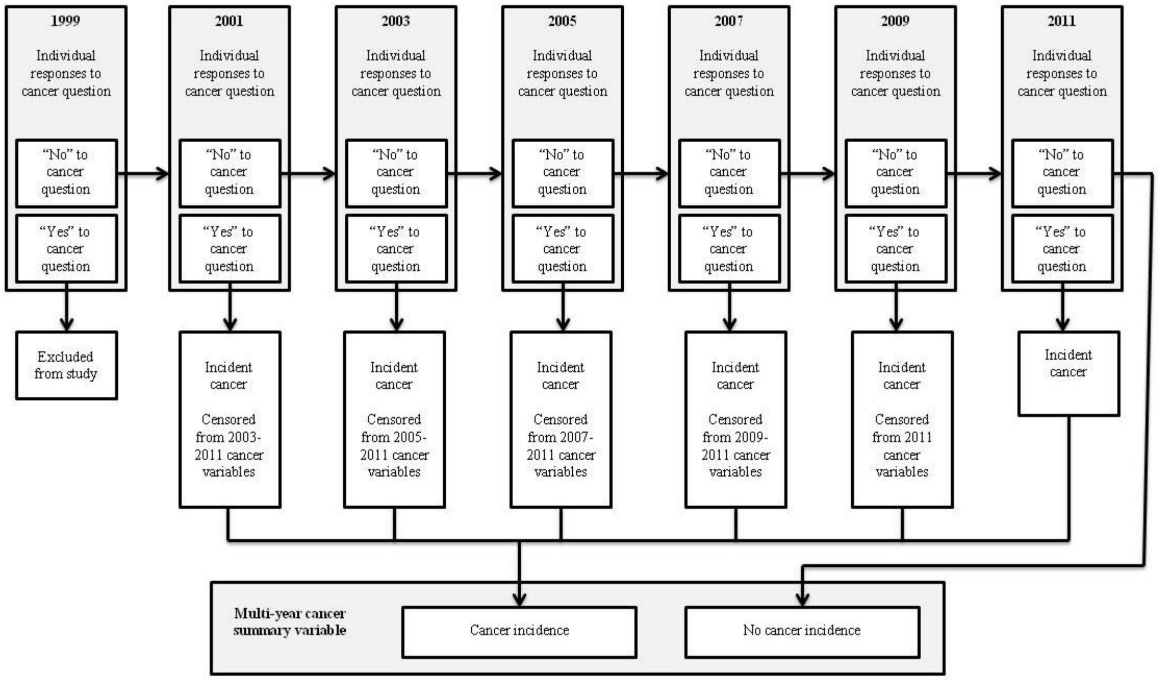
Definition and Construction of Current Study Cancer Incidence Multi-Year Summary Variable: Analogous to the variable construction for incident CVD, incident cancer (excluding skin cancers) was defined by an affirmative multi-year cancer summary variable. This multi-year summary variable was generated from participant responses to each year's PSID question on physician-diagnosed cancer. This PSID question on cancer outcomes is phrased as "Has a doctor ever told you that you have or had any of the following: cancer or a malignant tumor?" and requires a dichotomous response. An affirmative answer to the annual physician-diagnosed cancer question was defined as a positive cancer outcome in a given year. From these yearly cancer variables, a dichotomous multi-year summary variable will be constructed from participant responses between 2001 and 2011 (per the study's inclusion criteria, participants reporting physician-diagnosed cancer at baseline in 1999 were excluded) to represent cancer incidence over the study period. A schematic outlining the construction of the multi-year mental cancer summary variable is provided in Figure 3.

Data Analysis

Hypothesis 1

To distinguish the threshold values of average weekly work hours that increased the risk of poor SRGH status, incident CVD, and incident cancer after long-term exposure, we developed a series of univariate predictive models to identify a statistically optimized LWH cut point, which required evaluating the accuracy of a set of cut point candidates using measures of model fit, calibration, and discrimination. Because predictive models

Figure 3. Construction of incident cancer multi-year summary variable



incorporate a time dimension not required in diagnostic tests, model accuracy must be evaluated in terms of risk estimation (or, calibration) as well as risk classification (or, discrimination). In contrast, the evaluation of diagnostic models often relies primarily on the area under the receiver operating characteristic (ROC) curve, or AUC, to characterize a test's ability to discriminate diseased from non-diseased patients. However, the AUC has limited utility in assessing predictive or prognostic models.[139] In such models, calibration is an important component in evaluating risk factors, which may have a minimal impact on the AUC yet be important predictors of risk.[139,140] Calibration evaluates the agreement between predicted probabilities and actual observed risk.[139] Measures such as Somers' *D* statistics and risk ratios can be used as calibration metrics to assess the predictive power of the exposure on the outcome and estimate the probability of the outcome given the exposure.[141,142] Measures of discrimination assess the model's ability to classify the disease state and include such statistics as the Youden Index, AUC, sensitivity, specificity, and likelihood ratios.[141]

A comprehensive set of potential work hour cut points from 36 to 70 WH/w (in one-hour increments) was compiled in order to test the effect of each specific exposure threshold on the outcomes of interest. The definition of LWH differed as each cut point was tested, and participants were classified as exposed to LWH if their mean work hours per week across the study duration was greater than the dichotomized cut point being assessed. The relative frequencies of participants reporting the outcomes of interest at each cut point were determined. Small cell sizes precluded the analysis of cut points above 65 WH/w.

Administrative censoring of those participants reporting an outcome event occurred following the first report, and no subsequent data specific to that outcome were analyzed on those individuals. Additionally, work hour data for cases was censored in the year in which the incident case was reported and for all subsequent years. We reasoned that participants reporting outcome events were likely to have altered their work hour patterns in the last year as a result of their health status. Censoring in this way also preserved the temporality of the exposure–outcome relationship.

A series of univariate predictive models was developed to distinguish the threshold of average work hours at which the risk of an outcome event increased after long-term exposure to LWH. First, model fit of each exposure level cut point was assessed using the Bayesian

Information Criteria (BIC), which allows the comparison of models that are not nested, with lower BIC values indicating better fitting models.[143]

Next, Somers' *D* statistics and relative risks (RRs) were calculated to provide general indices of calibration for each cut point model. Somers' *D* statistics were used to determine the predictive power of the dichotomized exposure levels on the binary outcome by calculating the difference in the proportions of outcome events in the exposed and unexposed populations.[142] Univariate RRs were calculated with Poisson regression and used to determine the probability of the outcome of interest given the exposure of interest, which was defined as a value at or above a work hours threshold.[139] Because not all participants had equal exposure durations, an offset variable was constructed by dividing the number of waves of work hour data reported by the number of possible waves. For all Somers' *D* statistics and RRs, one-sided 95% confidence intervals (CIs) were calculated. One-sided analyses were performed because the study focused on possible increases in the risks of adverse health outcomes associated with working long hours; a protective effect of long work hours was not considered in this study and is not generally supported by the literature.[33] Although the Hosmer-Lemeshow statistic is often used as a measure of calibration, it was not an appropriate choice for these data, which were analyzed with probability weighting.[141]

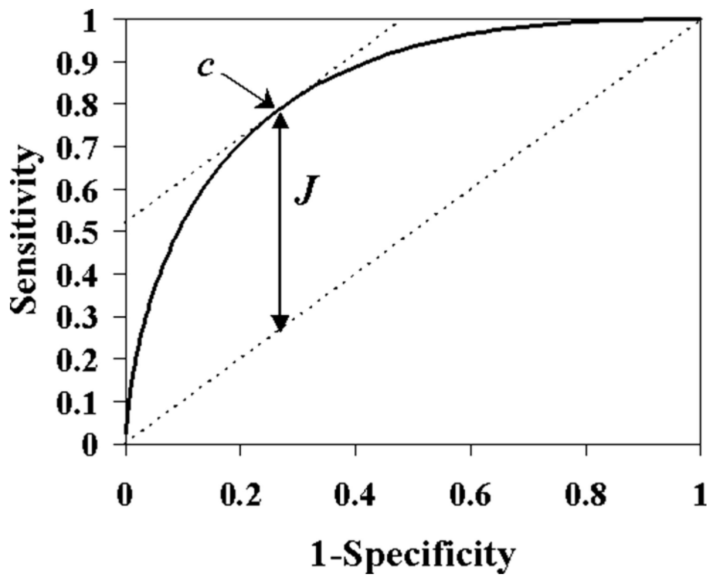
Finally, the Youden Index (*J*), AUC, sensitivity (s_n), and specificity (s_p) were calculated to provide information on the discrimination of each model. The Youden Index is a function of sensitivity and specificity and represents the maximum vertical distance between a ROC curve and the line of no discrimination.[144] When comparing cut points, the candidate corresponding to the largest magnitude of *J* is generally interpreted as the point on the curve farthest from chance and, thus, defined as the statistically optimal cut point.[144,145] Calculating the Youden Index for a given cut point produces a value ranging from 0 to 1, with 0 indicating that the cut point has no ability to differentiate between the diseased and the non-diseased and 1 indicating that no misclassification errors are present.[146] Additional measures of discrimination included positive and negative likelihood ratios (L_{pos} and L_{neg}, respectively) and positive and negative predictive values (PPV and NPV, respectively).

The statistically optimized cut point was defined *a priori* as the cut point that maximized the Youden Index, which is a commonly used measure of the overall effectiveness of a cut point dichotomized from a continuous exposure.[145] *J* is an identifiable value on a Receiver Operating Characteristic (ROC) curve, which plots sensitivity against 1−specificity for all possible dichotomized cut points of a continuous risk indicator and a dichotomous outcome. The diagonal line through the points (0,0) and (1,1) on a ROC curve represents the line of no discrimination, which is equivalent to the random assignment of outcome status.[147] The Youden Index represents the maximum vertical distance between the ROC curve and the line of no discrimination (Figure 4) and is defined as

$$J = \text{maximum}_c \{\text{sensitivity } (c) + \text{ specificity } (c) - 1\}$$

for all potential cut points c.[145] Thus, *J* is equivalent to the sum of the two proportions of correctly identified cases and controls, diminished by the value of one.[146] Calculating the Youden Index for a given cut point produces a value ranging from 0 to 1, with 0 indicating that the cut point has no ability to differentiate between the diseased and the controls and 1 indicating that no misclassification errors are present.[146]

Figure 4: ROC curve illustrating Youden Index J for cut point c



Source: Schisterman *et al*., 2005[145]

The cut point θ occurs at the value of J that maximizes the number of correctly classified individuals while giving equal weight to errors of sensitivity and specificity.[147] Thus, when deciding among cut point candidates, θ represents the cut point that maximizes risk difference between the probability of the outcome of interest for levels of exposure above and below the cut point and would be considered the optimal cut point.[144,145] In the present study, θ will be defined as the cut point for which J is closest to the value of 1.

To compensate for restricting the cut points in these analyses to natural numbers, a final analysis was performed treating WH/w as a continuous variable instead of dichotomizing it against a specific cut point. Using the CUTPT module in Stata with the Youden method selected, a statistically optimal cut point was generated for the continuous WH/w exposure variable versus each of the binary outcome variables. We had reasoned that a whole number cut point would be most useful for evaluating exposures to LWH in work settings. Although this step provided additional information on the best fitting cut points given these data, it was not intended to be a determining factor in identifying the statistically optimized cut points for any of the outcomes of interest.

Once the LWH cut points were identified, descriptive analyses (e.g., means, percentages) were conducted to define the study populations by demographic and occupational characteristics according to their exposure status. The distribution of LWH was evaluated using univariate analyses (e.g.. *t* tests, chi square tests) to examine unadjusted differences between participant groups by exposure status.

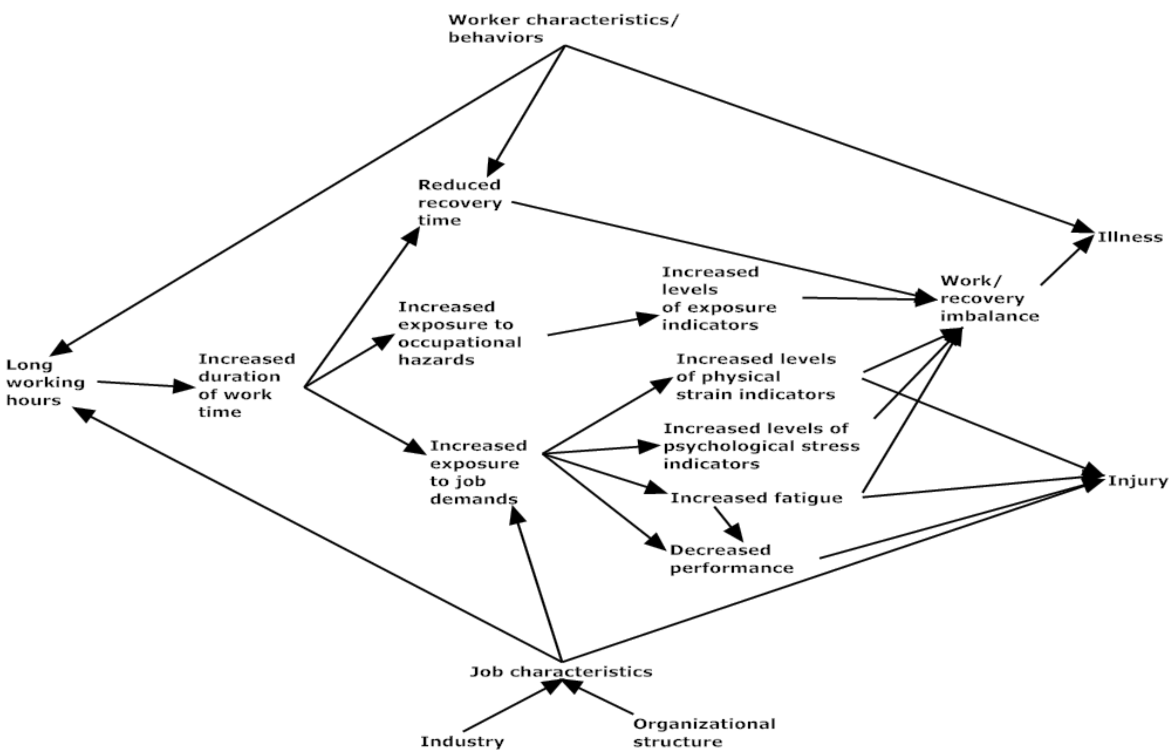
In order to generate cut points that could be applied to U.S. workers across industries and occupations, no covariates were included in these analyses.[148] Introducing variables other than the exposure and outcome into the analyses would have produced cut point values that were conditioned on those covariates, which would have limited their generalizability. The current approach calculated threshold values that can be used as general guidelines for U.S. workers. Further, omitting other covariates avoided the introduction of bias that might have occurred if additional covariates had been associated with the exposure and/ or the outcome of interest.

Hypothesis 2

To estimate the presence of a continuous dose-response function between work hours and CVD, we began by calculating descriptive statistics for selected socio-demographic and occupational characteristics at baseline by outcome status. Adjusted Wald tests and χ^2 statistics were used to compare the means of continuous covariates and prevalences of categorical covariates, respectively, to determine whether there were significant differences between the participants who developed CVD and the ones who did not.

In order to determine which covariates should be tested for inclusion in the multivariate model and the subsequent dose response modeling procedure, a directed acyclic graphic (DAG) was developed during a preliminary analysis to guide the generation of a parsimonious multivariate model (Figure 5). Potential covariates were drawn from several existing conceptual frameworks,[33,99,149,150] and a minimum sufficient adjustment set was derived based on the procedure outlined in Fleischer and Diez Roux 2008.[151] It was determined that the minimal sufficient adjustment set for mulitvariate modeling of LWH and illness was composed of workers' demographic and behavioral characteristics and job characteristics. A series of univariate sub-DAGS were constructed to assess the unadjusted exposure-covariate-outcome relationships. A comprehensive list of potential covariates corresponding to workers' demographic/ behavioral characteristics and job characteristics was compiled from the literature,[1,33,99,149] and a set of PSID variables were selected as possible covariates in the multivariate model. However, many of the DAG-identified risk factors were available in the data set, and some underfitting was expected due to the limitation of the available variables. Factors that were incident outcomes or collinear with LWH were eliminated from consideration, leaving variables that occurred prior to or at the same time as the exposure. Each remaining covariate was evaluated to determine whether it was associated with the outcome of interest. Covariates for which most – but not all – of the repeated measures over the study duration were associated with the outcome (e.g., occupation, paid hourly vs. salaried, health insurance status) were evaluated to determine

Figure 5. Proposed directed acyclic graph for exploring the relationship between LWH and illness



whether using the measure at baseline was as explanatory as using all years of measures. Many of the job-related variables demonstrated significant collinearity; as a result, multiple iterations of step-down variable selection were performed to determine the effect of removing each of the work-related covariates from the rest of the set. All regression models were adjusted for the same factors: age, sex, race/ethnicity, and pay status (salaried vs. hourly). Tests for interactions and effect measure modification were not significant.

Next, a series of nested logistic regression models was compared.[107] The models used to evaluate the dose-response relationship were (1) the logistic regression model; (2) the quadratic logistic regression model, in which a quadratic term for the exposure was added to the logistic model; (3) the step function, in which the exposure was categorized; (4) the restricted linear spline logistic regression model; and (5) the restricted cubic spline logistic regression model. Unlike the logistic or quadratic models, the spline models allowed flexibility in modeling the dose-response curve. Splines are functions defined by piecewise polyonomials connected at the edges; the slope of each polynomial segment can differ, thereby allowing for greater flexibility of the curve to fit an unknown functional form.[152] In spline regression, the exposure is divided into a number of line segments that are connected at junction points, or "knots". Compared to conventional methods, such as linear logistic regression or step function models, linear splines have advantages in modeling continuous predictors, including flexibility in modeling the underlying functional form and easily interpretable model parameters.[152] Quadratic splines, including cubic splines, offer some of the same advantages as linear splines, and they further smooth out the connections between the splines by allowing a parabolic curve to connect the line segments at the points of intersection.[153] Restricted regression splines were utilized to reduce the instability of the estimates at the tails.[154]

The step function required the categorization of the exposure, which was done by dividing LWH into quartiles. First quartile was the reference category for step function estimates. For the restricted linear spline regression, the optimal number of segments was selected by introducing an increasing number of knots at subsequent quintiles, and the best fitting model was defined as the one that minimized the Akaike Information Criterion (AIC).[155] The AIC is a penalized likelihood that takes into account the number of parameters estimated in the model. As such, it is a variable selection criterion that balances good model fit with model parsimony. There is evidence that the AIC performs poorly with small sample sizes or if too many parameters are introduced into the model relative to the size of the sample[156]; neither of these conditions was present in our study. The same approach was used to select the optimal number of segments for the restricted cubic spline regression, with two additional models generated using the mvrs and the rc_spline programs in Stata; all cubic spline models were evaluated in terms of AIC to identify the best fitting model.

The five logistic regression models (i.e., logistic, quadratic, step, restricted linear spline, and restricted cubic spline) were then compared to determine the best fitting model among the set, which was defined as the one that balanced parsimony and model performance. Parsimony was evaluated in terms of degrees of freedom, and model performance was evaluated in terms of the Hosmer Lemeshow goodness of fit statistic and AIC.

All analyses were performed using Stata/SE software (version 13.1; Stata Corp, College Station, TX) and took into account clustering and probability weighting of the PSID data.

Study Power

Study power calculations were completed prior to the execution of this research and demonstrated sufficient power (defined as power ≥ 0.8) to detect the anticipated measures of effect for each outcome-related cohort. The sample sizes, case numbers, and exposure proportions were based on preliminary analyses of the data. For these study power calculations, statistical power was set at 80 percent and Type I error at five percent. Because the statistically optimized cut point for LWH was unknown at that time, LWH was defined (and power was assessed) as two separate values: working ≥40 WH/w, on average, over the study duration, and working ≥50 WH/w, on average, over the study duration.

For each study, the exposure and outcome measures were defined, and the inclusion criteria were applied to define the specific outcome-related cohort to be evaluated. For Aims 1 and 2, the proportion of exposed and unexposed participants could only be determined after the analysis was complete. For all Aims, the relative risks of the exposure-outcome relationships were also known only following the statistical analysis. At the conclusion of each analysis, power calculations were re-performed to determine that sufficient statistical power was present to detect the measure of effect demonstrated between LWH and that outcome. For each outcome-specific cohort, it was determined that there was sufficient power to detect the measure of effect that was found, as shown in Table 5.

Table 5. Statistical power of measures of association between LWH and the outcomes of interest

Outcome	Sample Size	Proportion Exposed (≥52 work hours per week)	Relative Risk	Power
Poor SRGH	2,206	12.9%	1.28	0.8677
CVD	1,806	12.4%	1.42	0.9985
Cancer	2,196	11.1%	1.62	0.8505

Strengths and Limitations

The PSID data were not collected specifically to address the relationship between long work hours and health, and limitations due to this fact will be acknowledged during the execution of this study. Although PSID participants comprise a population-based, nationally-representative U.S. sample, the PSID cohort is not specifically sampled to produce a representative sample of U.S. workers. The PSID collects self-reported average work hours for the entirety of the previous year but lacks information on scheduling characteristics, such as shift work, non-standard work weeks, or frequency of schedule changes. Whether there will be issues in comprehensively assessing the exposure of interest given the lack of schedule-specific data will be carefully evaluated and, if necessary, other sources of occupational details that could be appended to this data set will be considered. Additionally, the potential for bias exists due to the fact that the exposure (LWH) and outcomes of interest (SRGH status, CHD, and mental health status) are self-reported; as a result, the data will be evaluated for common method variance to determine whether it is systematic or random and what procedures will best minimize the presence of such bias.[157] Some pertinent PSID questions and response choices have been added, removed, or altered over the 25 years of data to be analyzed for this project. Where possible, shifting questions or response choices will be recoding for consistency, which may increase the length of time necessary to complete these analyses. Moreover, the formatting of the data across and within years is also expected to require a significant investment of time in the early stages of the project.

However, the longitudinal nature of this study, its panel design, and the vast number of variables collected on a large, population-based sample over the study period are strengths of this study. The PSID is comprised of more than 40 consecutive years of occupational experience data from a nationally representative, U.S. sample, which may allow for the generalization of results from this project to the wider U.S. working population, given that participants are not selected based on work status. The employment and health statuses of this cohort over the life course provide valuable evidence that can be generalized to the U.S. working population, and the application of PSID sampling weights allows for the calculation of unbiased individual estimates. Although secondary data analysis studies often have lower validity based on limited access to and inadequate documentation of data, this is not an issue with the PSID because its data files, survey instruments, and codebooks are complete and accessible to the public.[118]

Human Subjects Considerations

No contact with study participants will occur. The dataset to be analyzed will be comprised of data collected by the Survey Research Center (SRC) at the University of Michigan. All identifying variables will be removed by the SRC, and only de-identified data will be provided to the researcher. All datasets will be housed on University of Texas School of Public Health servers with secure firewalls and accessed using secure passwords. The only potential risk to subjects is the possibility of a subject's confidentiality being violated, which is minimal. PSID has been granted institutional approval by the University of Michigan. This project has been granted exempt status by the University of Texas Health Science Center Institutional Review Board (study #HSC-SPH-13-0871, reference #103130 and #116765) (Appendix A). Per U.S. federal regulations (§45 CFR 46.115(b)), data will be retained for a minimum of three years following the conclusion of the study. After that time, the data will be destroyed if they are no longer being utilized for research and once it is determined that there is no reasonable possibility of receiving requests that the data be made available to other institutions or researchers.

CHAPTER III: THE IDENTIFICATION OF A LONG WORK HOURS CUT POINT FOR PREDICTING ELEVATED RISK OF POOR SELF-REPORTED GENERAL HEALTH

Name of Journal Proposed for Article Submission: Epidemiology

Abstract

Background. A growing body of evidence has demonstrated that long work hours (LWH) is associated with numerous adverse health outcomes. The existing definitions of LWH, however, are heterogeneous and not health risk-based, and no previous study has identified the LWH cut point that best predicts the risk of poor health. The objective of this study was to derive a statistically optimized LWH cut point associated with poor self-reported general health (SRGH) status.

Methods. A retrospective analysis of 25 years (1986-2011) of repeated work hour measures on a representative sample of 2,206 U.S. workers from the Panel Study of Income Dynamics. A comprehensive set of potential work hour cut points from 36 to 65 work hours per week (WH/w) over a minimum of 10 years was tested to examine the effect of each specific WH/w exposure threshold on SRGH. A statistically optimized LWH cut point was identified through a series of univariate tests of model fit, calibration, and discrimination.

Results. Our analyses indicate that the LWH threshold for the increased risk of poor SRGH in U.S. workers is an average of ≥52 WH/w over a minimum of 10 years. The dichotomized exposure level of 52 WH/w or greater, on average, over a minimum of 10 years was identified as the optimized cut point in terms of model fit, calibration, and discrimination. Approximately 13 percent of participants were classified as exposed to 52 WH/w, and they had a higher risk of poor SRGH (RR = 1.28; 95% CI = 1.06−1.53) than participants exposed to lower LWH levels.

Conclusion. To our knowledge, this is the first study to identify a statistically optimized LWH threshold beyond which the risk of poor general health increases. This provides a health-risk based alternative to past and current arbitrarily defined definitions of LWH. Further research is necessary to determine the utility and predictive power of this LWH on other health outcomes.

Background

A growing body of evidence has demonstrated that long work hours (LWH) can have deleterious consequences on an individual's physical and psychological well-being. Positive associations have been found between LWH and a variety of adverse health outcomes, including poor general health,[1] cardiovascular disease,[2] musculoskeletal disorders,[3] work-related injuries,[4,5] depression,[6] and sleep disruption.[7] However, the literature is far from comprehensive, and the more limited the evidence on the relationship between LWH and a specific health outcome, the more likely those results are inconsistent or contradictory.

Variations in the definition, operationalization, and manner of assessment of LWH are routinely identified as the primary factors behind the conflicting epidemiological evidence.[8-10] For example, some studies have defined LWH in terms of overtime[11,12] or extended shift hours[13] while others examined daily[14] or weekly[15] work hours. Subdivisions of exposure categories in terms of hours worked – representing lesser or greater levels of exposure – have varied similarly.[2,9,16] Further, most studies of the LWH−health relationship have focused on specific occupational groups or workplace settings rather than analyzing representative national databases[13,17-19]; those analyses that have employed representative samples typically have been performed on data from countries other than the U.S.[2,20,21]

In response, several scholars have proposed that a single definition of LWH be applied universally based on the convention of 40 hours of work per week or 8 hours of work per day.[9] The issue with such suggestions, however, is that the recommended cut-points are not selected to accurately distinguish risk categories for the effects of LWH. Instead, the proposed cut-points are defined based on labor practices, such as legally-defined or generally-accepted work week durations,[22] or other subjective selection criteria, such as by constructing exposure categories by adding eight- or 10-hour increments,[23,24] rather than based on evidence of health risk.

Based on a large, population-based sample of U.S. households, the purpose of this research was to evaluate and identify the statistically optimized LWH exposure threshold that best predicts future poor self-reported general health (SRGH).[25,26] Measures of SRGH provide a single, composite health score reflecting an individual's perception of his or her collective well-being and morbidity status, which is based on a combination of subjective and objective biological, psychological, and social aspects of health. [27, 28] Because working long hours may have a greater impact on some aspects of health than others, SRGH is an appropriate measure for evaluating a broad range of physiological outcomes that may be associated with the exposure through the use of a composite measure.

Methods

Study Population

We utilized data from the Panel Study of Income Dynamics (PSID), which is a representative national household panel survey comprised of more than 22,000 individuals residing in 9,000 households.[29] The PSID was performed annually from its inception in 1968 until 1997 and biennially since 1997 to present. Response rates have varied between 90.7 and 98.0 percent since 1970.[30] Data on participant work hours have been collected since 1968; questions about general health status have been included since 1984.

For the current study, we restricted the cohort to adult participants (i.e., aged 18 years or older) reporting non-zero work hour data for a minimum of 10 years between 1986 and 2011 (n = 3,596). We excluded participants lacking work hour data for the years prior to a report of poor SRGH (n = 27), participants with missing SRGH responses in more than five waves (n = 22), and those reporting poor SRGH at baseline (n = 300). Further, given that LWH refers to daily or weekly work durations exceeding a threshold generally defined in terms of the standard work hours of a full-time employee,[9] we excluded 1,041 participants whose average work hours over the study duration were less than full-time employment equivalent (i.e., average of at least 35 hours of work per week).[31,32] The final sample for our analyses included 2,206 workers.

Data Collection

In each interview year, participants were asked whether they had "any extra jobs in addition to your main job" and "on average, how many hours per week did you work" for up to five jobs in the previous year. If respondents indicated in a follow-up question that their reported work hours per week did not include actual overtime worked, they were asked "how many hours did that overtime amount to in the previous year". Participants were asked how many weeks out of the previous year they had worked on each job as well as how many weeks they had not worked, including time missed due to illness, vacation, strike, unemployment, or layoff. Respondents were also asked to provide the start and stop dates (month and year) for up to four employers for whom they worked in the previous year.

From these data, a composite measure of total annual hours spent working was created for each of the 19 waves comprising this study duration, which was calculated from the total number of weeks actually worked on each job multiplied by the average weekly work hours per job plus the average weekly overtime hours, if applicable, all in the previous year. For each study wave, we calculated the average number of hours worked per week in the previous year on all jobs by dividing the total annual hours spent working by 50 weeks of work. We then calculated the mean number of hours worked per week on all jobs across the study duration to produce a summary variable of average weekly work hours over a minimum of 10 years — and up to 25 years — of work time. Because our goal for this analysis was to identify a threshold of exposure to long-term work hours durations that increased participants' risk of poor SRGH, analyzing the arithmetic mean of the annual average work hours per week (WH/w) over the study duration was more appropriate than analyzing the sum of time at risk.

In each interview year, participants were asked to report if their health, in general, was "excellent, very good, good, fair, or poor". SRGH is used widely as an index of the overall health of an individual and has been shown to be a better predictor of mortality than medical diagnostic criteria.[27,28,33] For this analysis, we constructed a dichotomized SRGH variable, with good SRGH defined as participant responses of "excellent", "very good", and "good", and poor SRGH defined as responses of "fair" or "poor". Incident cases of poor SRGH were defined as first reports of poor SRGH following consistent reports of good SRGH.

Statistical Analysis

A comprehensive set of potential work hour cut points from 36 to 65 WH/w (in one-hour increments) was compiled in order to test the effect of each specific exposure threshold on SRGH. Small cell sizes precluded the analysis of cut points above 65 WH/w. The definition of LWH differed as each cut point was tested, and participants were classified as exposed to LWH if their mean work hours per week across the study duration was equal to or greater than the dichotomized cut point being assessed. The relative frequencies of participants reporting poor SRGH at each cut point were determined. Since participants reporting poor SRGH were likely to have altered their work hour patterns in the last year as a result of their poor health, participants were censored after they reported poor SRGH.

To distinguish the threshold of average work hours at which SRGH status declined after long-term exposure, we developed a series of univariate predictive models to identify a statistically optimized LWH cut point, which required evaluating the accuracy of the set of candidate cut point using calibration and discrimination measures. The accuracy of predictive models must be evaluated in terms of risk estimation (or, calibration, that is, the predictive power of the exposure on the outcome[26,34] with measures such as the risk or rate ratios; or the agreement between predicted probabilities and actual observed risk[35] including measures such as the Somers' *D* statistics) as well as risk classification (or, discrimination, that is, the model's ability to classify the disease state and include such statistics as the Youden Index, the receiver operating characteristic (ROC) curve (or, AUC), sensitivity, specificity, and likelihood ratios[26]). The accuracy of diagnostic models often relies on the AUC to assess the ability to discriminate patients from non-patients, but the AUC has limited utility in assessing predictive models.[35] In such models, calibration is an important component in evaluating risk factors, which may have a minimal impact on the AUC yet be important predictors of risk.[35,36]

To measure model fit for each exposure level cut point, we calculated the Bayesian Information Criteria (BIC), which allows the comparison of non-nested models, with lower BIC values indicate better fitting models.[37] As measures of calibration we used (1) Somers' *D* statistics to determine the predictive power of the dichotomized exposure levels on the binary outcome by calculating the difference in the proportions of outcome events in the exposed and unexposed populations [34]; and, (2) univariate relative risks (RRs) from Poisson regression to determine the probability of the outcome of interest given the exposure of interest (i.e., at or above a work hour threshold).[35] Because not all participants had equal exposure durations, an offset variable was constructed by dividing the number of waves of work hour data reported by the number of possible waves. For all *D* statistics and RRs, one-sided 95% confidence intervals (CIs) were calculated. One-sided analyses were performed because the study focused on possible increases in the risks of adverse health outcomes associated with LWH; a protective effect of LWH was not considered in this study and is not generally supported by the literature.[10] Although the Hosmer-Lemeshow goodness of fit test is often used as a measure of calibration and has an extension for use with survey data, it is not applicable to models fitted using Poisson regression.[38]

As measures of discrimination, we used (1) the AUC; (2) sensitivity (s_n) and specificity (s_p); and, (3) the Youden Index (J), which is a function of sensitivity and specificity and represents the maximum vertical distance between a ROC curve and the line of no discrimination.[39] When comparing cut points, the candidate corresponding to the largest magnitude of J is generally interpreted as the point on the curve farthest from chance and, thus, defined as the statistically optimal cut point.[25,39] Calculating the J statistic for a given cut point produces a value ranging from 0 to 1, with 0 indicating that the cut point has no ability to differentiate between the diseased and the non-diseased and 1 indicating that no misclassification errors are present.[40] Additional measures of discrimination included (4) positive and negative predictive values (PPV and NPV, respectively), which indicated the probability of disease given that participants have been identified as exposed or unexposed[41]; and (5) likelihood ratios (LR), both positive (LR_{pos}) and negative (LR_{neg}), which ranged from 0 to infinity and represented the probability of the risk factor of interest in participants with the outcome of interest versus the probability of the risk factor in participants without the outcome.[42] Like the *J* statistic, LRs may be used to select an optimal cut points from among several candidates; when comparing cut points, the candidate that maximizes LR_{pos} while minimizing LR_{neg} is determined to be the optimal cut point.[41,43]

To compensate for restricting the cut points in these analyses to natural numbers, a final analysis was performed treating WH/w as a continuous variable instead of dichotomizing it against a specific cut point. Using the CUTPT module in Stata with the Youden method selected, a statistically optimal cut point was generated for the continuous WH/w exposure variable versus the binary SRGH status outcome variable. We had reasoned that a whole number cut point would be most useful for evaluating exposures to LWH in work settings. Although this step provided additional information on the best fitting cut point given these data, it was not intended to be a determining factor in identifying the statistically optimized cut point.

Data were analyzed using Stata/SE software (version 13.1; Stata Corp, College Station, TX), taking into account clustering and probability weighting of the PSID data.

Results

The left side of Table 1 provides information on the model fit and calibration of LWH exposure at various cut points and the outcome of poor SRGH status. The *D* tests produced a statistically significant coefficient for the working hours cut point of 52 hours per week; no other cut point produced a statistically significant *D* coefficient. Univariate Poisson regression yielded analogous results, with the only statistically significant RR generated at 52 WH/w (RR: 1.27; 95% CI: 1.06 – 1.53); this cut point was shown to have adequate statistical power to detect these results (power: 0.958).

The right side of Table 1 presents the discriminatory ability of the cut point models. We observed that the *J* statistic was greatest at 52 WH/w, and the AUC was maximized at 51 and 52 WH/w. The 52 WH/w cut point simultaneously maximized the value of LR_{pos} and minimized the value of LR_{neg} as well as maximizing PPV and NPV. No single exposure level maximized sensitivity and specificity; sensitivity decreased and specificity increased as the work hour values increased.

Based on the results of the model fit, calibration, and discrimination analyses, the data suggest the identification of 52 WH/w as the LWH cut point that most accurately predicts an increase in the risk of poor SRGH status after long-term exposure. Participants working at least 52 hours per week, on average, for a minimum of 10 years had an approximately 28 percent increased risk of reporting poor SRGH than those who averaged fewer than 52 work hours per week over the same duration. The selection of 52 WH/w as the statistically optimized cut point is supported by the analysis of the exposure as a continuous variable, which identified the optimal cut point as 51.8 WH/w (J = 0.041; s_n = 16%; s_p = 88%; AUC = 0.52).

Following the selection of the statistically optimized cut point, descriptive data on the participants were calculated by work hour duration (Table 2). Participants in our study who reported working an average of 52 WH/w or more for a minimum of 10 years tended to be older, more likely to be male, and have attained higher levels of education than individuals working fewer hours. Race differed by LWH exposure status with fewer minorities reporting LWH exposure, but ethnicity (defined in terms of Hispanic descent) did not vary by LWH exposure status. Those exposed to LWH were more likely to be married, self-employed, and paid through non-traditional arrangements than unexposed participants. Occupation, industry, and the number of children residing in the household were not associated with LWH exposure.

Discussion

To our knowledge, this is the first study to have identified a statistically optimized threshold beyond which the risk of reporting poor general health increases, which is 52 hours of work per week or more, on average, for a duration of at least 10 years. Overall, this 52 WH/w cut point is the more accurate cut point in terms of calibration than discrimination compared to other cut points, lower or higher, showing good model calibration but limited model discrimination.

Models typically cannot maximize both calibration and discrimination measures, as increasing one construct generally decreases the other.[26] Previous studies have demonstrated that a model in which the predicted risk equals the observed risk for all subgroups, which would be considered perfectly calibrated, achieved maximum discriminatory power of an AUC equal to 0.83 and could not achieve an AUC of 1 under typical circumstances.[45] It has been argued that perfect calibration and perfect discrimination can be achieved only if the true risk as well as the estimated risk were either 0 or 1.[35] Although very large RRs have been shown to influence the AUC, more moderate ones have a limited effect.[26] This is because the AUC is strongly influenced by the degree of overlap between the distributions for cases and non-cases.[35] It has been proposed that an RR of approximately 16 per 2 standard deviations would be required to achieve adequate discrimination between cases and non-cases, which corresponded to an estimated AUC of 0.84.[36] Studies have shown that RRs of less than or equal to 3.0 per 2 standard deviations have not resulted in a substantial increase in the area under the ROC curve, which suggests that the RRs produced by this research would have little influence on the corresponding AUC.[35] Further, it has been demonstrated that a single factor – such as LWH – rarely achieves adequate levels of discrimination, which would require a substantial difference in the means of cases and non-cases.[36] This is evident in the relatively small magnitudes of *J* and the AUC, which indicate that the 52 WH/w cut point has poor distinguishing ability.[40]

Our results are consistent with the hypothesis that a LWH risk threshold exists beyond which increases in the risk of poor health can be measured. Based on his review of the literature, Harrington (1994) proposed that increased risk might occur from working more than 48 to 56 hours per week.[46] A similar conclusion was reached in a meta-analysis by Sparks *et al.* (1997), whose evidence indicated that exceeding a threshold of 48 hours of work per week might result in the onset of adverse health symptoms.[47] Spurgeon and colleagues (1997) found that weekly work hours exceeding 50 WH/w were associated with an increase in occupational stress but ultimately concluded that existing data were not sufficient for determining a LWH threshold of effect.[48] More recently, Caruso and colleagues (2006) argued that flawed methodological approaches had precluded the identification of a valid threshold for safe work hours.[10]

Our findings also support the hypothesis that LWH may be a risk factor for poor SRGH. Overall, we found that working an average of 52 hours per week or greater for more than 10 years was associated with an increase in the risk of poor SRGH of approximately 28 percent. Several previous studies have also reported a positive association between LWH and poor SRGH,[1,14,19,22,49] although the finding is not universal.[23,50] In the largest study to date of the LWH–poor SRGH relationship (n = 15,583), which was performed in a representative sample of European workers, results were significant for certain subgroups of European males working 40 to 60 WH/w, who were 1.4 to 9.9 times more likely to report poor SRGH than males who worked fewer hours. Similar results were found for certain subgroups of European females, who were 2.2 to 3.4 times more likely to report poor SRGH than females who worked fewer hours.[49]

However, all previous studies of LWH and SRGH status employed multivariate models. We reasoned that a univariate examination of the predictive ability of dichotomized work hour cut points on poor SRGH would yield the cut point that was the strongest prognostic indicator of the outcome of interest across all members of our cohort. We considered this analysis a first step towards building a multivariate model that categorized work hour exposure based on health risk. An additional objective was to produce a cut point that could be applied to U.S. workers across various socio-demographic characteristics, industries, and occupations, which required the omission of covariates.

A limitation of our study is that the PSID data were not collected specifically to address the relationship between work hours and health. Although PSID participants comprise a population-based, nationally-representative U.S. sample, the PSID cohort is not specifically sampled to produce a representative sample of U.S. workers and its work hours and health data are entirely self-reported. However, the application of PSID sampling weights allows for the calculation of unbiased individual estimates, which increases generalizability to the U.S. working population. Validation studies have demonstrated a strong positive correlation (Pearson's r = 0.64) between cross-sectional PSID self-reported total hours worked in the previous year and weekly work hours reported by employers.[51] Validation studies of SRGH have consistently shown it to be a reliable and valid measure of general health among adults, and it is highly correlated with objective measures of physical and mental health.[52]

In conclusion, this study provides an alternative to the current definitions of LWH, which typically characterize LWH in terms of one or more arbitrarily defined cut points or categories. We derived a single cut point to define LWH relative to poor SRGH but further analyses may identify alternative statistically optimized LWH cut points for other health outcomes. Also, further research may explore whether or not a range of work hour cut points may be appropriate for identifying a gradient of at-risk individuals.

Acknowledgements

Partial funding for this project was received from an Occupational Epidemiology Traineeship and a Pilot Projects Research Training Grant, both from the National Institute for Occupational Safety and Health (NIOSH)/ Centers for Disease Control and Prevention to the University of Texas School of Public Health Southwest Center for Occupational and Environmental Health, a NIOSH Education and Research Center (Grant No. 5T42OH008421).

Table 1. Measures[a] of model fit, calibration and model discrimination for long work hours and self-reported general health: Panel Study of Income Dynamics, 1986-2011.

		Measures of model fit and calibration				Model discrimination							
Cut points (weekly work hours)	**Proportion above cut point (%)**	**Bayesian Information Criteria**	**RR [b]**	**(95% CIs) [b]**	**Somers' *D* Statistics**	**Youden Index (*J*)**	**Sensitivity (%)**	**Specificity (%)**	**AUC (*c*) [c]**	**LR [d] (+)**	**LR [d] (-)**	**% PPV [e]**	**% NPV[f]**
36	95.5	5.47E+07	0.87	(0.64 - 1.19)	-0.0062	0.000	95.0	4.3	0.50	0.99	1.15	34.6	62.0
37	88.4	5.46E+07	0.77	(0.63 - 0.94)	-0.0277	0.000	87.5	11.1	0.50	0.98	1.13	34.3	62.5
38	81.5	5.47E+07	0.88	(0.74 - 1.05)	-0.0138	0.003	81.7	18.6	0.50	1.00	0.98	34.8	65.7
39	74.0	5.47E+07	0.90	(0.77 - 1.06)	-0.0134	0.000	73.1	25.5	0.50	0.98	1.05	34.3	64.1
40	65.2	5.47E+07	0.91	(0.78 - 1.05)	-0.0141	0.000	63.8	34.1	0.50	0.97	1.06	34.0	63.9
41	57.1	5.46E+07	0.85	(0.74 - 0.99)	-0.0333	0.000	54.7	41.6	0.50	0.94	1.09	33.2	63.3
42	50.3	5.46E+07	0.87	(0.75 - 1.01)	-0.0263	0.000	47.8	48.4	0.50	0.93	1.08	33.0	63.6
43	44.8	5.47E+07	0.95	(0.82 - 1.10)	0.0006	0.000	43.9	54.8	0.50	0.97	1.03	34.0	64.7
44	40.0	5.47E+07	0.94	(0.81 - 1.08)	-0.0054	0.000	39.3	59.6	0.50	0.97	1.02	34.1	64.9
45	35.5	5.47E+07	0.92	(0.79 - 1.07)	-0.0103	0.000	34.3	63.9	0.50	0.95	1.03	33.6	64.7
46	31.3	5.47E+07	0.90	(0.78 - 1.05)	-0.0148	0.000	30.3	68.2	0.50	0.95	1.02	33.6	64.8
47	27.7	5.47E+07	0.94	(0.81 - 1.09)	-0.0052	0.006	28.1	72.5	0.50	1.02	0.99	35.2	65.5
48	24.0	5.47E+07	1.00	(0.86 - 1.17)	0.0097	0.022	25.5	76.8	0.51	1.10	0.97	36.8	66.0
49	21.0	5.47E+07	1.01	(0.86 - 1.19)	0.0093	0.026	22.7	79.9	0.51	1.13	0.97	37.5	66.0
50	17.7	5.47E+07	1.05	(0.88 - 1.25)	0.0135	0.029	19.6	83.3	0.51	1.17	0.97	38.4	66.1
51	15.0	5.46E+07	1.12	(0.94 - 1.34)	0.0223	0.032	17.1	86.1	0.52	1.23	0.96	39.6	66.2
52	**12.9**	**5.45E+07**	**1.28**	**(1.06 - 1.53)**	**0.0400***	**0.040**	**15.5**	**88.5**	**0.52**	**1.35**	**0.95**	**41.8**	**66.3**
53	10.9	5.46E+07	1.16	(0.95 - 1.42)	0.0212	0.023	12.4	89.9	0.51	1.22	0.97	39.4	65.9
54	9.5	5.46E+07	1.21	(0.98 - 1.51)	0.0251	0.021	10.8	91.3	0.51	1.24	0.98	39.7	65.8
55	7.8	5.46E+07	1.20	(0.95 - 1.52)	0.0197	0.017	8.9	92.9	0.51	1.24	0.98	39.8	65.7
56	6.8	5.46E+07	1.22	(0.97 - 1.54)	0.0200	0.021	8.1	94.0	0.51	1.34	0.98	41.6	65.8
57	5.8	5.46E+07	1.21	(0.95 - 1.56)	0.0170	0.017	6.9	94.8	0.51	1.33	0.98	41.4	65.7
58	4.6	5.46E+07	1.22	(0.92 - 1.61)	0.0135	0.013	5.5	95.8	0.51	1.32	0.99	41.2	65.6
59	3.9	5.47E+07	1.15	(0.84 - 1.57)	0.0081	0.008	4.4	96.4	0.50	1.23	0.99	39.5	65.5
60	3.4	5.47E+07	1.02	(0.71 - 1.47)	0.0019	0.004	3.7	96.7	0.50	1.12	1.00	37.3	65.4
61	3.0	5.47E+07	1.05	(0.71 - 1.54)	0.0024	0.003	3.1	97.2	0.50	1.10	1.00	36.9	65.4
62	2.4	5.47E+07	1.13	(0.75 - 1.71)	0.0049	0.000	2.3	97.6	0.50	1.00	1.00	34.6	65.3
63	2.0	5.47E+07	0.98	(0.60 - 1.60)	-0.0003	0.000	1.7	97.8	0.50	0.79	1.00	29.5	65.2
64	1.8	5.47E+07	0.89	(0.52 - 1.51)	-0.0023	0.000	1.4	98.0	0.50	0.71	1.01	27.5	65.2
65	1.5	5.47E+07	1.02	(0.60 - 1.74)	0.0009	0.000	1.3	98.3	0.50	0.78	1.00	29.4	65.2

[a] All calculations take into account clustering and probability weighting; [b] Rate ratio and corresponding 95% Confidence Interval from univariate Poisson regression. [c] Area under the receiver operating characteristic (ROC) curve; [d] Likelihood Ratio; [e] Positive Predictive Value; [f] Negative Predictive Value.

* P value < 0.05; ** P value < 0.01.

References

1. Ettner S, Grzywacz J. Workers' perceptions of how job affect health: A social ecological perspective. *J Occ Health Psychol*. 2001;6:101-113.

2. Virtanen M, Heikkila K, Jokela M, *et al*. Long working hours and coronary heart disease: A systematic review and meta-analysis. *Am J Epidemiol*. 2012;176(7):586-596.

3. Lipscomb J, Trinkoff A, Geiger-Brown J, Brady B. Work-schedule characteristics and reported musculoskeletal disorders of registered nurses. *Scand J Work Environ Health*. 2002;28:394-401.

4. Simpson C, Severson R. Risk of injury in African American hospital workers. *J Occ Env Med*. 2000;42:1035-1040.

5. Dembe A, Erickson J, Delbos R, Banks S. The impact of overtime and long work hours on occupational injuries and illnesses: New evidence from the United States. *Occ Env Med*. 2005;62:588-597.

6. Hilton M, Whiteford H, Sheridan J, *et al*. The prevalence of psychological distress in employees and associated occupational risk factors. *J Occup Environ Med*. 2008;50:746-757.

7. Kivisto M, Harma M, Sallinen M, Kalimo R. Work-related factors, sleep debt and

insomnia in IT professionals. *Occup Med (Lond)*. 2008;58:138-140.

8. van der Hulst M. Long workhours and health. *Scand J Work Environ Health*. 2003;29(3):171-188.

9. Bannai A, Tamakoshi A. The association between long working hours and health: A systematic review of epidemiological evidence. *Scand J Work Environ Health*. 2014:40(1):5-18.

10. Caruso C, Bushnell T, Eggerth D, *et al*. Long working hours, safety, and health: Toward a national research agenda. *Am J Ind Med*. 2006;49:930-942.

11. Theorell T, Rahe R. Behavior and life satisfaction characteristics of Swedish subjects with myocardial infarction. *J Chronic Dis*. 1972;25(3):139-147.

12. Kawakami N, Akachi K, Shimizu H, *et al*. Job strain, social support in the workplace, and haemoglobin A1c in Japanese men. *Occup Environ Med*. 2000;57:805-809.

13. Proctor S, White R, Robins T, Echeverria D, Rocskay A. Effect of overtime work on cognitive function in automotive workers. *Scand J Work Environ Health*. 1996;22:124-132.

14. Jex S, Bliese P. Efficacy beliefs as a moderator of the impact of work-related stressors: a multilevel study. *J Appl Psychol*. 1999;84:349-361.

15. Kroenke C, Spiegelman D, Manson J, Schernhammer E, Colditz G, Kawachi I. Work characteristics and incidence of type 2 diabetes in women. *Am J Epidemiol*. 2007;165:175-183.

16. van der Hulst M. Long workhours and health. *Scand J Work Environ Health*. 2003;29:171-188.

17. Borg V, Kristnsen T. Psychosocial work environment and mental health among travelling salespeople. *Work Stress*. 1999;13:132-143.

18. Nakanishi N, Nishina H, Yoshida H, *et al*. Hours of work and the risk of developing impaired fasting glucose or type 2 diabetes mellitus in Japanese male office workers. *Occup Environ Med*. 2001;58:569-74.

19. Baldwin P, Wrate R. Young doctors' health -- I. How do working conditions affect attitutdes, health and performance. *Soc Sci Med*. 1997;45(1):35-40.

20. Tarumi K, Hagihara A, Morimoto K. A prospective observation of onsets of health defects associated with working hours. *Ind Health*. 2003;41(2):101-108.

21. Holtermann A, Mortensen O, Burr H, *et al*. Long work hours and physical fitness: 30-year risk of ischaemic heart disease and all-cause mortality among middle-aged Caucasian men. *Heart*. 2010;96(20):1638-1644.

22. Song JT, Lee G, Kwon J, Park JW, Choi H, Lim S. The association between long working hours and self-rated health. *Ann Occ Environ Med*. 2014;26(2):1-12.

23. Artazcoz L, Cortes I, Borrell C, Escriba-Aguir V, Cascant L. Gender perspective in the analysis of the relationship between long workhours, health and health-realted behavior. *Scand J Work Environ Health*. 2007;33(5):344-350.

24. Ribet C, Derriennic F. Age, working conditions, and sleep disorders: a longitudinal analysis in the French cohort E.S.T.E.V. *Sleep*. 1999;22(4):491-504.

25. Schisterman E, Perkins N, Liu A, Bondell H. Optimal cut-point and its corresponding Youden Index to discriminate individuals using pooled blood samples. *Epidemiol*. 2005;16:73-81.

26. Cook N. Statistical evaluation of prognostic versus diagnostic models: Beyond the ROC curve. *Clinical Chemistry*. 2008;54(1):17-23.

27. Kuhn R, Rahman O, Menken J. Survey Measures of Health: How Well Do Self-Reported and Observed Indicators Measure Health and Predict Mortality? In: Cohen B, Menken J, eds. *Aging in Sub-Saharan Africa: Recommendation for Furthering Research*. Wasghington, DC: National Academies Press; 2006.

28. Jylha M. What is self-rated health and why does it predict mortality? Towards a unified

conceptual model. *Soc Sci Med*. 2009;69:307-316.

29. University of Michigan Institute for Social Research. PSID: Data on Families Since 1968. http://psidonline.isr.umich.edu/Guide/Brochures/PSID.pdf. Accessed Apr 11, 2014.

30. McGonagle K, Schloeni R, Sastry N, Freedman V. The Panel Study of Income Dynmaics: Overview, recent innovations, and potential for life course research. *Longi Life Course Stud*. 2012;3(2):268-284.

31. Ricci J, Chee E, Lorandeau M, Berger J. Fatigue in the U.S. workeforce: Prevalence and implications for lost productive work time. *JOEM*. 2007;49(1):1-10.

32. Bureau of Labor Statistics, United States Department of Labor Division of Labor Force Statistics. Labor Force Statistics from the Current Population Survey. Dec 08, 2014. Available at: http://www.bls.gov/cps/lfcharacteristics.htm#fullpart. Accessed Jan 05, 2015.

33. Idler E, Benyamini Y. Self-rated health and mortality: A review of twenty-seven community studies. *J Health Soc Behav*. 1997;38:21-37.

34. Newson R. Parameters behind "nonparametric" statistics: Kendall's tau, Somers' D and median differences. *The Stata Journal*. 2002;2(1):45-64.

35. Cook N. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*. 2007;115:928-935.

36. Pepe M, Janes J, Longton G, Leisearing W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol*. 2004;159:882-890.

37. Hirose K, Kawano S, Konishi S, Ichikawa M. Bayesian information criterion and selection of the number of factors in factor analysis models. *J Data Sci*. 2011;9:243-259.

38. Archer K, Lemeshow S. Goodness-of-fit test for a logistic regression model fitted using survey sample data. *Stata J*. 2006;6(1):97-105.

39. Perkins N, Schisterman E. The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve. *Am J Epidemiol*. 2006;163:670-675.

40. Youden W. Index for rating diagnostic tests. *Cancer*. 1950;3(1):32-35.

41. Feinstein A. Clincal Epidemiology: The Architecture of Clinical Research. Philadelphia: WB Saunders; 1987.

42. McGee S. Simplifying likelihood ratios. *J Gen Intern Med*. 2002;17(8):647-650.

43. Radack K, Rouan G, Hedges J. The likelihood ratio. *Arch Parhol Lab Med*. 1986;110:689-693.

44. Higgins PB, Gower BA, Hunger GR, Goran MI. Defining helath-related obesity in prepubertal children. *Obesity Research*. 2001;9(4):233-240.

45. Diamond G. What price perfeection? Calibration and discrimination of clinical prediction models. *J Clin Epidemiol*. 1992;45:85-89.

46. Harrington J. Shift work and health -- A critical review of the literature on working hours. *Annals of the Academy of Medicine, Singapore*. 1994;23:699-705.

47. Sparks K, Cooper C, Fried Y, Shirom A. The effects of hours of work on health: A meta-analytic review. *J Occup Organ Psychol*. 1997;70:391-408.

48. Spurgeon A, Harrington J, Cooper C. Health and safety problems associated with long working hours: A review of the current position. *Occup Environ Med*. 1997;54:367-375.

49. Artozcoz L, Cortes I, Escriba-Aguir V, Bartoll X, Basart H, Borrel C. Long working hours and health status among employees in Europe: Between-country differences. *Scan J Work Environ Health*. 2013;39(4):369-378.

50. Artazcoz L, Cortes I, Escriba-Agur V, Cascant L, Villegas R. Understanding the

relationship of long working hours with health status and health-related behaviors. *J Epidemiol Community Health*. 2009;63:521-527.

51. Bound J, Brown C, Duncan G, Rodgers W. *Measurment error in cross-sectional and longitudinal labor market surveys: Results from validation studies*. Cambridge, MA 1989. Working Paper No. 2884.

52. McGee D, Liao Y, Cao G, Cooper R. Self-reported health status and mortality in a multiethnic US cohort. *Am J Epidemiol*. 1999;149(1):41-46.

CHAPTER IV: APPLICATION OF A LONG WORK HOURS CUT POINT FOR PREDICTING ELEVATED RISKS OF CARDIOVASCULAR DISEASE AND CANCER

Name of Journal Proposed for Article Submission: Epidemiology

Abstract

Background. Existing studies of the influence of work hours on health overwhelmingly define work hours in terms of 40 hours per week, but there is no evidence 40 hours of work per week is a threshold of health risk. The purpose of this study was to apply a previously derived methodology to assess the predictive ability of a set of dichotomized work hour cut points on the outcomes of incident cardiovascular disease (CVD) and incident cancer.

Methods. A retrospective analysis of 25 years (1986-2011) of repeated work hour measures was performed on a representative sample from the Panel Study of Income Dynamics of U.S. workers aged 18 years or older at baseline. Two study cohorts were constructed to evaluate the LWH–CVD relationship (n=1,806) and the LWH–cancer relationship (n=2,196). A comprehensive set of potential work hour cut points from 36 to 65 work hours per week (WH/w) was assess through a series of univariate tests of model fit, calibration, and discrimination, and a statistically optimized cut point for long work hours was identified for each outcome.

Results. The dichotomized exposure cut point of 52 WH/w or greater, on average, over a minimum of 10 years was identified as the optimized cut point for both outcomes in terms of model fit, calibration, and discrimination. In the CVD analysis, approximately 12 percent of participants were classified as exposed to 52 WH/w or greater, and they had a higher risk of incident CVD (RR = 1.42; 95% CI = 1.24−1.63) than participants exposed to lower LWH levels. Similar results were found in the cancer analysis, in which approximately 11 percent of participants were classified as exposed to 52 WH/w or greater, with an increased risk of incident cancer (RR = 1.62; 95% CI = 1.22−2.17) compared to participants exposed to lower LWH levels.

Conclusions. LWH is related to increased risk of incident CVD and incident cancer independently of other risk factors. These findings suggest that work hour durations equal to or longer than 52 hours per week, on average, over at least 10 years adversely affect health.

Background

Many studies have investigated the relationship between long work hours (LWH) and Many studies have investigated the relationship between long work hours (LWH) and health outcomes as varied as all-cause mortality, [1] occupational injury, [2, 3] mental health conditions, [4] and cognitive function, [5] with inconsistent findings in terms of the direction and magnitude of association. [6] The potential impact of LWH on health has been highlighted by researchers who believe that methodological weakness contributes to the equivocal results. [6, 7] Specifically, variations in the definition, operationalization, and assessment of LWH have been employed, [6, 7] as a single, accepted approach does not currently exist. As a result, more stringent frameworks for the evaluation of the effects of LWH on health have been proposed,[7-10] and the consistent application of a validated threshold of effect for LWH has been called for by numerous researchers.[7,11-14]

Cardiovascular disease (CVD) is the leading cause of death worldwide, and has been characterized as having a substantial negative effect on quality of life in industrialized countries. [15] Several work-related exposures (e.g., work stress, high job demand, low job control) are known etiological factors that contribute to the development CVD,[16-18] and a recent meta-analysis reported that LWH was associated with an estimated 37 percent increase in the risk of CVD (95% CI: 1.11 − 1.70). [19]

Less is known about the LWH–cancer relationship, however. Although several studies have demonstrated an association between shift work and breast cancer, [20, 21] there is a paucity of research on the effect of long-term exposure to LWH on cancer. To our knowledge, this is the first study to assess the relationship between hours of work and cancer incidence in a sample of workers that was not restricted to women.

The majority of studies of the LWH-health relationship employs the duration of 40 work hours per week to demarcate exposure to LWH, based on the conventional definition of a full-time work week in many industrialized nations, including the U.S. However, the use of a labor convention is of questionable validity for evaluating the health risks associated with LWH, as there is no evidence that this method of attributing work hours risk accurately measures that risk. Limited research has suggested possible cut points or ranges of LWH that are believed to have a threshold effect on poor health.[12, 13] However, this prior body of work did not focus on identifying the specific, statistically optimized work hour threshold that best predicts a particular health outcome. As previously detailed in this dissertation, we sought to develop a methodology for identifying a statistically optimized threshold or cut-point of work hours associated with an increased risk of poor self-reported general health (SRGH). In those analyses, we found consistent evidence that an average of 52 hours of work per week over a minimum of 10 years increased the risk SRGH.

The purpose of this study is to apply the methodological approach utilized in our prior study to examine the associations between LWH with incident CVD and incident cancer. More specifically, we aimed to assess the predictive ability of a set of dichotomized work hour cut points on these two outcomes of interest. We also sought to compare the optimized cut points observed in our study of SRGH of 52 hours per week with those identified for CVD and cancer. We then described the demographic and occupational characteristics of workers who exceeded the identified LWH threshold associated with these two health outcomes.

Methods

Study Population

The Panel Study of Income Dynamics has been described in detail elsewhere.[22, 23] In brief, the PSID is an on-going longitudinal survey of a representative sample of U.S. households.[23] Initiated by the U.S. Census Bureau in 1966 as the Survey of Economic Opportunity (SEO), the PSID was acquired and expanded by the Survey Research Center (SRC) at the University of Michigan in 1968 and has been ongoing to date.[24] The sample size is currently more than 22,000 individuals in 9,000 families with response rates greater than 90 percent.[22, 23]

Two study cohorts were constructed to evaluate the LWH-CVD relationship and the LWH–cancer relationship [hereafter referred to as the CVD cohort and the cancer cohort, respectively]. Participants were included in these analyses if they were at least 18 years of age or older at study baseline (1986) and reported non-zero work hours for a minimum of 10 years between 1986 and 2011. Due to differences observed in the baseline health of workers by number of work hours, we restricted our study sample to include those who worked at least 35 hours per week over the study duration. For inclusion in the outcome-specific cohorts, participants also had to have a complete set of responses to the health questions pertaining to the outcome of interest for their time of observation in the study. Prevalent cases of CVD and cancer at the baseline health study (1999) were excluded for purposes of examining the risk or onset of the outcomes post-exposure.

A total of 3,596 of 60,214 PSID participants met our initial study inclusion criteria. The CVD cohort was comprised of 1,806 participants, with 777 cases of incident CVD (or, 43.0%), after excluding 276 participants due to missing CVD data, 658 prevalent cases of CVD, and 856 participants reporting average weekly work hours of less than 35 hours.

For the outcome of cancer, 2,196 participants met our criteria, of whom 263 (12.0%) reported incident cancer, after excluding 116 participants for missing information on cancer status, 205 prevalent cases of cancer, and 1,079 participants reporting average weekly work hours of less than 35 hours.

Data Collection

The work hour exposure variable was constructed from information provided by participants in each interview year on the number of jobs they worked; the number of hours worked per week, on average, per job; the amount of overtime worked per week, on average; the number of weeks they had not worked due to illness, vacation, strike, unemployment, or layoff; and the start and stop dates for up to four jobs, all for the previous year. Total annual hours spent working was calculated from this information for each wave. For each year in the study, the average number of hours worked per week in the previous year on all jobs was calculated by dividing the total annual hours spent working by 50 weeks of work, per the U.S. Bureau of Labor Statistics' method for calculating incidence rates.[25] Finally, we calculated the mean of the average number of hours worked per week across the study duration.

The outcome of CVD was constructed using interview data starting in 1999. Participants were asked to report if a doctor had "ever told you that you have or had" any of the following: heart attack, high blood pressure, hypertension, coronary heart disease, angina, congestive heart failure or stroke. A dichotomous CVD index variable was generated for each year based on participant responses, with affirmative answers to any CVD-related question defined as a positive CVD outcome in a given year. From these index variables, a dichotomous CVD multi-year summary variable was constructed to identify whether individuals had ever reported incident CVD. Incident CVD was defined as first reports of any of the CVD-related outcomes (heart attack occurrence; coronary heart disease, angina, or congestive heart failure diagnoses; hypertensive status; or stroke occurrence) following at least one negative response to the CVD multi-year summary variable. Validation studies have demonstrated strong agreement between most self-reported symptoms and events related to CVD and medical records.[26-28]

Cancer cases were defined by an affirmative response to the PSID survey question "has a doctor ever told you that you have or had cancer or a malignant tumor, excluding skin cancer," which was asked in each wave beginning in 1999. From these variables, a dichotomous cancer multi-year summary variable was constructed to identify whether individuals had ever reported incident cancer. Incident cancer was defined as first reports of cancer following at least one previous report of not having ever been diagnosed with cancer. Previous studies have demonstrated excellent correlation between self-reported cancer status and cancer registry data or medical records.[28, 29]

Statistical Analysis

A series of univariate tests of model fit, calibration, and discrimination were used to evaluate a comprehensive set of potential work hour cut points from 36 to 65 work hours per week (WH/w). Candidates were defined as exposed to LWH if the arithmetic mean of their hours worked per week across the study duration was equal to or greater than the cut point being tested for the outcome being examined.

Within the CVD cohort, participants reporting incident cases of CVD were administratively censored following their first report, and no subsequent CVD data were analyzed on those individuals. Work hour data for cases was censored in the year in which the incident case was reported as well as for all subsequent years because it seemed likely that individuals with incident CVD may have changed their work hour patterns in the last year due to their health status. Incident cancer cases were treated equivalently within the cancer cohort.

Model fit was assessed for each work hour cut point using the Bayesian Information Criteria (BIC), which was selected because it does not require that models being compared are also nested. Model calibration was evaluated using Somers' *D* statistics and relative risks (RRs) for each univariate cut point model. *D* statistics provided a measure of the performance of each LWH cut point as a predictor of the outcomes of interest.[30] Univariate RRs, which were calculated using Poisson regression, indicated the probability of the outcome of interest given the exposure of each LWH cut point.[31] To adjust for unequal exposure durations, an offset variable was constructed by dividing the number of waves of work hour data that were reported by the number of possible waves. One-sided 95%

confidence intervals (CIs) were calculated for all *D* statistics and RRs; one-sided analyses were performed because the study question focused on the potential effects of a positive association between LWH and adverse health outcomes. Model discrimination was assessed by calculating the Youden Index (*J*) for each LWH exposure cut point model, which is maximized at the work hours threshold that has optimized differentiating ability in terms of the outcome of interest, giving equal weight to sensitivity and specificity. The area under the Receiver Operating Characteristics (ROC) curve (AUC), sensitivity (s_n), specificity (s_p), positive and negative likelihood ratios (LR_{pos} and LR_{neg}, respectively) and positive and negative predictive values (PPV and NPV, respectively) were calculated for each LWH cut point, as well, to provide additional evidence of each cut point's discriminatory ability. One alternative to the *J* statistic for identifying statistically optimized cut points is to determine those that maximize the LR_{pos} and/ or minimize the LR_{neg}, and agreement between the likelihood ratios and the *J* statistic was noted.[32] Additionally, an analysis of the exposure as a continuous variable was performed to determine the statistically optimized cut point had it not been restricted to a natural number.

Finally, we calculated descriptive statistics for selected socio-demographic and occupational characteristics at baseline by work hour duration for a combined CVD-cancer cohort, which was generated by merging the two cohorts. These analyses were performed after the identification of the best predicting cut point for each of the outcomes, as the LWH threshold was not known *a priori*. To determine whether there were significant differences between participants who worked long hours − defined as working equal to or greater than the LWH threshold − versus those who worked less, the means of continuous covariates and the prevalences of categorical covariates were compared using adjusted Wald tests and χ^2 statistics, respectively.

All analyses were performed in both cohorts (e.g., the CVD cohort and the cancer cohort) using Stata/SE software (version 13.1; Stata Corp, College Station, TX), and all calculations accounted for clustering and probability weighting.

Results

On the left side of Table 1, measures of model fit and calibration for each of the LWH cut points relative to the outcome of incident CVD are presented. The BIC was minimized for the work hour cut points of 52 through 56 WH/w. The *D* statistic was maximized at 52 WH/w, with positive associations seen for exposure cut points from 49 to 62 WH/w. Univariate Poisson regression yielded similar results, with significant increases in the risk of incident CVD demonstrated from 50 to 63 WH/w.

On the right side of Table 1, measures of model discrimination for the LWH cut points relative to incident CVD are presented. The *J* statistic was maximized at 52 WH/w, and the AUC was maximized for all values between 51 and 54 WH/w except 53 WH/w. No single cut point maximized sensitivity and specificity; instead, sensitivity decreased and specificity increased as the cut point values increased. In this analysis, LR_{pos} and PPV were maximized at 56 WH/w; LR_{neg} was minimized at 51 through 54 WH/w, excluding 53 WH/w; and NPV was maximized at 36 WH/w.

The left side of Table 2 presents measures of model fit and calibration for each of the LWH exposure cut points versus incident cancer. Minimized BIC values were seen for the work hour cut points of 52, 54, and 56 WH/w. The cut point of 52 WH/w maximized the *D* statistic, and positive associations were shown for the cut points of 48 WH/w and 50 through 57 WH/w. RR calculations demonstrated significant increases in the risk of incident cancer for cut points from 50 to 57 WH/w.

The right side of Table 2 presents the measures of model discrimination for the LWH cut points relative to incident cancer. Maximization of the *J* statistic and the AUC occurred at 52 WH/w; the AUC value was equivalent to 52 WH/w at 51 and 54 WH/w. Analogous to the results in the CVD cohort, sensitivity decreased and specificity increased in the cancer cohort as work hours increased. Additionally, L_{pos} was maximized at 62 WH/w; L_{neg} was minimized and NPV was maximized at 52 WH/w, with equivalent NPV values at 51 and 54 WH/w; and PPV was maximized at 63 WH/w.

For both outcomes, the results of the model fit, calibration, and discrimination analyses suggest the identification of 52 WH/w as the LWH cut point that most accurately predicts an increase in the risk of incident CVD and incident cancer after long-term exposure. These conclusions are supported by the analyses of the exposure as a continuous variable, which identified the optimal LWH cut point relative to incident CVD as approximately 52.2 WH/w (J = 0.060; s_n = 15%; s_p = 91%; AUC = 0.53), and the optimal LWH cut point relative to incident cancer as approximately 52.1 WH/w (J = 0.058; s_n = 16%; s_p = 90%; AUC = 0.53). However, in both the incident CVD and the incident cancer analyses, small values of *J* and AUC were produced, which suggests that the univariate cut point models have poor distinguishing ability.[33, 34]

Descriptive data on the participants are presented in Table 3 by average work hours per week, with the ≥52 WH/w category representing the exposure of LWH. When combined, the CVD and cancer cohorts represented 2,332 participants, of whom 1,670 were included in both the CVD and the cancer cohorts. Because a substantial proportion of participants were shared across both cohorts, the socio-demographic variables in Table 3 were evaluated using the combined sample of 2,332 participants. Participants in the combined cohort who reported working 52 WH/w or more, on average, for a minimum of 10 years were more likely to be male, older, married, and have attained higher levels of education. Although race varied by LWH exposure status, with minorities less likely to report LWH, ethnicity (defined in terms of Hispanic descent) was not associated with LWH. Those exposed to LWH were more likely to be self-employed or paid through non-traditional pay structures (defined as not paid hourly or salaried) than those not reporting LWH. Participants reporting LWH were more likely to report incident CVD or incident cancer. Occupation, industry, and the number of children residing in the household were not associated with LWH exposure.

Discussion

We evaluated a range of LWH cut points to identify the LWH thresholds that most accurately predicted an increased risk in incident CVD or incident cancer. We observed that an average of 52 hours of work per week for a minimum of 10 years was the statistically optimized threshold beyond which the risk of incident CVD or incident cancer increased.

This conclusion was consistent with our prior analyses in which we examined LWH and SRGH.[35]

Our analyses also suggest that the use of LWH cut points that have not been optimized for the prediction of health risks may be contributing to the inconsistent results and generally weak measures of effect seen in other studies of the LWH−health relationship. Although the number of publications investigating the health effects of LWH is increasing, scientific rationale to support the use of a specific quantity of work hours as an indicator of health risk is lacking. To our knowledge, there have been no published evaluations of the predictive ability of a specific LWH cut point for different adverse health outcomes.

Studies of LWH often define work hours in terms of 40 hours per week, either as a dichotomized cut-point or as a category end-point in a multilevel exposure categorization. This is most likely due to the fact that 40 work hours per week is the generally accepted labor standard for full-time work in the U.S. and several other industrialized nations. Other cut points used to define LWH have tended to add hours in five- or ten-hour increments (e.g. Artazcoz *et al.* 2007 exposure categories of <30, 30-40, >40 work hours per week)[36] or in additional 8-hour durations (e.g. Ribet and Derriennic 1999 exposure categories of ≤48 and >48 hours per week).[37] It is often unclear whether such selections are arbitrary, rooted in convention, or based on scientific evidence.

The longitudinal nature of this study in a large, representative sample of U.S. households allowed us to identify the statistically optimized cut points for LWH using repeated measures of working hours. However, the PSID cohort is not specifically sampled to produce a representative sample of U.S. workers, and PSID data are not collected primarily to address the relationship between work hours and health. PSID sampling weights were applied in these analyses to produce unbiased individual estimates, which increases the generalizability of these findings across U.S. workers.

In the present study, we were able to address methodological issues identified in previous research and provide a foundation for future examinations of LWH through the use of this health risk-derived cut point. The reliance on a general industry standard − e.g., the 40 hour workweek – or other arbitrarily-defined LWH cut points or categories as a health risk threshold is problematic because they have been neither created nor validated for such a purpose. Further research using other representative data sets of U.S. workers and adverse health outcomes is necessary to assess the predictive value of the exposure cut point of 52 WH/w as an independent risk factor of LWH for detrimental health conditions.

Acknowledgements

Partial funding for this project was received from an Occupational Epidemiology Traineeship and a Pilot Projects Research Training Grant, both from the National Institute for Occupational Safety and Health (NIOSH)/ Centers for Disease Control and Prevention to the University of Texas School of Public Health Southwest Center for Occupational and Environmental Health, a NIOSH Education and Research Center (Grant No. 5T42OH008421).

Table 1. Measures[a] of model fit, calibration and model discrimination for long work hours and incident coronary heart disease: Panel Study of Income Dynamics, 1986-2011.

Cut points (weekly work hours)	Proportion above cut point (%)	Measures of model fit and calibration				Model discrimination							
		Bayesian Information Criteria	RR [b]	(95% CIs) [b]	Somers' *D* Statistics	Youden Index (*J*)	Sensitivity (%)	Specificity (%)	AUC (*c*) [c]	LR [d] (+)	LR [d] (-)	% PPV [e]	% NPV [f]
36	94.6	5.66e+07	1.30	(0.94 - 1.81)	0.0225	0.015	95.5	6.0	0.51	1.02	0.75	46.1	61.1
37	87.4	5.66e+07	1.13	(0.92 - 1.39)	0.0278	0.028	89.7	13.1	0.51	1.03	0.78	46.6	60.2
38	79.6	5.66e+07	1.14	(0.96 - 1.34)	0.0439	0.032	82.4	20.8	0.52	1.04	0.85	46.7	58.4
39	71.9	5.66e+07	1.14	(0.99 - 1.31)	0.0538	0.046	75.7	28.9	0.52	1.06	0.84	47.3	58.5
40	63.4	5.66e+07	1.14	(1.01 - 1.30)	0.0623	0.056	68.7	36.9	0.53	1.09	0.85	47.9	58.3
41	55.6	5.66e+07	1.06	(0.94 - 1.19)	0.0395	0.025	58.3	44.2	0.51	1.04	0.94	46.8	55.7
42	48.5	5.66e+07	1.07	(0.96 - 1.20)	0.0413	0.027	52.3	50.5	0.51	1.06	0.95	43.7	52.2
43	42.7	5.66e+07	1.09	(0.97 - 1.22)	0.0494	0.028	46.6	56.2	0.51	1.06	0.95	47.3	55.5
44	37.9	5.66e+07	1.04	(0.93 - 1.17)	0.0312	0.024	41.6	60.8	0.51	1.06	0.96	47.2	55.2
45	34.2	5.66e+07	1.02	(0.90 - 1.14)	0.0191	0.009	36.6	64.4	0.50	1.03	0.99	46.4	54.6
46	30.1	5.66e+07	1.01	(0.89 - 1.14)	0.0158	0.012	31.9	69.3	0.51	1.04	0.98	46.7	54.7
47	26.3	5.66e+07	1.03	(0.91 - 1.16)	0.0196	0.023	29.1	73.2	0.51	1.08	0.97	47.8	55.0
48	22.8	5.66e+07	1.09	(0.97 - 1.24)	0.0387	0.033	26.0	77.3	0.52	1.15	0.96	49.1	55.3
49	19.6	5.66e+07	1.12	(0.98 - 1.27)	0.0416*	0.036	23.0	80.6	0.52	1.19	0.96	50.0	55.4
50	16.2	5.66e+07	1.21*	(1.06 - 1.38)	0.0594	0.045	20.2	84.3	0.52	1.28	0.95	52.0	55.6
51	13.4	5.65e+07	1.28**	(1.12 - 1.47)	0.0672**	0.051	17.8	87.3	0.53	1.40	0.94	54.1	55.7
52	**11.1**	**5.64e+07**	**1.42****	**(1.24 - 1.63)**	**0.0827****	**0.058**	**15.7**	**90.1**	**0.53**	**1.59**	**0.94**	**57.3**	**55.9**
53	9.7	5.64e+07	1.41**	(1.21 - 1.63)	0.0699**	0.048	13.3	91.5	0.52	1.57	0.95	56.9	55.6
54	8.3	5.64e+07	1.45**	(1.25 - 1.69)	0.0695**	0.051	12.1	93.1	0.53	1.74	0.94	59.5	55.6
55	6.7	5.64e+07	1.50**	(1.28 - 1.77)	0.0629**	0.050	10.3	94.7	0.52	1.94	0.95	62.0	55.6
56	5.6	5.64e+07	1.49**	(1.26 - 1.76)	0.0563**	0.046	9.1	95.4	0.52	2.00	0.95	62.8	55.5
57	4.7	5.66e+07	1.37**	(1.12 - 1.66)	0.0344*	0.029	6.7	96.2	0.51	1.76	0.97	59.8	55.0
58	3.6	5.66e+07	1.35*	(1.09 - 1.67)	0.0272*	0.022	5.5	96.6	0.51	1.64	0.98	58.1	54.8
59	3.1	5.66e+07	1.44**	(1.15 - 1.80)	0.0285*	0.023	4.8	97.5	0.51	1.91	0.98	61.7	54.8
60	2.6	5.66e+07	1.37*	(1.07 - 1.76)	0.0222*	0.018	4.1	97.7	0.51	1.81	0.98	60.4	54.7
61	2.4	5.66e+07	1.40*	(1.09 - 1.80)	0.0221*	0.018	3.9	97.9	0.51	1.87	0.98	61.2	54.7
62	2.1	5.66e+07	1.42*	(1.08 - 1.87)	0.0174*	0.013	3.0	98.4	0.51	1.82	0.99	60.5	54.6
63	1.8	5.66e+07	1.38*	(1.02 - 1.87)	0.0137	0.011	2.6	98.5	0.51	1.69	0.99	58.8	54.5
64	1.7	5.66e+07	1.29	(0.91 - 1.82)	0.0089	0.006	2.1	98.6	0.50	1.46	0.99	55.2	54.4
65	1.5	5.66e+07	1.38	(0.97 - 1.96)	0.0101	0.007	1.9	98.8	0.50	1.62	0.99	57.7	54.4

[a] All calculations take into account clustering and probability weighting; [b] Rate ratio and corresponding 95% Confidence Interval from univariate Poisson regression. [c] Area under the receiver operating characteristic (ROC) curve; [d] Likelihood Ratio; [e] Positive Predictive Value; [f] Negative Predictive Value.

* *P* value < 0.05; ** *P* value < 0.01.

Table 2. Measures[a] of model fit, calibration and model discrimination for long work hours and incident cancer: Panel Study of Income Dynamics, 1986-2011.

Cut points (weekly work hours)	Proportion above cut point (%)	Measures of model fit and calibration				Model discrimination							
		Bayesian Information Criteria	RR [b]	(95% CIs) [b]	Somers' *D* Statistics	Youden Index (*J*)	Sensitivity (%)	Specificity (%)	AUC (*c*) [c]	LR [d] (+)	LR [d] (-)	% PPV [e]	% NPV [f]
36	94.6	3.62E+07	0.69	(0.45 - 1.04)	-0.0192	0.000	91.6	5.0	0.50	0.96	1.67	11.6	81.5
37	87.4	3.62E+07	0.88	(0.64 - 1.22)	-0.0055	0.000	85.9	12.4	0.50	0.98	1.13	11.8	86.6
38	79.6	3.62E+07	0.86	(0.65 - 1.13)	-0.0145	0.000	78.3	20.2	0.50	0.98	1.07	11.8	87.3
39	71.9	3.62E+07	0.82	(0.64 - 1.05)	-0.0316	0.000	70.3	27.9	0.50	0.98	1.06	11.7	87.4
40	63.4	3.62E+07	0.97	(0.77 - 1.23)	0.0054	0.009	64.3	36.7	0.50	1.01	0.97	12.1	88.3
41	55.6	3.62E+07	1.06	(0.85 - 1.34)	0.0294	0.021	57.4	44.6	0.51	1.04	0.95	12.4	88.5
42	48.5	3.62E+07	1.00	(0.80 - 1.25)	0.0123	0.002	48.7	51.5	0.50	1.00	1.00	12.0	88.1
43	42.7	3.62E+07	0.99	(0.79 - 1.24)	0.0093	0.000	42.6	57.3	0.50	1.00	1.00	12.0	88.0
44	37.9	3.62E+07	1.04	(0.83 - 1.31)	0.0229	0.005	38.4	62.1	0.50	1.01	0.99	12.1	88.1
45	34.2	3.62E+07	1.06	(0.85 - 1.34)	0.0279	0.000	34.2	65.8	0.50	1.00	1.00	12.0	88.0
46	30.1	3.62E+07	1.15	(0.91 - 1.45)	0.0462	0.017	31.6	70.1	0.51	1.06	0.98	12.6	88.3
47	26.3	3.62E+07	1.22	(0.96 - 1.55)	0.0564	0.030	28.9	74.1	0.51	1.11	0.96	13.2	88.4
48	22.8	3.62E+07	1.27	(1.00 - 1.63)	0.0611*	0.035	25.9	77.7	0.52	1.16	0.95	13.6	88.5
49	19.6	3.62E+07	1.16	(0.89 - 1.51)	0.0355	0.023	21.7	80.7	0.51	1.12	0.97	13.2	88.3
50	16.2	3.62E+07	1.32*	(1.01 - 1.74)	0.0560*	0.036	19.4	84.2	0.52	1.23	0.96	14.3	88.5
51	13.4	3.61E+07	1.45*	(1.10 - 1.92)	0.0657*	0.051	17.9	87.2	0.53	1.40	0.94	16.0	88.6
52	**11.1**	**3.60E+07**	**1.62****	**(1.22 - 2.17)**	**0.0742***	**0.055**	**16.0**	**89.5**	**0.53**	**1.53**	**0.94**	**17.2**	**88.7**
53	9.7	3.61E+07	1.57*	(1.14 - 2.16)	0.0612*	0.045	13.7	90.8	0.52	1.49	0.95	16.9	88.6
54	8.3	3.60E+07	1.71**	(1.23 - 2.38)	0.0650*	0.052	12.9	92.3	0.53	1.68	0.94	18.6	88.6
55	6.7	3.61E+07	1.70*	(1.18 - 2.46)	0.0543*	0.036	9.9	93.7	0.52	1.58	0.96	17.7	88.4
56	5.6	3.60E+07	1.86**	(1.27 - 2.70)	0.0564*	0.041	9.1	94.9	0.52	1.80	0.96	19.7	88.5
57	4.7	3.61E+07	1.67*	(1.12 - 2.48)	0.0383*	0.033	7.6	95.7	0.52	1.75	0.97	19.2	88.4
58	3.6	3.62E+07	1.40	(0.89 - 2.22)	0.0182	0.015	4.9	96.5	0.51	1.43	0.98	16.3	88.2
59	3.1	3.62E+07	1.54	(0.96 - 2.46)	0.0212	0.017	4.6	97.1	0.51	1.57	0.98	17.6	88.2
60	2.6	3.62E+07	1.46	(0.86 - 2.46)	0.0157	0.013	3.8	97.5	0.51	1.53	0.99	17.2	88.2
61	2.4	3.62E+07	1.62	(0.97 - 2.73)	0.0193	0.016	3.8	97.8	0.51	1.75	0.98	19.2	88.2
62	2.1	3.62E+07	1.65	(0.97 - 2.83)	0.0162	0.016	3.4	98.1	0.51	1.84	0.98	20.0	88.2
63	1.8	3.62E+07	1.45	(0.78 - 2.69)	0.0097	0.010	2.7	98.3	0.51	1.61	0.99	17.9	88.1
64	1.7	3.62E+07	1.49	(0.80 - 2.75)	0.0102	0.011	2.7	98.4	0.51	1.66	0.99	18.4	88.1
65	1.5	3.62E+07	1.57	(0.81 - 3.05)	0.0099	0.009	2.3	98.7	0.50	1.70	0.99	18.8	88.1

[a] All calculations take into account clustering and probability weighting; [b] Rate ratio and corresponding 95% Confidence Interval from univariate Poisson regression. [c] Area under the receiver operating characteristic (ROC) curve; [d] Likelihood Ratio; [e] Positive Predictive Value; [f] Negative Predictive Value.

* *P* value < 0.05; ** *P* value < 0.01.

Table 3. General description of the combined study populations by selected characteristics at study baseline (1986) by average work hours per week: Panel Study of Income Dynamics, 1986-2011.

	Average work hours per week [a]		*P*-value [b]
	35−51 hours n = 1,954 (86.7%)	**≥ 52 hours n = 378 (13.3%)**	
Sex			>0.001[c]
Male	1,161 (61.3)	300 (88.5)	
Female	793 (38.7)	78 (11.5)	
Age			0.021[d]
Mean (Standard Error)	33.6 (0.25)	35.4 (0.71)	
Educational level (highest completed)			0.007[c]
Did not complete high school	166 (7.0)	16 (5.3)	
High school diploma	1,245 (61.6)	140 (50.6)	
College degree	543 (31.4)	110 (44.1)	
Race			0.019[c]
White	1,425 (89.9)	235 (94.2)	
Non-white	513 (10.1)	35 (5.8)	
Ethnicity			0.972[c]
Hispanic	36 (2.6)	4 (2.6)	
Non-Hispanic	1,899 (97.4)	262 (97.4)	
Marital status			>0.001[c]
Married/ cohabitating	1,492 (75.6)	230 (86.5)	
Not married or cohabitating	462 (24.4)	148 (13.5)	
Number of children in the household			0.088[d]
Mean (Standard Error)	1.06 (0.04)	1.20 (0.08)	
Employment status			>0.001[c]
Self-employed	180 (12.5)	80 (33.9)	
Employed by others	1,579 (87.5)	179 (66.1)	
Industry			0.488[c]
Services	1,174 (68.1)	177 (70.6)	
Non-services	588 (31.9)	81 (29.4)	
Occupation			0.612[c]
Manual	587 (29.1)	88 (30.9)	
Non-manual	1,201 (70.9)	172 (69.1)	
Pay status			>0.001[c]
Salaried	746 (46.4)	138 (54.7)	
Hourly	826 (40.0)	40 (11.9)	
Other arrangement	216 (13.6)	82 (33.4)	
Incident cardiovascular disease			>0.001[c]
Yes	1,075 (54.7)	173 (67.2)	
No	832 (45.3)	92 (32.8)	
Incident cancer			0.001[c]
Yes	220 (12.5)	56 (21.1)	
No	1,734 (87.5)	322 (78.9)	

[a] Proportions adjusted for probability weighting.
[b] All calculations take into account clustering and probability weighting.
[c] X^2 test.
[d] Adjusted Wald test.

References

1. Virtanen M, Ferrie J, Singh-Manoux A, *et al.* Overtime work and incident coronary heart disease: The Whitehall II prospective cohort study. *Eur Heart J*. 2010;31(14):1737-1744.

2. Simpson C, Severson R. Risk of injury in African American hospital workers. *J Occ Env Med*. 2000;42:1035-1040.

3. Dembe A, Erickson J, Delbos R, Banks S. The impact of overtime and long work hours on occupational injuries and illnesses: New evidence from the United States. *Occ Env Med*. 2005;62:588-597.

4. Hilton M, Whiteford H, Sheridan J, *et al.* The prevalence of psychological distress in employees and associated occupational risk factors. *J Occup Environ Med*. 2008;50:746-757.

5. Virtanen M, Singh-Manoux A, Ferrie J, Gimeno D, Marmot MEM, al. e. Long working hours and cognitive function: The Whitehall II Study. *Am J Epidemiol*. 2009;41(2):596-605.

6. Bannai A, Tamakoshi A. The association between long working hours and health: A systematic review of epidemiological evidence. *Scand J Work Environ Health*. 2014:40(1):5-18.

7. Caruso C, Bushnell T, Eggerth D, *et al.* Long working hours, safety, and health: Toward a national research agenda. *Am J Ind Med*. 2006;49:930-942.

8. Choi G, Schnall P, Dobson M, *et al.* Very long (>48 hours) shifts and cardiovascular strain in firefighters: A theoretical framework. *Ann Occup Environ Med*. 2014;26(5):1-9.

9. World Health Organization. *WHO healthy workplace framework and model: Synthesis report* 2010.

10. National Institute for Organizational Safety and Health, Centers for Disease Control and Prevention. *The Changing Organization of Work and the Safety and Health of Working People*. Cincinnati, OH 2002. Report No. 2002-116.

11. Spurgeon A, Harrington J, Cooper C. Health and safety problems associated with long working hours: A review of the current position. *Occup Environ Med*. 1997;54:367-375.

12. Sparks K, Cooper C, Fried Y, Shirom A. The effects of hours of work on health: A meta-analytic review. *J Occup Organ Psychol*. 1997;70:391-408.

13. Harrington J. Shift work and health -- A critical review of the literature on working hours. *Annals of the Academy of Medicine, Singapore*. 1994;23:699-705.

14. Folkard S. A critique of J.M. Harrington. *Unpublished manuscript*. October 1994.

15. Mathers C, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med*. 2006;3:e442.

16. Kuper H. Systematic review of prospective cohort studies of psychosocial factors in the etiology and prognosis of coronary heart disease. *Semin Vasc Med*. 2002;2:267-314.

17. Kivimäki M, Virtanen M, Elovainio M, Kouvonen A, Väänänen A, Vahtera J. Work stress in the etiology of coronary heart disease: Systematic review and meta-analysis of prospective cohort studies. *Scand J Work Environ Health*. 2006;32:431-442.

18. Eller N, Netterström B, Gyntelberg F, *et al.* Work-related psychosocial factors and the development of ischemic heart disease. *Cardiol Rev*. 2009;17:83-97.

19. Kang MY, Park H, Seo JC, *et al.* Long working hours and cardiovascular disease: A meta-analysis of epidemiologic studies. *JOEM*. 2012;54(5):532-537.

20. Davis S, Mirick D, Stevens R. Night-shift work, light at night, and risk of breast cancer. *J Natl Cancer Inst*. 2001;93:1557-1562.

21. Schernhammer E, Laden F, Speizer F, al e. Rotating night shifts and risk of breast cancer

in women participating in the Nurses' Health Study. *J Natl Cancer Inst*. 2001;93:1563-1568.

22. University of Michigan Institute for Social Research. PSID: Data on Families Since 1968. http://psidonline.isr.umich.edu/Guide/Brochures/PSID.pdf. Accessed Apr 11, 2014.

23. McGonagle K, Schloeni R, Sastry N, Freedman V. The Panel Study of Income Dynmaics: Overview, recent innovations, and potential for life course research. *Longi Life Course Stud*. 2012;3(2):268-284.

24. Hill M. *The Panel Study of Income Dynamics: A User's Guide*. Vol Volume 2 of Guides to Major Social Science Data Bases: SAGE Publications, Inc.; 1991.

25. United States Bureau of Labor Statistics. Appendix C. How to Compute Your Firm's Incidence Rate for Safety Management. http://www.bls.gov/iif/oshwc/osh/os/osh06_appc.pdf. Accessed Nov. 21, 2014.

26. Okura Y, Urban L, Mahoney D, Jacobsen S, Rodeheffer R. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *J Clin Epidemiol*. 2004;57(10):1096-1103.

27. Kehoe R, Wu SY, Leske M, Cheylack L. Comparing self-reported and physician-reported medical history. *Am J Epidemiol*. 1994;139(8):813-818.

28. Paganini-Hill A, Chao A. Accuracy of recall of hip fracture, heart attack, and cancer: A comparison of postal survey data and medical records. *Am J Epidemiol*. 1993;138(2):101-106.

29. Bergmann M, Calle E, Mervis C, Miracle-McMahill M, Tun M, Health C. Validity of self-reported cancers in a prospective cohort study in comparison with data from state cancer registries. *Am J Epidemiol*. 1998;147(6):556-562.

30. Newson R. Confidence intervals for rank statistics: Somers' D and extensions. *The Stata Journal*. 2006;6(3):309-334.

31. Cook N. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*. 2007;115:928-935.

32. Higgins PB, Gower BA, Hunger GR, Goran MI. Defining helath-related obesity in prepubertal children. *Obesity Research*. 2001;9(4):233-240.

33. Youden W. Index for rating diagnostic tests. *Cancer*. 1950;3(1):32-35.

34. Cook N. Statistical evaluation of prognostic versus diagnostic models: Beyond the ROC curve. *Clinical Chemistry*. 2008;54(1):17-23.

35. Conway S, Pompeii L, Roberts R, Follis J, Gimeno D. The identification of a threshold beyond which long work hours elevates the risk of poor self-reported general health. *In preparation*.

36. Artazcoz L, Cortes I, Borrell C, Escriba-Aguir V, Cascant L. Gender perspective in the analysis of the relationship between long workhours, health and health-realted behavior. *Scand J Work Environ Health*. 2007;33(5):344-350.

37. Ribet C, Derriennic F. Age, working conditions, and sleep disorders: a longitudinal analysis in the French cohort E.S.T.E.V. *Sleep*. 1999;22(4):491-504.

CHAPTER V: DOSE-RESPONSE RELATION BETWEEN WORK HOURS AND CARIDOVASCULAR DISEASE: FINDINGS FROM THE PANEL STUDY OF INCOME DYNAMICS

Name of Journal Proposed for Article Submission: Journal of Occupational and Environmental Medicine

Abstract

Background. Working long hours has been linked to cardiovascular disease (CVD); however, an association between a dose response relationship between the number of hours worked and CVD has not been previously reported.

Methods. We conducted a retrospective cohort study using historical work hour and CVD indicator data (1986-2011) among a representative sample of 1,926 U.S. individuals from the Panel Study of Income Dynamics who were employed for a minimum of 10 years over the study duration. Logistic regression was employed to examine the mean number of work hours per year over the study period and CVD. To examine the dose response relationship we employed a series of nested logistic regression models. Hierarchical regression analyses using restricted cubic spline regression was used to estimate the dose-response relationship of work hours and CVD.

Results. There were 822 incident cases of CVD (42.7%), and weekly work hours ranged from 1.56 to 90.95, on average, over the study duration. The risk of CVD was associated with working longer hours. For each additional hour worked, participants were at a 2% elevated risk of CVD (OR: 1.02; 95% CI: 1.00 − 1.03; *p*-value: 0.036). Similar results were demonstrated by a subgroup analysis of full-time workers but were not seen in a subgroup analysis of part-time workers. The dose-response model produced by restricted cubic spline regression demonstrated that increasing work hours were associated with increased risk of CVD from 46 work hours per week and above. Odds ratios were calculated to compare specific work hour durations using the equation for the cubic spline regression, which was generated from the regression coefficients. From this, we found a 33 percent increase in the odds of developing CVD in participants working an average of 55 hours per week for ten years or more compared to those working 44 hours per week for the same duration (OR: 1.33; 95% CI: 1.02 − 1.73).

Conclusions. These findings suggest that longer work hour durations place workers at increased risk for CVD.

Background

Evidence that working long hours increases the risk of adverse health conditions has been documented in the literature for more than 50 years.[1] By definition, working long hours entails work hours beyond those of a full-time employee,[2] which is believed to increase the risk of exposure to occupational stress and reduce the amount of recovery time available to those exposed.[3,4] Extended work hours also have been associated with behavioral factors known to increase the risk of ill health, including smoking, increased alcohol consumption, and lower levels of physical activity. [5-8] These work-related and lifestyle factors are known to contribute to the development of cardiovascular disease (CVD), [9-11] with several recent studies reporting a positive association between long work hours and increased risk of CVD.[2,12] CVD is the leading cause of mortality worldwide and accounts for almost one-third of all deaths.[13] Although the number of CVD-related deaths has been declining over the past 40 years,[13] it is predicted to be a leading cause of morbidity and mortality for at least the next several decades.[14] However, to our knowledge, no study has focused on describing and modeling the dose-response relationship between work hour durations and CVD.

The purpose of this study was to assess the association between work hours and incident CVD, including the presence of a dose-response relationship, in a large, representative sample of U.S. workers who are participants of the Panel Study of Income Dynamics (PSID). Examining the dose-relationship relationship between work hours and health outcomes can be challenging given the possible inter-correlation between hours worked and demographic and occupational characteristics, as well as job demands.[15] We employed hierarchical regression models, including spline regression, because they provided advantages over conventional regression models when adjusting for covariates that are strongly interrelated[16] as well as when examining possible non-linear relationships between exposure and risk. [17]

Methods

Study Population

We conducted a retrospective cohort study utilizing existing data collected from participants enrolled in the PSID, which has been described in detail elsewhere.[18,19] Briefly, the PSID is a longitudinal survey of a representative sample of U.S. households, which was performed annually from 1968 until 1997, and biennially since 1999 to present.[19] Participants have been surveyed on their work hours since 1968 and on specific health outcomes, including CVD, since 1999. There are currently more than 22,000 individuals in 9,000 families in the PSID sample,[18] and response rates have surpassed 90 percent since 1970.[19]

In the current study, we included PSID participants aged 18 years or older who were employed for at least 10 years over the study duration. More specifically, they were included if they reported non-zero work hours for at least 10 years between 1986 and 2011 (n = 3,596). The cohort was further restricted by excluding those missing CVD data (n = 364), those reporting CVD at the baseline health study in 1999 (n = 658), and those reporting other prevalent chronic health conditions or disabilities at baseline (n = 529) or missing those data (n = 119). Our final sample for these analyses included 1,926 workers.

Data Collection

For each year in the study, the total annual hours of work was calculated from self-reported data for the previous year on the number of jobs worked; the number of hours worked per week, on average, per job; the amount of overtime worked per week, on average; the number of weeks of worked missed for any reason (i.e., illness, vacation, strike, unemployment, or layoff); and the start and stop dates for up to four jobs. The total annual hours of work for each study year was divided by 50 weeks of work to generate the number of work hours per week in the previous year on all jobs. The work hour variable included in these analyses was the arithmetic mean of work hours per week across the study duration.

For the outcome of CVD, participants were asked to report if they had ever been told by a physician that they had angina, coronary heart disease, congestive heart failure, a heart attack, high blood pressure/ hypertension, or a stroke. From this information, a dichotomized CVD variable was constructed to identify individuals who did and did not report incident CVD. Incident CVD was defined as the first affirmative report of any CVD-related outcome.

Covariate occupational data were collected in each survey wave, including employment status, industry, occupation, pay status, and whether they were self-employed. These values were dichotomized for the purpose of this analysis: industry (i.e., service vs. non-service), occupation (i.e., manual vs. non-manual), pay status (i.e., hourly vs. salary), and self-employment (i.e., self-employed vs. employed by other). Socio-demographic data were also collected, including age, sex, marital status, education level, and race/ethnicity. In addition, household membership data were collected, including number of children in the household, household and individual income (e.g., annual household income, annual individual income by job, annual overtime income), household and individual expenses (e.g., monthly rent/mortgage payment, annual child care expenses), and health insurance status.

Statistical Analysis

We employed descriptive analyses for the purpose of describing selected socio-demographic and occupational characteristics at baseline by outcome status. Adjusted Wald tests and χ^2 statistics were used to compare the means of continuous covariates and prevalences of categorical covariates, respectively, to determine whether there were significant differences between the participants who developed CVD and the ones who did not.

A list of potential covariates for inclusion in the multivariate model was compiled from the literature,[2,20-22] and those factors were matched to PSID variables. Many of the possible risk factors were not available in the data set, however, and some underfitting was expected due to the limitation of the available variables. Factors that were incident outcomes or collinear with work hour durations were eliminated from consideration, leaving variables that occurred prior to or at the same time as the exposure. Each remaining covariate was evaluated to determine whether it was associated with the outcome of interest. Covariates for which most – but not all – of the repeated measures over the study duration were associated with the outcome (e.g., occupation, paid hourly vs. salaried, health insurance status) were evaluated to determine whether using the measure at baseline was as explanatory as using all years of measures. Many of the job-related variables demonstrated significant collinearity; as a result, multiple iterations of step-down variable selection were performed to determine the effect of removing each of the work-related covariates from the rest of the set. Variables capturing income, expenses, household composition, and insurance status were not statistically significant and were excluded from the model. All regression models were adjusted for the same factors: age, sex, race/ethnicity, and pay status (salaried vs. hourly). Tests for interactions and effect measure modification were not significant. A stratified analysis of part-time (defined as participants averaging <35 WH/w over the study duration) and full-time (defined as participants averaging ≥35 WH/w over the study duration) workers was performed.

To analyze the dose-response relationship for work hours (expressed as a continuous measure of average work hours per week over the study duration) and CVD, a series of nested logistic regression models was compared. [23] The models used to evaluate the dose-response relationship were (1) the linear logistic regression model; (2) the quadratic logistic regression model, in which a quadratic term for the exposure was added to the logistic model; (3) the step function, in which the exposure was categorized; (4) the restricted linear spline logistic regression model; and (5) the restricted cubic spline logistic regression model. Unlike the logistic or quadratic models, the spline models allowed flexibility in modeling the dose-response curve. Splines are functions defined by piecewise polyonomials connected at the edges; the slope of each polynomial segment can differ, thereby allowing for greater flexibility of the curve to fit an unknown functional form.[24] In spline regression, the exposure was divided into a number of line segments that are connected at junction points, or "knots". Compared to conventional methods, such as linear logistic regression or step function models, linear splines have advantages in modeling continuous predictors, including flexibility in modeling the underlying functional form and easily interpretable model parameters.[24] Quadratic splines, including cubic splines, offer some of the same advantages as linear splines, and they further smooth out the connections between the splines by allowing a parabolic curve to connect the line segments at the points of intersection.[25] Restricted regression splines were utilized to reduce the instability of the estimates at the tails.[26] In order to determine odds ratios (ORs) for participants with different mean work hours, the equation for the cubic spline regression was generated using the regression coefficients, and ORs comparing specific work hour durations were calculated.

The step function required the categorization of the exposure, which was done by dividing work hours into quartiles. The quartile representing the lowest number of work hours was the reference category for step function estimates. For the restricted linear spline regression, the optimal number of segments was selected by introducing an increasing number of knots at subsequent quintiles, and the best fitting model was defined as the one that minimized the Akaike Information Criterion (AIC).[27] The AIC is a penalized likelihood that takes into account the number of parameters estimated in the model. As such, it is a variable selection criterion that balances good model fit with model parsimony. The AIC has been shown to perform poorly with small sample sizes or in the presence of too many parameters in relation to the size of the sample[28]; neither of these conditions affected our study. The same approach was used to select the optimal number of segments for the restricted cubic spline regression, with two additional models generated using the mvrs and the rc_spline programs in Stata; all cubic spline models were evaluated in terms of AIC to identify the best fitting model.

The five logistic regression models (i.e., logistic, quadratic, step, restricted linear spline, and restricted cubic spline) were then compared to determine the best fitting model among the set, which was defined as the one that balanced parsimony and model performance. Parsimony was evaluated in terms of degrees of freedom, and model performance was evaluated in terms of the Hosmer Lemeshow goodness of fit statistic and the AIC. All analyses were performed using Stata/SE software (version 13.1; Stata Corp, College Station, TX) and took into account clustering and probability weighting of the PSID data.

Results

At study baseline, those who would develop CVD were older, more likely to be racial minorities, and reported a greater number of children residing in the household compared to non-cases; they were also more likely to be self-employed (Table 1). Work hours per week ranged from 1.56 to 90.95 (data not shown), and differences in mean work hours were noted between the groups, with cases reporting an average work week that was an hour longer than that of non-cases (cases: 39.6 hours; standard error (S.E.): 0.4; non-cases: 38.5 hours; S.E.: 0.3; *p*-value: 0.023). No differences between groups were found in terms of educational level, Hispanic ethnicity, and marital status. Industry, occupational, and pay status categories were also not significantly different between cases and non-cases.

In the univariate model, we found a 1 percent increase in the risk of CVD for each additional hour of work per week, on average, but the result was not significant (95% C.I.: 1.00 − 1.02; *p*-value: 0.092) (Table 2). In the multivariate logistic model, which was adjusted for age, sex, race/ethnicity, and pay status, there was evidence of a 2 percent increase in the odds of CVD for each additional hour worked per week, on average, for a minimum of 10 years (OR: 1.02; 95% CI: 1.00 − 1.03; *p*-value: 0.036).

When we stratified our analyses by part- and full-time work status (Table 3), we observed no significant association between average weekly work hours and CVD among part-time workers in the crude or the adjusted models (aOR: 0.98; 95% CI: 0.94 − 1.03). For full-time workers, the univariate model demonstrated a 1.02-fold increase in the risk of CVD per additional hour of weekly work time (95% CI: 1.00 − 1.04; *p*-value: 0.048). Similar results were seen in the mulitvariate model, which demonstrated that the odds of CVD among full-time workers increased by 3 percent per additional hour of work time per week.

In these analyses, we discovered significant evidence of departure from linearity (*p*-value: 0.03), and the restricted cubic spline model was determined to provide the best fit among the models considered. In our examination of a dose-response relationship, we observed a decrease in the risk of CVD as work hours increased among those who worked fewer than 30 hours per week (Figure 1). CVD risk then increased as weekly work hours approached 40 hours per week before declining again between 40 and 46 WH/w. Beginning at 46 WH/w, increasing work hours were progressively associated with increased risk of CVD. Using the equation for the cubic spline regression, which was generated using the regression coefficients, we found a 33 percent increase in the odds of developing CVD in participants working an average of 55 hours per week for ten years or more compared to those working 44 hours per week for the same duration (OR: 1.33; 95% CI: 1.02 − 1.73).

Discussion

To our knowledge, this is the first report of findings pertaining to an assessment of a dose-response relationship between work hours and CVD. We found that the risk of CVD increased as average weekly work hours increased. The logistic regression model indicated that each additional hour of work increased the odds of CVD by approximately 2 percent. Extrapolating this result to a five-hour increase in average weekly work hours for a minimum of 10 years would increase the odds of incident CVD by 20 percent; a ten-hour increase for the same duration would increase the odds of incident CVD by 40 percent. More specifically, we observed a difference in the risk profiles of part-time and full-time workers, as no association was seen between hours worked and CVD in part-time workers. Further, the shape of the dose-response curve suggested that the most substantial risk was among those working 46 hours per week or greater, on average, for at least ten years.

There is a long history of research on the relationship between occupational stressors and CVD,[29] and this study agrees with the findings of two recent meta-analyses, which presented evidence that the risk of CVD increased as working hours increased.[12, 30] Kang *et al.* (2012) reported a 37 percent increase in the odds of CVD among those exposed to long work hours (the definition of which varied by study) compared to those unexposed (95% CI: 1.11 − 1.70), which was not influenced by age, geographical location, or study year. A subgroup analysis of participants working ≤55 hours per week attenuated the relationship between work hour duration and CVD (OR: 1.28; 95% CI: 0.85 − 1.91).[30] Virtanen *et al.* (2012) evaluated the more limited outcome of coronary heart disease (CHD) and demonstrated that working long hours (definition varied by study) was associated with an estimated 80 percent increase in the risk of CHD (RR: 1.80; 95% CI: 1.42 − 2.29). In a series of subgroup analyses, they reported that studies defining LWH as a threshold above 50 WH/w had substantially higher RRs than those studies defining LWH as a threshold equal to or below 50 WH/w ($RR_{\leq 50\ WH/w}$: 1.41; 95% CI: 1.14 − 1.74 vs. $RR_{>50WH/w}$: 2.37; 95% CI: 1.56 − 3.59).[12]

The dose-response relationship described in the current study is consistent with existing evidence of a positive association between work hours and CVD, particularly at longer work hour durations. Several mechanisms of effect have been hypothesized to explain the impact of working long hours on CVD. Extended work hour durations are believed to interfere with physiological recovery mechanisms because the exposure requires longer time at work, thereby reducing the amount of rest and recovery time, particularly in terms of sleep hours.[3,4] Insufficient sleep has been identified as an independent risk factor for ischemic heart disease,[31] acute myocardial infarction,[31] and coronary artery disease.[32,33] Another hypothesis is that long work hours are associated with individual behaviors that increase the risk of poor health outcomes.[5-7] Working long hours has been cited as a barrier to physical activity[34] and has been associated with increased odds for unhealthy weight gain,[35,36] smoking,[37] increased alcohol use,[35,37,38] and decreased physical activity,[37,39] each of which has been positively associated with increased CVD risk above certain exposure levels.[40,41] Additionally, the relationship between extended work hours and CVD is believed to be mediated by psychosocial stress, including occupational stress. Long work hour durations have been associated with significant increases in job stress,[6,42] and job stress has been associated with CVD. [43-45] Personality type may exacerbate the influence of work stress on health; type A personality, characterized as a drive to achieve coupled with a sense of urgency and a predisposition to competitiveness and irritability, has been associated with both LWH and increased risk of CVD.[46] However, it is unknown whether work hour durations are a causal risk factor or only an indicator of increased CVD risk.[12]

An advantage of these analyses is that we were able to construct a hierarchical set of nested logistic regression models to identify the model that best captured patterns of exposure and odds of outcome among the data. [23] In these analyses, restricted cubic spline regression offered the advantage of parsimony as well as smoothness and flexibility in the shape of the dose response curve and was considered superior to the other models under consideration. However, one drawback of cubic spline models is that the coefficients associated with the spline segments are not interpretable.[47] As an alternative, specific work hour durations can be compared by entering those values into the regression equation. Using this method, we found evidence that a ten hour increase from 45 to 55 WH/w increased the odds of CVD by 33 percent. These results were not attenuated by sex, and no evidence of effect measure modification was found in terms of sex, industry, or occupation.

Significant relationships between hours worked and CVD were found in each of the models under consideration. Compared to the cubic spline model, however, the other models may have lacked the flexibility to adequately capture the relationship between hours worked and CVD. Because cubic spline modeling accounts for variation among the data within categories as well as across categories, the resulting dose response curve does not assume linearity, as is the case with the logistic model, and it may be less sensitive to category choice than models such as the step model.[48]

Our findings should be interpreted within the context of the study limitations. PSID data lacked several potentially influential covariates, including many known risk factors for CVD, which may have resulted in residual confounding. Our measure of interest (i.e., work hours, CVD status) were self-reported, which may have introduced bias into the analysis. However, validation studies have demonstrated strong agreement between medical records and most self-reported CVD-related symptoms and events [49-51] as well as strong positive correlation between cross-sectional PSID self-reported total hours worked in the previous year and weekly work hours reported by employers.[52] Further, the selection of the best fitting model involved some subjective criteria. Assigning categories or spline knots to produce better fitting models would have required *a priori* knowledge that is currently unavailable. Instead, an iterative process was undertaken to determine the best placement of the knots given the data, and the shape of the curve was assessed.

These analyses represent the first characterization of the dose-response relationship between hours worked and CVD. Given that a recent study suggested that adding information on work hours to the Framingham risk score improved the predictive power of the model in low-risk working populations,[53] future research on the health impacts of work hour durations should focus on furthering our understanding of the potential for a threshold of effect of hours worked.

Acknowledgements

Partial funding for this project was received from an Occupational Epidemiology Traineeship and a Pilot Projects Research Training Grant, both from the National Institute for

Occupational Safety and Health (NIOSH)/ Centers for Disease Control and Prevention to the University of Texas School of Public Health Southwest Center for Occupational and Environmental Health, a NIOSH Education and Research Center (Grant No. 5T42OH008421).

Table 1. General description of the study population by selected characteristics at study baseline (1986) by outcome category: Panel Study of Income Dynamics, 1986-2011.

	Incident cardiovascular disease[a]		***P*-value**[b]
	Cases n = 822 (42.7%)	**Non-cases** n = 1,104 (57.3%)	
Sex			0.109[c]
Male	446 (56.3)	533 (51.5)	
Female	376 (43.7)	571 (48.5)	
Age (years)			<0.001[d]
Mean (Standard Error)	35.5 (0.4)	30.9 (0.3)	
Educational level (highest completed)			0.325[c]
Did not complete high school	79 (7.3)	80 (7.1)	
High school diploma	491 (58.2)	696 (59.4)	
College degree	252 (34.5)	328 (33.5)	
Race			<0.001[c]
White	571 (89.3)	904 (94.4)	
Non-white	246 (10.7)	197 (5.6)	
Ethnicity			0.756[c]
Hispanic	14 (2.4)	18 (2.1)	
Non-Hispanic	801 (97.6)	1,078 (97.9)	
Marital status			0.223[c]
Married/ cohabitating	664 (79.8)	927 (82.6)	
Not married or cohabitating	158 (20.2)	177 (17.4)	
Number of children in the household			0.043[d]
Mean (Standard Error)	1.18 (0.06)	1.04 (0.05)	
Employment status			0.045[c]
Self-employed	104 (18.4)	112 (13.7)	
Employed by others	617 (81.6)	806 (86.3)	
Industry			0.357[c]
Services	497 (69.7)	662 (72.4)	
Non-services	224 (30.3)	262 (27.6)	
Occupational social class			0.347[c]
Manual	228 (28.2)	260 (25.6)	
Non-manual	499 (71.8)	671 (74.4)	
Pay status			0.285[c]
Salaried	290 (44.9)	405 (49.2)	
Hourly	335 (39.2)	402 (37.8)	
Other arrangement	103 (15.9)	124 (13.0)	
Work hours per week, on average			0.023[d]
Mean (Standard Error)	39.6 (0.4)	38.5 (0.3)	

[a] Proportions adjusted for probability weighting.
[b] All calculations take into account clustering and probability weighting.
[c] Pearson X^2 test.
[d] Adjusted Wald test.

Table 2. Association between work hour duration and incident cardiovascular disease: Panel Study of Income Dynamics, 1986-2011.

	Univariate Model[a]		**Multivariate Model**[a,b]	
Variable	OR	(95% C.I.)	OR	(95% C.I.)
Work Hours per Week	1.01	(1.00 − 1.02)	1.02	(1.00 − 1.03)*
Age	1.05	(1.04 – 1.07)	1.06	(1.04 − 1.07)
Sex				
Female	1.00	Referent	1.00	Referent
Male	1.22	(0.96 – 1.54)	1.05	(0.76 − 1.44)
Race/ Ethnicity				
White, Non-Hispanic	1.00	Referent	1.00	Referent
Black, Non-Hispanic	2.18	(1.54 – 3.09)	2.51	(1.62 − 3.88)
Hispanic	1.77	(0.84 – 3.73)	2.01	(0.85 − 4.76)
Other	2.70	(1.09 – 6.71)	2.39	(0.89 − 6.38)
Pay Status				
Salaried	1.00	Referent	1.00	Referent
Hourly	1.19	(0.92 – 1.53)	1.37	(1.04 − 1.80)

[a] All calculations take into account clustering and probability weighting.

[b] Model adjusted for age, sex, race/ethnicity, and pay status (hourly vs. salaried).

* *P* value < 0.05

Table 3. Association between work hour duration and incident cardiovascular disease in part-time workers (<35 work hours per week, on average) and full-time (≥35 work hours per week, on average): Panel Study of Income Dynamics, 1986-2011.

	Part-Time Workers (n = 592)				**Full-Time Workers (n = 1,334)**			
	Univariate Model[a]		**Multivariate Model[a,b]**		**Univariate Model[a]**		**Multivariate Model[a,b]**	
Variable	OR	(95% C.I.)	OR	(95% C.I.)	OR	(95% C.I.)	OR	(95% C.I.)
Work Hours per Week	0.99	(0.96 − 1.03)	0.98	(0.94 − 1.03)	1.02	(1.00 − 1.04)*	1.03	(1.01 − 1.05)
Age	1.04	(1.02 − 1.06)	1.06	(1.04 − 1.09)	1.05	(1.04 − 1.08)	1.06	(1.04 − 1.08)
Sex								
Female	1.00	Referent	1.00	Referent	1.00	Referent	1.00	Referent
Male	1.38	(0.76 − 2.53)	0.84	(0.42 − 1.70)	1.16	(0.85 − 1.58)	1.07	(0.74 − 1.55)
Race/ Ethnicity								
White, Non-Hispanic	1.00	Referent	1.00	Referent	1.00	Referent	1.00	Referent
Black, Non-Hispanic	2.97	(1.47 − 6.00)	3.40	(1.40 − 8.28)	1.90	(1.22 − 2.96)	2.45	(1.46 − 4.11)
Hispanic	1.90	(0.64 − 5.62)	1.40	(0.38 − 5.16)	1.75	(0.63 − 4.85)	3.17	(0.86 − 11.71)
Other	2.65	(0.47 − 15.09)	3.46	(0.75 − 15.97)	2.70	(0.90 − 8.09)	2.30	(0.63 − 8.43)
Pay Status								
Salaried	1.00	Referent	1.00	Referent	1.00	Referent	1.00	Referent
Hourly	1.24	(0.74 − 2.09)	1.42	(0.80 − 2.53)	1.19	(0.88 − 1.60)	1.32	(0.97 − 1.82)

[a] All calculations take into account clustering and probability weighting.
[b] Model adjusted for age, sex, race/ethnicity, and pay status (hourly vs. salaried).
* *P* value < 0.05

Figure 1. Restricted cubic spline model for the relationship between work hour duration and incident cardiovascular disease: Panel Study of Income Dynamics, 1986-2011.

References

1. Sparks K, Cooper C, Fried Y, Shirom A. The effects of hours of work on health: A meta-analytic review. *J Occup Organ Psychol*. 1997;70:391-408.

2. Bannai A, Tamakoshi A. The association between long working hours and health: A systematic review of epidemiological evidence. *Scand J Work Environ Health*. 2014:40(1):5-18.

3. Meijman T, Mulder G. Psychological aspects of workload. In: Drenth P, Thierry H, De Wolff C, eds. *Handbook of Work and Organizational Psychology, Vol. 2*. Hove (U.K.): Psychology Press; 1998.

4. Totterdell P, Spelten E, Smith L, Baron J, Folkard S. Recovery from work shifts: How long does it take? *J Appl Psychol*. 1995;80:43-57.

5. House J, Strecher V, Metzner H, Robins C. Occupational stress and health among men and women in the Tecumseh community health study. *J Health Soc Behav*. 1986;27:65-77.

6. Maruyama S, Morimoto K. Effects of long workhours on lifestyle, stress and quality of life among intermediate Japanese managers. *Scand J Work Enviorn Health*. 1996;22:353-359.

7. Raggatt P. Work stress among long-distance coach drivers: A survey and correlational study. *J Organ Behav*. 1991;12:565-579.

8. Steptoe A, Kivimaki M. Stress and cardiovascular disease: An update on current knowledge. *Annu Rev Public Health*. 2013;34:337-354.

9. Kuper H. Systematic review of prospective cohort studies of psychosocial factors in the etiology and prognosis of coronary heart disease. *Semin Vasc Med*. 2002;2:267-314.

10. Kivimäki M, Virtanen M, Elovainio M, Kouvonen A, Väänänen A, Vahtera J. Work stress in the etiology of coronary heart disease: Systematic review and meta-analysis of prospective cohort studies. *Scand J Work Environ Health*. 2006;32:431-442.

11. Eller N, Netterström B, Gyntelberg F, *et al*. Work-related psychosocial factors and the development of ischemic heart disease: A systematic review. *Cardiol Rev*. 2009;17:83-97.

12. Virtanen M, Heikkila K, Jokela M, *et al*. Long working hours and coronary heart disease: A systematic review and meta-analysis. *Am J Epidemiol*. 2012;176(7):586-596.

13. Organization WH. Fact sheet No 317: Cardiovascular diseases (CVDs). Jan 2015.

http://www.who.int/mediacentre/factsheets/fs317/en/. Accessed Mar 22, 2015.

14. Mathers C, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med*. 2006;3:e442.

15. Hughes E, Parkes K. Work hours and well-being: The roles of work-time control and work-family interference. *Work Stress*. 2007;21(3):264-278.

16. Greenland S. Methods for epidemiologic analyses of multiple exposures: A review and comparative study of maximum-likelihood, preliminary testing, and empirical-Bayes regression. *Stat Med*. 1993;12:717-736.

17. Witte J, Greenland S. A nested approach to evaluating dose-response and trend. *Ann Epidemiol*. 1997;7:188-193.

18. University of Michigan Institute for Social Research. PSID: Data on Families Since 1968. http://psidonline.isr.umich.edu/Guide/Brochures/PSID.pdf. Accessed Apr 11, 2014.

19. McGonagle K, Schloeni R, Sastry N, Freedman V. The Panel Study of Income Dynmaics: Overview, recent innovations, and potential for life course research. *Longi Life Course Stud*. 2012;3(2):268-284.

20. Caruso C, Bushnell T, Eggerth D, *et al*. Long working hours, safety, and health: Toward

a national research agenda. *Am J Ind Med*. 2006;49:930-942.

21. National Institute for Organizational Safety and Health, Centers for Disease Control and Prevention. *The Changing Organization of Work and the Safety and Health of Working People*. Cincinnati, OH 2002. Report No. 2002-116.

22. Choi G, Schnall P, Dobson M, *et al*. Very long (>48 hours) shifts and cardiovascular strain in firefighters: A theoretical framework. *Ann Occup Environ Med*. 2014;26(5):1-9.

23. International Council on Mining & Metals. *Good Practice Guidance on Occupational Health Risk Assessment* 2009.

24. de Boor C. *A practical guide to splines*. New Your 1978. Volume 27, Appl Math Sci.

25. Witte J, Greenland S. A nested approach to evalutating dose-response and trend. *Ann Epidemiol*. 1997;7(3):188-193.

26. Harrell F, Lee K, Pollock B. Regression models in clinical studies: Determining relationships between predictors and response. *J Natl Cancer Inst*. 1988;80:1198-1202.

27. Akaike H. Information theory and an extension of the maximum likelihood principle. Paper presented at: Second International Symposium on Information Theory,

1974; Budapest.

28. Sugiura N. Further analysis of the data by Akaike's information criterion and the finite corrections. *Commun Statist Theory Meth*. 1978;7:13-26.

29. Russek H, Zohman B. Relative significance of heredity, diet and occupational stress in coronary heart disease of young adults; based on an analysis of 100 patients between the ages of 25 and 40 years and a similar group of 100 normal control subjects. *Am J Med Sci*. 1958;235(3):266-277.

30. Kang MY, Park H, Seo JC, *et al*. Long working hours and cardiovascular disease: A meta-analysis of epidemiologic studies. *JOEM*. 2012;54(5):532-537.

31. Liu Y, Tanaka H. Overtime work, insufficient sleep, and risk of non-fatal acute myocardial infarction in Japanese men. *Occup Environ Med*. 2002;59:447-451.

32. Wingard D, Berkman L. Mortality risk associated with sleeping patterns amond adults. *Sleep*. 1983;6:102-107.

33. Partinen M, Putkonen P, Kaprio J, Koskenvuo M, Hilakivi I. Sleep disorders in relation to coronary heart disease. *Acta Med Scand Suppl*. 1982;660:69-83.

34. Taris T, Ybema J, Beckers D, Werheijden M, Geurts S, Kompier M. Investigating the associations among overtime work, health behaviors, and health: A longitudinal

study among full-time employees. *Int J Behav Med*. 2011;18:352-360.

35. Shields M. Long working hours and health. *Health Rep*. 1999;11:33-48.

36. Nakamura K, Shimai S, Kikuchi S, *et al*. Increases in body mass index and waist circumference as outcomes of working overtime. *Occup Med (Lond)*. 1998;48:169-173.

37. Song JT, Lee G, Kwon J, Park JW, Choi H, Lim S. The association between long working hours and self-rated health. *Ann Occ Environ Med*. 2014;26(2):1-12.

38. Trinkoff A, Storr C. Work schedule characteristics and substance use in nurses. *Am J Ind Med*. 1998;34:266-271.

39. Mizoue T, Reijula K, Andersson K. Environmental tobacco smoke exposure and overtime work as risk factors for sick building syndrome in Japan. *Am J Epidemiol*. 2001;154:803-808.

40. Mozaffarian D, Wilxon P, Kannel W. Beyond established and novel risk factors: Lifestyle risk factors for cardiovascular disease. *Circ*. 2008;117:3031-3038.

41. Husain K, Ansari R, Ferder L. Alcohol-induced hypertension: Mechanism and prevention. *World J Cardiol*. 2014;6(5):245-252.

42. Park J, Yi Y, Kim Y. Weekly work hours and stress complaints of workers in Korea. *Am J Ind Med*. 2010;53:1135-1141.

43. Uchiyama S, Kurasawa T, Sekizawa T, Nakatsuka H. Job strain and risk of cardiovasculare events in treated hypertensive Japanese workers: Hypertension follow-up group study. *J Occup Health*. 2005;23:102-111.

44. Johnson J, Stewart W, Hall E, Fedlund P, Theorell T. Long-term psychosocial work environment and cardiovascular mortality among Swedish men. *Am J Public Health*. 1996;86:324-331.

45. Cooper C, Marshall J. Occupational sources of stress: A review of the literature relating to coronary heart disease and mental ill health. *J Occup Psych*. 1976;49:11-28.

46. Friedman M, Rosenman R. Type A behavior pattern -- Its association with coronary heart-disease. *Ann Clin Res*. 1971;3:300-312.

47. Desquilbet L, Mariotti F. Dose-response analyses using restricted cubic spline functions in public health research. *Statist Med*. 2010;29:1037-1057.

48. Greenland S. Dose-response and trend analysis in epidemiology: Alternatives to categorical analysis. *Epidemiology*. 1995;6:356-365.

49. Okura Y, Urban L, Mahoney D, Jacobsen S, Rodeheffer R. Agreement between self-

report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *J Clin Epidemiol*. 2004;57(10):1096-1103.

50. Kehoe R, Wu SY, Leske M, Cheylack L. Comparing self-reported and physician-reported medical history. *Am J Epidemiol*. 1994;139(8):813-818.

51. Paganini-Hill A, Chao A. Accuracy of recall of hip fracture, heart attack, and cancer: A comparison of postal survey data and medical records. *Am J Epidemiol*. 1993;138(2):101-106.

52. Bound J, Brown C, Duncan G, Rodgers W. *Measurment error in cross-sectional and longitudinal labor market surveys: Results from validation studies*. Cambridge, MA 1989. Working Paper No. 2884.

53. Kivimaki M, Batty G, Hamer M, *et al*. Using additional information on working hours to predict coronary heart disease. *Ann Intern Med*. 2011;154:457-463.

CHAPTER VI: CONCLUSIONS

This research was motivated by a lack of evidence on the influence of specific working durations on adverse health outcomes, to which large numbers of U.S. workers are exposed each year. This project was intended to improve our understanding of the LWH-health relationship by offering evidence to enhance current approaches to defining and operationalizing LWH.

For each of our outcomes of interest, our analyses demonstrated that the same threshold of LWH − an average of 52 WH/w for a minimum of 10 years − best predicted poor SRGH, incident CVD, and incident cancer in each of their respective cohorts. This suggested that patterns of work hours may have been more consistent across the cohorts than was anticipated. To determine whether this was due to significant overlap among the cohorts, we calculated that a total of 2,414 participants met the inclusion criteria for at least one of the cohorts from which a statistically optimized LWH cut point was derived. Of those 2,414 participants, 1,627 (or, 67.4%) were included in all three cohorts. However, the participants reporting the outcomes of interest did not overlap to the same degree. Of the 1,369 participants reporting any one of the outcomes of interest over the study duration, 274 participants (or, 20.0%) reported both poor SRGH and CVD, 108 participants (or, 7.9%) reported both poor SRGH and cancer, 90 participants (or, 6.6%) reported both CVD and cancer, and 35 participants (or, 2.6%) reported all three outcomes of interest. Because a relatively small proportion of individuals were cases in more than one analysis, this suggests that the optimized threshold value of 52 WH/w was derived from the working hours of a unique set of case data for each of the different cohorts.

The analyses of LWH and the three adverse health outcomes studied herein – poor SRGH, incident CVD, and incident cancer – lay the groundwork for other studies to examine LWH-disease relationships by applying a health risk-based exposure level threshold. Future investigations of LWH and health should include validating the threshold value of 52 WH/w for a minimum of 10 years in other large, representative data sets (e.g., NHANES) and with other health outcomes (e.g., mental health conditions). Further investigations into dose-response relationships with outcomes other than CVD would advance our understanding of the potential effects of LWH at various exposure thresholds.

As average work hours in the U.S. continue to lengthen, it is critical that we improve our understanding of the relationship between work hours, illness, and injury. Given that the U.S. is the only industrialized nation lacking federally-mandated maximum work hours for individuals over the age of 16 years in general industry, these findings suggest that the regulation of hours of work may be a fundamental − yet under-addressed − policy issue. Working-time arrangements directly impact the health and well-being of workers, and work time standards reflect a society's perception of its workers' rights and protections.[158] Although numerous legal instruments exist for addressing working time, national legislation limiting weekly work time is widely considered the most effective measure for reducing excessively long working hours. [158] This research suggests that there may be value in pursuing a statutory limit on working hours with the goal of reducing the greatest risk of adverse health conditions while balancing the needs of industry.

APPENDICES

Appendix A: University of Texas Health Science Center Institutional Review Board Outcome Notices

UTHealth
The University of Texas
Health Science Center at Houston

Committee for the Protection of Human Subjects

6410 Fannin Street, Suite 1100
Houston, Texas 77030

Dr. Sarai Conway
School of Public Health

January 03, 2014

<u>HSC-SPH-13-0871</u> - *Health Effects of Long Working Hours*

The above named project is determined to qualify for exempt status according to 45 CFR 46.101(b)

CATEGORY #4 : *Research, involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified directly or through identifiers linked to the subjects.*

Health Insurance Portability and Accountability Act:
Exempt from HIPAA

CHANGES: Should you choose to make any changes to the protocol that would involve the inclusion of human subjects or identified data from humans, please submit the change via iRIS to the Committee for the Protection of Human Subjects for review.

STUDY CLOSURES: Upon completion of your project, submission of a study closure report is required. The study closure report should be submitted once all data has been collected and analyzed.

Should you have any questions, please contact the Office of Research Support Committees at 713-500-7943.

UTHealth
The University of Texas
Health Science Center at Houston

Committee for the Protection of Human Subjects
6410 Fannin Street, Suite 1100
Houston, Texas 77030

NOTICE OF APPROVAL TO IMPLEMENT REQUESTED CHANGES

November 18, 2014

HSC-SPH-13-0871 - Health Effects of Long Working Hours
PI: Dr. Sarai Conway

Reference Number: 116765

PROVISIONS; Unless otherwise noted, this approval relates to the research to be conducted under the above referenced title and/or to any associated materials considered at this meeting, e.g. study documents, informed consent, etc.

APPROVED: By Expedited Review and Approval

CHANGE APPROVED: Revisions to research aims and methodology

REVIEW DATE: November 18, 2014

APPROVAL DATE: November 18, 2014

CHAIRPERSON: John C. Ribble, MD John C. Ribble

Upon receipt of this letter, and subject to any provisions noted above, you may now implement the changes approved.

CHANGES: The principal investigator (PI) must receive approval from the CPHS before initiating any changes, including those required by the sponsor, which would affect human subjects, e.g. changes in methods or procedures, numbers or kinds of human subjects, or revisions to the informed consent document or procedures. The addition of co-investigators must also receive approval from the CPHS. ALL PROTOCOL REVISIONS MUST BE SUBMITTED TO THE SPONSOR OF THE RESEARCH.

INFORMED CONSENT: Informed consent must be obtained by the PI or designee(s), using the format and procedures approved by the CPHS. The PI is responsible to instruct the designee in the methods approved by the CPHS for the consent process. The individual obtaining informed consent must also sign the consent document. **Please note that if revisions to the informed consent form were made and approved, then old blank copies of the ICF MUST be destroyed. Only copies of the appropriately dated, stamped approved informed consent form can be used when obtaining consent.**

UNANTICIPATED RISK OR HARM, OR ADVERSE DRUG REACTIONS: The PI will immediately inform the CPHS of any unanticipated problems involving risks to subjects or others, of any serious harm to subjects, and of any adverse drug reactions.

RECORDS: The PI will maintain adequate records, including signed consent documents if required, in a manner that ensures subject confidentiality.

REFERENCES

1. Bannai A, Tamakoshi A. The association between long working hours and health: A systematic review of epidemiological evidence. *Scand J Work Environ Health*. 2014:40(1):5-18.

2. International Labour Organization. *Measurement of working time*. Geneva, Switzerland: 18th International Conference of Labour Statisticians; 2008.

3. United States Department of Labor. Office of the Assistant Secretary for Policy - elaws - Fair Labor Standards Act Advisor. 2013. Available at: http://www.dol.gov/elaws/faq/esa/flsa/011.htm. Accessed October 15, 2013.

4. Spurgeon A, Harrington J, Cooper C. Health and safety problems associated with long working hours: A review of the current position. *Occup Environ Med*. 1997;54:367-375.

5. Wang XS, Armstrong M, Cairns B, Key T, Travis R. Shift work and chronic disease: The epidemiological evidence. *Occ Med*. 2011;61(2):78-89.

6. Organisation for Economic Co-operation and Development. *Employment Outlook 1998. Working hours: Latest trends and policy initiativies.* Paris 1998. Chapter 5 in Employment Outlook.

7. Lee S, McCann D, Messenger J. *Working time around the world: Trends in working hours, laws and policies in a global comparative perspective*. Geneva 2013.

8. United States Department of Labor. Wage and Hour Division: Overtime Pay. 2014. Available at: http://www.dol.gov/whd/overtime_pay.htm. Accessed Mar 20, 2014.

9. Bureau of National Affairs I. *Staffing and Scheduling Strategies* 1994. Personal Policies Forum Survey No. 152.

10. Caruso C. Possible broad impacts of long work hours. *Indus Health*. 2006;44:531-536.

11. Johnson J, Lipscomb J. Long working hours, occupational health and the changing nature of work. *Am J Ind Med*. 2006;49:921-929.

12. Harrington J. Health effects of shift work and extended hours of work. *Occup Environ Med*. 2001;58:68-72.

13. European Parliament, Council of the European Union. Directive 2003/88/EC of the European Parliament and of the Council of 4 November 2003 concerning certain aspects of the organisation of working time. http://eur-lex.europa.eu/legal-content/EN/ALL/;jsessionid=JCbNTwZJVTpvLf6sgvkmP1hkHvJpqpbz0t2T2mwn85mhJZyMnbm5!435512023?uri=CELEX:32003L0088. Accessed Mar

23, 2014.

14. European Foundation for the Improvement of Living and Working Conditions. *Working Time Developments - 2010*. Dublin, Ireland 2011.

15. Ministry of Government Legislation. *Labor Standards Act: Laws on Green Growth and Economic Investment in Korea* 1997. English translation by the Korean Ministry of Labor.

16. Organisation for Economic Co-Operation and Development. StatExtracts average annual hours actually worked per worker. Available at: http://stats.oecd.org/index.aspx?DatasetCode=AVE_HRS. Accessed Jan 13, 2014.

17. Freyssinet J, Michon F. Overtime in Europe. *eiro online: European Industrial Observatory On-line*. July 1, 2003. Available at: http://www.eurofound.europa.eu/eiro/2003/02/study/tn0302101s.htm. Accessed April 13, 2014.

18. Foner P. *1947*. New York, NY: International Publishers; History of the labor movement in the United States from colonial times to the founding of the American Federation of Labor.

19. Hunnicutt B. The end of shorter hours. *Labor Hist*. 1984;25:373-404.

20. The New York Times. [Ford] Gives $10,000,000 To 26,000 Employees. *The New York Times*. Jan 1914.

21. Grossman J. *Fair Labor Standards Act of 1938: Maximum struggle for a minimum wage* 1978. Monthly Labor Review.

22. Bosch G. Working time: Tendencies and emerging issues. *Int Labor Reveiw*. 1999;138:131-148.

23. Mishel L. *Vast majority of wage earners are working harder, and for not much more: Trends in U.S. work hours and wages over 1979-2007* 2013. Issue Brief #348.

24. United States, Department of Interior, Census Office. Report on the statistics of wages in manufacturing industries. *1880 Census*. 1883;20.

25. United States Senate. *Senate Report 1394, Fifty-Second Congress, Second Session*. Washington 1893. Wholesale Prices, Wages, and Transportation.

26. Owen J. Work-time reduction in the United States and Western Europe. *Monthly Labor Review*. 1988;111.

27. Whaples R. The shortening of the American work week: An economic and historical

analysis of its context, causes, and consequences. *Ph.D. dissertation*. 1990.

28. Huberman M. *Working hours of the world unite? New international evidence of worktime, 1870-2000*. Montreal 2002. CIRANO Scientific Series 2002s-77.

29. Organisation for Economic Co-operation and Development. Average annual hours actually worked per worker. *OECD.StatExtracts*. Available at: https://stats.oecd.org/Index.aspx?DataSetCode=ANHRS. Accessed Apr 4, 2014.

30. Saad L. The "40-hour" workweek is actually longer -- by seven hours: Full-time U.S. workers, on average, report working 47 hours weekly. *Gallup*. Aug 29, 2014. Available at: http://www.gallup.com/poll/175286/hour-workweek-actually-longer-seven-hours.aspx. Accessed Oct 28, 2014.

31. Grosch J, Carusso C, Rosa R, Suter S. Long work hours in the U.S.: Associations with demographic and organizational characteristics, psychosocial working conditions and health. *Am J Ind Med*. 2006;49:943-952.

32. Schor J. Working hours and time pressure: The controversy about trends in time. *Working time: International trends*: Routledge: Theory and Policy Perspectives; 2000.

33. Caruso C, Bushnell T, Eggerth D, *et al*. Long working hours, safety, and health: Toward a national research agenda. *Am J Ind Med*. 2006;49:930-942.

34. Sparks K, Cooper C, Fried Y, Shirom A. The effects of hours of work on health: A meta-analytic review. *J Occup Organ Psychol*. 1997;70:391-408.

35. Ettner S, Grzywacz J. Workers' perceptions of how job affect health: A social ecological perspective. *J Occ Health Psychol*. 2001;6:101-113.

36. Virtanen M, Heikkila K, Jokela M, *et al*. Long working hours and coronary heart disease: A systematic review and meta-analysis. *Am J Epidemiol*. 2012;176(7):586-596.

37. Lipscomb J, Trinkoff A, Geiger-Brown J, Brady B. Work-schedule characteristics and reported musculoskeletal disorders of registered nurses. *Scand J Work Environ Health*. 2002;28:394-401.

38. Simpson C, Severson R. Risk of injury in African American hospital workers. *J Occ Env Med*. 2000;42:1035-1040.

39. Dembe A, Erickson J, Delbos R, Banks S. The impact of overtime and long work hours on occupational injuries and illnesses: New evidence from the United States. *Occ Env Med*. 2005;62:588-597.

40. Hilton M, Whiteford H, Sheridan J, *et al*. The prevalence of psychological distress in employees and associated occupational risk factors. *J Occup Environ Med*. 2008;50:746-757.

41. Kivisto M, Harma M, Sallinen M, Kalimo R. Work-related factors, sleep debt and insomnia in IT professionals. *Occup Med (Lond)*. 2008;58:138-140.

42. van der Hulst M. Long workhours and health. *Scand J Work Environ Health*. 2003;29:171-188.

43. Artozcoz L, Cortes I, Escriba-Aguir V, Bartoll X, Basart H, Borrel C. Long working hours and health status among employees in Europe: Between-country differences. *Scan J Work Environ Health*. 2013;39(4):369-378.

44. Proctor S, White R, Robins T, Echeverria D, Rocskay A. Effect of overtime work on cognitive function in automotive workers. *Scand J Work Environ Health*. 1996;22:124-132.

45. Borg V, Kristnsen T. Psychosocial work environment and mental health among travelling salespeople. *Work Stress*. 1999;13:132-143.

46. Nakanishi N, Nishina H, Yoshida H, *et al*. Hours of work and the risk of developing impaired fasting glucose or type 2 diabetes mellitus in Japanese male office

workers. *Occup Environ Med*. 2001;58:569-74.

47. Baldwin P, Wrate R. Young doctors' health -- I. How do working conditions affect attitutdes, health and performance. *Soc Sci Med*. 1997;45(1):35-40.

48. Tarumi K, Hagihara A, Morimoto K. A prospective observation of onsets of health defects associated with working hours. *Ind Health*. 2003;41(2):101-108.

49. Holtermann A, Mortensen O, Burr H, al. e. Long work hours and physical fitness: 30-year risk of ischaemic heart disease and all-cause mortality among middle-aged Caucasian men. *Heart*. 2010;96(20):1638-1644.

50. Harma M. Are long workhours a health risk? *Scand J Work Environ Health*. 2003;29:167-169.

51. Sorensen G, Pirie P, Folsom A, al e. Sex differences in the relationship between work and health: The Minnesota Heart Survey. *J Health Soc Behav*. 1985;26:379-394.

52. Kawakami N, Araki S, Haratani T, al e. Relations of work stress to alcohol use and drinking problems in male and female employees of a computer factory in Japan. *Environ Res*. 1993;1993:314-324.

53. Artazcoz L, Cortes I, Escriba-Agur V, Cascant L, Villegas R. Understanding the

relationship of long working hours with health status and health-related behaviors. *J Epidemiol Community Health*. 2009;63:521-527.

54. Caruso C, Hitchcock E, Dick R, al e. *Overtime and extended work shifts: Recent findings on illnesses, injuries and health behaviors*. Cincinnati 2004.

55. Theorell T, Rahe R. Behavior and life satisfaction characteristics of Swedish subjects with myocardial infarction. *J Chronic Dis*. 1972;25(3):139-147.

56. Kawakami N, Akachi K, Shimizu H, *et al*. Job strain, social support in the workplace, and haemoglobin A1c in Japanese men. *Occup Environ Med*. 2000;57:805-809.

57. Jex S, Bliese P. Efficacy beliefs as a moderator of the impact of work-related stressors: a multilevel study. *J Appl Psychol*. 1999;84:349-361.

58. Kroenke C, Spiegelman D, Manson J, Schernhammer E, Colditz G, Kawachi I. Work charactersitics and incidence of type 2 diabetes in women. *Am J Epidemiol*. 2007;165:175-183.

59. Vyas M, Garg A, Iansavichus A, *et al*. Shift work and vascular events: systematic review and meta-analysis. *BMJ*. 2012;345:1-11.

60. United States Department of Labor, Bureau of Labor Statistics. BLS Information:

Glossary. 2013. Available at: http://www.bls.gov/bls/glossary.htm#F. Accessed November 19, 2013.

61. Kuhn R, Rahman O, Menken J. Survey Measures of Health: How Well Do Self-Reported and Observed Indicators Measure Health and Predict Mortality? In: Cohen B, Menken J, eds. *Aging in Sub-Saharan Africa: Recommendation for Furthering Research*. Wasghington, DC: National Academies Press; 2006.

62. Jylha M. What is self-rated health and why does it predict mortality? Towards a unified conceptual model. *Soc Sci Med*. 2009;69:307-316.

63. Idler E, Benyamini Y. Self-rated health and mortality: A review of twenty-seven community studies. *J Health Soc Behav*. 1997;38:21-37.

64. Murray C, Salomon J, Mathers C. A critical examination of summary measures of population health. *Bulletin of the World Health Organization*. 2000;78(8):981-994.

65. Song JT, Lee G, Kwon J, Park JW, Choi H, Lim S. The association between long working hours and self-rated health. *Ann Occ Environ Med*. 2014;26(2):1-12.

66. Artazcoz L, Cortes I, Borrell C, Escriba-Aguir V, Cascant L. Gender perspective in the analysis of the relationship between long workhours, health and health-

realted behavior. *Scand J Work Environ Health*. 2007;33(5):344-350.

67. Baldwin P, Dodd M, Wrate R. Young doctors' health - I. How do working conditions affect attitudes, health and performance? *Soc Sci Med*. 1997;45(1):35-40.

68. Ross C, Mirowsky J. Does employment affect health. *J Health Soc Behav*. 1995;36:230-243.

69. Won J, Kim S, Lim H, Kim H. *A study on approval criteria of cardiovascular disease from overwork*. Seoul, Korea 2008.

70. Taris T, Ybema J, Beckers D, Werheijden M, Geurts S, Kompier M. Investigating the associations among overtime work, health behaviors, and health: A longitudinal study among full-time employees. *Int J Behav Med*. 2011;18:352-360.

71. Scheff T. *On Being Mentally Ill: A Sociobiological Theory*. London: Weidenfeld & Nicholson; 1966.

72. Hosmer D, Lemeshow S. *Model-Building Strategies and Methods for Logistic Regression*. 2nd ed: John Wiley & Sons, Inc.; 2000.

73. Kirkeleit J, Riise J, Bjorge T, al e. The healthy worker effect in cancer incidence studies. *Am J Epidemiol*. 2013;177(11):1218-1224.

74. Naimi A, Richardson D, Cole S. Causal inference in occupational epidemiology: Accounting for the healty worker effect by suing structural nested models. *Am J Epidemiol*. 2013;178(12):1681-1686.

75. Robins J. A new approach to casual inference in mortality studies with a sustained exposure period -- application to control of the healthy worker survivor effect. *Math Model*. 1986;7(9-12):1393-1512.

76. Beckers D, van Hooff M, van der Linden D, Kompier M, Taris T, Geurts S. A diary study to open up the black box of overtime work amont university faculty members. *Scand J Work Environ Health*. 2008;34(3):213-223.

77. Mitchell D. *Extended working hours in the Southeastern logging industry*: Auburn University; 2008. Ph.D. dissertation.

78. Pasicott P, Murphy G. Effect of work schedule design on productivity of mechanised harvesting operations in Chile. *New Zeal J For Sci*. 2013;43(2):1-10.

79. Siegrist J, Rodel A. Work stress and health risk behavior. *Scand J Work Environ Health*. 2006;32(6):473-481.

80. Taris T, Yberma J, Beckers D, Verheijden M, Geurts S, Kompier M. Investigating the associations among overtime work, health behaviors, and health: A longitudinal study among full-time employees. *Int J Behav Med*.

2011;18(4):352-360.

81. Meijman T, Mulder G. Psychological aspects of workload. In: Drenth P, Thierry H, De Wolff C, eds. *Handbook of Work and Organizational Psychology, Vol. 2*. Hove (U.K.): Psychology Press; 1998.

82. Totterdell P, Spelten E, Smith L, Baron J, Folkard S. Recovery from work shifts: How long does it take? *J Appl Psychol*. 1995;80:43-57.

83. McEwen B. Stress, adaptation, and disease: Allostatsis and allostatic load. *Ann NY Acad Sci*. 1998;840:33-44.

84. Frankenhaueser M. Coping with job stress -- A psychobiological approach. In: Gardell B, Johansson G, eds. *Working life: A social science contribution to work reform*. Chichester: John Wiley; 1981.

85. Rissler A. Stress reactions at work and after work during a period of quantitative overload. *Ergonomics*. 1977;20:13-16.

86. House J, Strecher V, Metzner H, Robins C. Occupational stress and health among men and women in the Tecumseh community health study. *J Health Soc Behav*. 1986;27:65-77.

87. Maruyama S, Morimoto K. Effects of long workhours on lifestyle, stress and quality

of life among intermediate Japanese managers. *Scand J Work Enviorn Health*. 1996;22:353-359.

88. Raggatt P. Work stress among long-distance coach drivers: A survey and correlational study. *J Organ Behav*. 1991;12:565-579.

89. Shields M. Long working hours and health. *Health Rep*. 1999;11:33-48.

90. Nakamura K, Shimai S, Kikuchi S, *et al*. Increases in body mass index and waist circumference as outcomes of working overtime. *Occup Med (Lond)*. 1998;48:169-173.

91. Trinkoff A, Storr C. Work schedule characteristics and substance use in nurses. *Am J Ind Med*. 1998;34:266-271.

92. Mizoue T, Reijula K, Andersson K. Environmental tobacco smoke exposure and overtime work as risk factors for sick building syndrome in Japan. *Am J Epidemiol*. 2001;154:803-808.

93. Lundberg U, Palm K. Workload and catecholamine excretion in parents of preschool children. *Work Stress*. 1989;3:255-260.

94. Sluiter J, Frings-Dresen M, van der Beek A, Meijman F, Heisterkamp S. Neuroendocrine reactivity and recovery from work with different physical and

mental demands. *Scand J Work Environ Health*. 2000;49:306-316.

95. Kasl S. Mental health and the work environment. *J Occup Med*. 1973;15:509-518.

96. Cooper C, Marshall J. Occupational sources of stress: A review of the literature reating to coronary heart disease and mental ill health. *J Occup Psych*. 1976;49:11-28.

97. Leino P. Symptoms of stress predict musculoskeletal disorders. *J Epidemiol Community Health*. 1989;43:293-300.

98. House J, McMichael A, Wells J. Occupational stress and health among factory workers. *J Health Soc Behav*. 1979;20:139-160.

99. National Institute for Organizational Safety and Health, Centers for Disease Control and Prevention. *The Changing Organization of Work and the Safety and Health of Working People*. Cincinnati, OH 2002. Report No. 2002-116.

100. Barton J, Spelten E, Totterdell P, Smith L, Folkard S, Costa G. The standard shiftwork index -- A battery of questionnaires for assessing shiftwork-related problems. *Work Stress*. 1995;9(1):4-30.

101. Sauter S, Brightwell W, Colligan M, *et al*. *The changing organization of work and safety and health of working people: Knowledge gaps and research directions*.

Cincinnati, OH 2002. DHHS (NIOSH) Publication No. 2002-116.

102. Barnett R. Toward a review and reconceptualization of the work/family literature. *Genetic, Social & General Psychology Monographs*. 1998;124(2):125-128.

103. Olds D, SP C. The effect of work hours on adverse events and errors in health care. *J Safey Res*. 2010;41(2):153-162.

104. Occupational Safety & Health Administration, United States Department of Labor. Permissible Exposure Limits (PELs). 2014. Available at: https://www.osha.gov/dsg/topics/pel/. Accessed Apr 2, 2014.

105. National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention. NIOSH Pocket Guide to Chemical Hazards. 2013. Available at: http://www.cdc.gov/niosh/npg/pgintrod.html. Accessed Apr 2, 2014.

106. American Conference of Governmental Industrial Hygienists. *2014 TLVs and BEIs*. Cincinnati, OH: American Conference of Governmental Industrial Hygienists; 2014.

107. International Council on Mining & Metals. *Good Practice Guidance on Occupational Health Risk Assessment* 2009.

108. Jahoda M. *Employment and unemployment: A social-psychological analysis*.

Cambridge: Cambridge University Press; 1982.

109. Landsbergis P. The changing organization of work and the health and safety of working people: A commentary. *JOEM*. 2003;45:61-72.

110. Blair A, Stewart P, Lubin J, Forastiere F. Methodological issues regarding confounding and exposure misclassification in epidemiological studies of occupational exposures. *Am J Ind Med*. 2007;50:199-207.

111. Bodnar L, Siega-Riz A, Simhan H, Diesel J, Abrams B. The impact of exposure misclassification on associations between prepregnancy body mass index and adverse pregnancy outcomes. *Obesity*. 2010;18(11):2148-2190.

112. Pierce N, Checkoway H, Kriebel D. Bias in occupational epidemiology studies. *Occup Environ Med*. 2007;64(8):562-568.

113. Copeland K, Checkoway H, McMichael A, Holbrook R. Bias due to misclassification in the estimation of relative risk. *Am J Epidemiol*. 1977;105:488-493.

114. Harrington J. Shift work and health -- A critical review of the literature on working hours. *Annals of the Academy of Medicine, Singapore*. 1994;23:699-705.

115. Folkard S. A critique of J.M. Harrington. *Unpublished manuscript*. October 1994.

116. Kuhn P, Lozano F. *The expanding workweek? Understanding trends in long work hours among U.S. men, 1979-2002*. Working Paper: http://ftp.iza.org/dp1924.pdf 2005.

117. United States Bureau of Labor Statistics. Occupational outlook handbook. 2005. http://www.bls.gov/ooh/. Accessed November 3, 2013.

118. McGonagle K, Schloeni R, Sastry N, Freedman V. The Panel Study of Income Dynmaics: Overview, recent innovations, and potential for life course research. *Longi Life Course Stud*. 2012;3(2):268-284.

119. Hill M. *The Panel Study of Income Dynamics: A User's Guide*. Vol Volume 2 of Guides to Major Social Science Data Bases: SAGE Publications, Inc.; 1991.

120. University of Michigan Institute for Social Research. PSID: Data on Families Since 1968. http://psidonline.isr.umich.edu/Guide/Brochures/PSID.pdf. Accessed Apr 11, 2014.

121. Ko HJ. Figure 1: Steady state panel schematic. *Military Life Course - Military Life Course Resources - Panel Study of Income Dynamics*. Aug 5, 2013. Available at: https://sharepoint.oregonstate.edu/sites/militarylifecourse/Shared%20Documents/PSID-Documents/PSID_sampleframe.bmp. Accessed Jul 7, 2014.

122. Fitzgerald J, Gottschalk P, Moffitt R. An analysis of sample attrition in panel data. *J Hum Resour*. 1998;33:251-299.

123. Fitzgerald J. Attrition in models of intergenerational links in health and economic status in the PSID. *Be J Econ Anal Poli*. 2011;11(3):Article 2.

124. Panel Study of Income Dynamics. Panel Study of Income Dynamics: Wave 38 - 2013 - Questionnaire. October 2013. ftp://ftp.isr.umich.edu/pub/src/psid/questionnaires/q2013.pdf. Accessed May 3, 2014.

125. Andreski P, McGonagle K, Schoeni R. *An analysis of the quality of the health data in the Panel Study of Income Dynamics* 2007. Technical Series Paper #07-04.

126. Ricci J, Chee E, Lorandeau M, Berger J. Fatigue in the U.S. workeforce: Prevalence and implications for lost productive work time. *JOEM*. 2007;49(1):1-10.

127. Bureau of Labor Statistics, United States Department of Labor Division of Labor Force Statistics. Labor Force Statistics from the Current Population Survey. Dec 08, 2014. Available at: http://www.bls.gov/cps/lfcharacteristics.htm#fullpart. Accessed Jan 05, 2015.

128. Bound J, Brown C, Duncan G, Rodgers W. *Measurment error in cross-sectional and longitudinal labor market surveys: Results from validation studies*. Cambridge,

MA 1989. Working Paper No. 2884.

129. Kriebel D, Checkoway H, Pearce N. Exposure and dose modelling in occupational epidemiology. *Occup Environ Med*. 2007;64:492-498.

130. Faraggi D, Simon R. A simulation study of cross-validation for selecting an optimal cutpoint in univariate survival analysis. *Statist Med*. 1996;15:2203-2213.

131. Heliovaara M, Aromaa A, Klaukka T, Knekt P, Joukamaa M, Impivaara O. Reliability and validity of interview data on chronic diseases - The mini-Finland health survey. *J Clin Epidemiol*. 1993;46(2):181-191.

132. Martin L, Leff M, Calonge N, Garrett C, Nelson D. Validation of self-reported chronic conditions and health services in a managed care populations. *Am J Prev Med*. 2000;18(3):215-218.

133. Okura J, Urban L, Mahoney D, Jacobsen S, Rodeheffer R. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *J Clin Epidemiol*. 2004;57(10):1096-1103.

134. Cavanaugh K, Merkin S, Plantinga L, Fink N, Sadler J, Powe N. Accuracy of patients' reports of comorbid disease and their association with mortality in ESRD. *Am*

J Kidney Dis. 2008;52(1):118-127.

135. Merkin S, Cavanaugh K, Longenecker J, Fink N, Levey A, Powe N. Agreement of self-reported comorbid conditions with medical and physician reports varied by disease among end-stage renal disease patients. *J Clin Epidemiol*. 2007;60(6):634-642.

136. Sanchez-Villegas A, Schlatter J, Ortuna F, *et al*. Validity of a self-reported diagnosis of depression among participants in a cohort study using the Structured Clinical Interview for DSM-IV (SCID-I). *BMC Psychiatry*. 2008;8(43).

137. Desai M, Bruce M, Desai R, Druss B. Validity of self-reported cancer history: A comparison of health interview data and cancer registry records. *Am J Epidemiol*. 2001;153(3):299-306.

138. Kriegsman D, Penninx B, Van Eijk J, Boeke A, Deeg D. Self-reports and general practitioner information on the presence of chrnic diseases in community dwelling elderly: A study on the accuracy of patients' self-reports and on determinants of inaccuracy. *J Clin Epidem*. 1996;49(12):1407-1417.

139. Cook N. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*. 2007;115:928-935.

140. Pepe M, Janes J, Longton G, Leisearing W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol*. 2004;159:882-890.

141. Cook N. Statistical evaluation of prognostic versus diagnostic models: Beyond the ROC curve. *Clinical Chemistry*. 2008;54(1):17-23.

142. Newson R. Parameters behind "nonparametric" statistics: Kendall's tau, Somers' D and median differences. *The Stata Journal*. 2002;2(1):45-64.

143. Hirose K, Kawano S, Konishi S, Ichikawa M. Bayesian information criterion and selection of the number of factors in factor analysis models. *J Data Sci*. 2011;9:243-259.

144. Perkins N, Schisterman E. The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve. *Am J Epidemiol*. 2006;163:670-675.

145. Schisterman E, Perkins N, Liu A, Bondell H. Optimal cut-point and its corresponding Youden Index to discriminate individuals using pooled blood samples. *Epidemiol*. 2005;16:73-81.

146. Youden W. Index for rating diagnostic tests. *Cancer*. 1950;3(1):32-35.

147. Klotsche J, Ferger D, Pieper L, Brhm J, Wittchen HU. A novel nonparametric approach for estimating cut-offs in continuous risk indicators with application to diabetes epidemiology. *BMC Med Res Methodol*. 2009;9(63).

148. Ginty F, Adak S, Can A, *et al*. The relative distribution of membranous and cytoplasmic met is a prognostic indicator in stage I and stage II colon cancer. *Clin Cancer Res*. 2008;14(12):3814-3822.

149. Choi G, Schnall P, Dobson M, *et al*. Very long (>48 hours) shifts and cardiovascular strain in firefighters: A theoretical framework. *Ann Occup Environ Med*. 2014;26(5):1-9.

150. World Health Organization. *WHO healthy workplace framework and model: Synthesis report* 2010.

151. Fleischer N, Diez Roux A. Using directed acyclic graphs to guide analysis of neighbourhood health effects: An introduction. *J Epidemiol Community Health*. 2008;62:842-846.

152. de Boor C. *A practical guide to splines*. New Your 1978. Volume 27, Appl Math Sci.

153. Witte J, Greenland S. A nested approach to evalutating dose-response and trend. *Ann Epidemiol*. 1997;7(3):188-193.

154. Harrell F, Lee K, Pollock B. Regression models in clinical studies: Determining relationships between predictors and response. *J Natl Cancer Inst*. 1988;80:1198-1202.

155. Akaike H. Information theory and an extension of the maximum likelihood principle. Paper presented at: Second International Symposium on Information Theory, 1974; Budapest.

156. Sugiura N. Further analysis of the data by Akaike's information criterion and the finite corrections. *Commun Statist Theory Meth*. 1978;7:13-26.

157. Podsakoff P, MacKenzie S, Lee JY, Podsakoff N. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *J Appl Psychol*. 2003;88(5):879-903.

158. International Labour Organization. *Working time in the twenty-first century*. Geneva 2011. Report for discussion at the Tripartite Meeting of Experts on Working-time Arrangements.