The Open Public Health Journal

Content list available at: https://openpublichealthjournal.com

RESEARCH ARTICLE

Impact of Physical Activity on Disease and Mortality: Evidence from National Health and Nutrition Examination Survey

Sarah Yildirim[1,*]

[1]*Choate Rosemary Hall Wallingford, CT, USA*

Abstract:

Background:

The Centers for Disease Control and Prevention in the United States recommends the amount and the types of physical activity needed to maintain and improve overall health and reduce the risk of chronic disease. Previous studies have shown that multiple objective physical activities obtained from continuously monitored accelerometer data are associated with mortality.

Objective:

This paper first analyzes the individual characteristics of the participants in the National Health and Nutrition Examination Survey (NHANES) who wore an accelerometer between 2003-2006.

Methods:

It studies the association between objectively measured physical activity intensity and several prevalent diseases and mortality.

Results:

Key findings include: (1) physical activity is positively associated with education level and negatively associated with obesity, diabetes, chronic heart failure, and smoking. (2) physical activity is associated with a lower mortality rate after controlling the current health status and smoking habits.

Conclusion:

In this regard, this study contributes to the public health literature on physical activities by incorporating the individual characteristics of the national health survey participants into the analyses and linking those characteristics to the relationship between physical activity and different types of major diseases, which was less addressed in previous studies.

Keywords: Physical activity, Diseases, Mortality, Public health, Smoking habits, NHANES.

Article History	Received: April 13, 2023	Revised: May 27, 2023	Accepted: June 06, 2023

1. INTRODUCTION

The diseases mentioned throughout this study are extraordinarily expensive for society. According to The Centers for Disease Control and Prevention (CDC), in 2020, the cost of heart disease, cancer, and diabetes was estimated at 715 billion dollars [1]. Physical activity (PA) is one of the key modifiable risk factors that could prevent and reduce the burden of these diseases [2]. The CDC recommends the amount and types of physical activity needed to maintain and improve overall health and reduce the risk of chronic disease [3]. Indeed, multiple studies document that an increase in PA is associated with a decrease in all-cause and all-specific mortality [4, 5]. In particular, mortality due to heart failure [6 - 8], cancer [2, 9, 10], and type 1 and 2 diabetes [11] was reported to be associated with lower levels of physical activity.

This paper further explores the association between objectively measured physical activity using accelerometers and all-cause and cause-specific mortality using NHANES.

* Address correspondence to this author at the Choate Rosemary Hall Wallingford, CT, USA; E-mail: syildirim23@choate.edu

2. METHODS

2.1. Data

The National Health and Nutrition Examination Survey (NHANES) is the primary data used in this study [12 - 14]. The NHANES is specifically designed to measure the health and nutritional status of adults and children in the United States. The survey is comprehensive in terms of methodology as it combines physical examinations and interviews for the same subject (*i.e.*, individual participants). Due to the scientific importance and confidentiality clause of the data, all the survey process and protocol, from the design of the questionnaires to data collection, is approved by CDC as well as the National Center for Health Statistics (NCHS) Ethics Review Board (ERB). Furthermore, the informed consent for the survey is publicly available on the NHANES website.

More specifically, it is a national-level data set that was started in 1999, with over 10,000 individuals randomly selected in each two-year assessment cycle. The random section is conducted through a statistical process using the U.S. Census information. Briefly describing the sample selection procedure, all the counties in the U.S. are divided into 15 groups based on their characteristics. From each of the 15 selected groups, between 20 and 24 neighborhoods are also selected based on the neighborhood characteristics. Then, a sample of about 30 households is chosen within each neighborhood. Finally, the randomly selected individuals are invited to an interview. After the interview, individuals go through a health examination. As specified in the Helsinki Declaration, confidentiality of personal information is strictly protected, and therefore personal information, such as names and addresses, are not disclosed in the data. In 2005-2006, the NHANES also released individual-level accelerometer data in addition to their health status information. This study particularly uses the information collected until December 31, 2015, for an average follow-up time of 10 years.

2.2. Physical Activity

NHANES collects objective physical activity data using hip-worn accelerometers, and it is publicly available from the CDC's webpage. The self-reported data covers two time periods, between 2003-2004 and 2005-2006, and is organized in the R-package "rnhanesdata" [15]. Following household interviews, participants are invited to have a health examination in an NHANES mobile examination center. The participants wore the ActiGraph AM-7164 device for seven consecutive days from the day of the NHANES examination and removed it during sleep and water-related activities. The NHANES data reports both the wear and non-wear time. Specifically, the non-wear time was identified as an interval of at least 60 consecutive minutes of 0 activity counts. The physical activity was summarized as an activity count at the minute level weekly for each participant. Therefore, individuals with high PA time are considered to move actively throughout the day.

2.3. Analysis

The NHANES 2003-2006 survey participants were followed up for mortality status, and their mortality status was available from NHANES *via* a link to the National Mortality Index [16]. This study specifically uses the data collected until December 31, 2015, for an average follow-up time of a 10-year period.

To distinguish individuals' physical activity levels, the total per week physical activity amount for each individual was calculated, and then indicator variables were created for the analysis, based on the 25th, 50th, or 75th percentile of total physical activity to identify low, medium, and high PA individuals to examine the PA levels. After deleting the missing information, the sample size became 11,170 participants with accelerometry data between the 2003-2004 and 2005-2006 cycles. This study first analyzes the dependent variables of high and low physical activity levels with different diseases and covariates such as age, gender, and education level to study the active and less active individuals' characteristics. Then, by running nine logit models, the following section of the study analyzes the impact of different physical activity levels on mortality.

3. RESULTS

3.1. Descriptive Statistics

Table **1** summarizes all variables in the final dataset, separating the data between alive (*i.e.*, assigned as 0) and dead (deceased, assigned as 1). The first column is for the entire sample. Participants with a critical underlying disease are associated with a high 10-year mortality rate in the sample. Those underlying diseases include diabetes, heart disease, and cancer. For example, 410 individuals out of 1057 with diabetes, 252 out of 464 with coronary heart disease (CHD), 250 out of 377 with congestive heart failure (CHF), 246 out of 404 with stroke, and 398 out of 992 with cancer died within ten years. Moreover, smokers are low-mobility individuals who are also at high risk. The data confirms that 731 individuals out of 2620 former smokers and 1036 out of 2184 with mobility difficulty also did not survive within ten years.

Table 1. Summary statistics between 2003 and 2006.

Characteristic	Overall, N = 11,170[1]	0, N = 9,302[1]	1, N = 1,868[1]
Age	44 (28, 64)	39 (25, 55)	75 (63, 83)
Gender	-	-	-
Male	5,365 (48%)	4,328 (47%)	1,037 (56%)
Female	5,805 (52%)	4,974 (53%)	831 (44%)
Education	-	-	-

(Table 1) contd.....

Characteristic	Overall, N = 11,170[1]	0, N = 9,302[1]	1, N = 1,868[1]
Less than 9th grade	1,353 (14%)	927 (11%)	426 (23%)
9-11th grade	1,525 (15%)	1,183 (14%)	342 (19%)
High school grad/GED	2,447 (24%)	1,978 (24%)	469 (26%)
Some College or AA degree	2,769 (28%)	2,377 (29%)	392 (21%)
College graduate or above	1,891 (19%)	1,695 (21%)	196 (11%)
Refused	11 (0.1%)	4 (0.1%)	7 (0.4%)
Don't know	12 (0.1%)	5 (0.1%)	7 (0.4%)
Unknown	1,162	1,133	29
Race	-	-	-
White	5,490 (49%)	4,324 (46%)	1,166 (62%)
Mexican American	2,344 (21%)	2,040 (22%)	304 (16%)
Other Hispanic	338 (3.0%)	310 (3.3%)	28 (1.5%)
Black	2,526 (23%)	2,203 (24%)	323 (17%)
Other	472 (4.2%)	425 (4.6%)	47 (2.5%)
BMI Category	-	-	-
Normal	3,308 (32%)	2,789 (32%)	519 (32%)
Underweight	206 (2.0%)	170 (1.9%)	36 (2.2%)
Overweight	3,490 (33%)	2,908 (33%)	582 (36%)
Obese	3,419 (33%)	2,947 (33%)	472 (29%)
Unknown	747	488	259
Diabetes	-	-	-
No	9,949 (89%)	8,543 (92%)	1,406 (75%)
Yes	1,057 (9.5%)	647 (7.0%)	410 (22%)
Borderline	156 (1.4%)	107 (1.2%)	49 (2.6%)
Refused	0 (0%)	0 (0%)	0 (0%)
Don't know	8 (0.1%)	5 (0.1%)	3 (0.2%)
CHD	-	-	-
No	9,489 (95%)	7,938 (97%)	1,551 (84%)
Yes	464 (4.6%)	212 (2.6%)	252 (14%)
Refused	0 (0%)	0 (0%)	0 (0%)
Don't know	54 (0.5%)	18 (0.2%)	36 (2.0%)
Unknown	1,163	1,134	29
CHF	-	-	-
No	9,592 (96%)	8,025 (98%)	1,567 (85%)
Yes	377 (3.8%)	127 (1.6%)	250 (14%)
Refused	0 (0%)	0 (0%)	0 (0%)
Don't know	38 (0.4%)	16 (0.2%)	22 (1.2%)
Unknown	1,163	1,134	29
Cancer	-	-	-
No	9,100 (91%)	7,664 (94%)	1,436 (78%)
Yes	892 (8.9%)	494 (6.0%)	398 (22%)
Refused	0 (0%)	0 (0%)	0 (0%)
Don't know	15 (0.1%)	10 (0.1%)	5 (0.3%)
Unknown	1,163	1,134	29
Stroke	-	-	-
Yes	404 (4.0%)	158 (1.9%)	246 (13%)
No	9,587 (96%)	8,005 (98%)	1,582 (86%)
Refused	0 (0%)	0 (0%)	0 (0%)
Don't know	16 (0.2%)	5 (¡0.1%)	11 (0.6%)
Unknown	1,163	1,134	29
Mobility Problem	-	-	-
No Difficulty	7,823 (78%)	7,020 (86%)	803 (44%)
Any Difficulty	2,184 (22%)	1,148 (14%)	1,036 (56%)
Unknown	1,163	1,134	29

(Table 1) contd.....

Characteristic	Overall, N = 11,170[1]	0, N = 9,302[1]	1, N = 1,868[1]
Drink Status	-	-	-
Moderate Drinker	4,914 (57%)	4,339 (61%)	575 (39%)
Non-Drinker	3,139 (36%)	2,324 (32%)	815 (55%)
Heavy Drinker	586 (6.8%)	495 (6.9%)	91 (6.1%)
Unknown	2,531	2,144	387
Smoke Cigs	-	-	-
Never	5,156 (52%)	4,405 (54%)	751 (41%)
Former	2,620 (26%)	1,889 (23%)	731 (40%)
Current	2,219 (22%)	1,864 (23%)	355 (19%)
Unknown	1,175	1,144	31
Cause of death	-	-	-
Accidents	64 (3.4%)	0 (NA%)	64 (3.4%)
All other causes	653 (35%)	0 (NA%)	653 (35%)
Alzheimer's disease	64 (3.4%)	0 (NA%)	64 (3.4%)
Cerebrovascular disease	93 (5.0%)	0 (NA%)	93 (5.0%)
Chronic lower respiratory dis.	101 (5.4%)	0 (NA%)	101 (5.4%)
Diabetes mellitus	56 (3.0%)	0 (NA%)	56 (3.0%)
Diseases of the heart	356 (19%)	0 (NA%)	356 (19%)
Influenza and pneumonia	46 (2.5%)	0 (NA%)	46 (2.5%)
Malignant neoplasms	388 (21%)	0 (NA%)	388 (21%)
Nephritis	46 (2.5%)	0 (NA%)	46 (2.5%)
Unknown	9,303	9,302	1
PA High	2,372 (25%)	1,946 (24%)	426 (30%)
Unknown	1,672	1,240	432
PA Low	2,375 (25%)	2,139 (27%)	236 (16%)
Unknown	1,672	1,240	432
PA Medium	4,751 (50%)	3,977 (49%)	774 (54%)
Unknown	1,672	1,240	432
Weekly Use Mins	5,209 (3,903, 6,048)	5,171 (3,798, 6,020)	5,460 (4,484, 6,191)
Unknown	1,672	1,240	432

Note: PA High, PA Low, and PA Medium are dummy variables created from the original data, based on the subject belonging to the 25th, 50th, or 75th percentile for weekly wearable device usage.

[1]Median (IQR); n (%)

High Body Mass Index (BMI) is another potential cause of mortality. Among the deceased group within the BMI group in Table **1**, the obese and overweight subjects have a high mortality rate. Another observation from Table **1** is that 22% of the deceased sub- jects have diabetes, 14% have coronary heart disease (CHD), 14% have congestive heart failure (CHF), 22% have cancer, 36% are overweight, 56% have mobility problems, 39% are moderate drinkers, 40% are former smokers, and 19% are current smokers. Amongst all the deceased subjects, college graduates have the lowest mortality rate compared to other education categories. Although the data used for this study does not include information on nutrition and eating habits based on education level, the analysis will later proxy the intensity of physical activity for college graduates to see the impact on the mortality rate. The mortality data merged with CDS's data also have the cause of death. However, the cause of death could be due to complications of other preexistent diseases. Therefore, this paper focuses on physical activity and its association with disease and mortality.

The basic descriptive statistics show that the individuals in the high-PA group are associated with a higher percentage of mortality compared to the low-PA group, even though each group has an almost identical number of observations. This implies that other variables should be factored in to examine the relationship between PA and mortality status.

3.2. Physical Activity Intensity

In Table **2**, a series of logistic regression models were analyzed to understand the survey data beyond descriptive statistics and study the association between objectively measured physical activity intensity and several prevalent diseases. In particular, the models reported in this table identify who is more/less engaged in physical activity. Specifically, three models (cancer, CFH, and CHD) are examined for high physical activity. Then similar models are discussed for low physical activity while keeping the diabetes dummy variable in all models. The results confirm that physically active individuals significantly decrease diabetes, chronic heart failure, and obesity. Individuals currently smoking are found to be associated with fewer PA times.

Regarding educational achievement, the highly educated group showed a significantly high level of physical activity. The odds of physical activity being high with a college education were 1.87 (*i.e.*, exp(0.63)=1.87) times higher than with someone without.

Table 2. Logistic regression with dependent variables of high and low physical activity.

-		PA High=1	-	-	PA Low=1	-
-	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
(Intercept)	−2.38***	−2.39***	−2.36***	0.36***	0.37***	0.36***
-	(0.10)	(0.10)	(0.10)	(0.09)	(0.09)	(0.09)
Cancer (Yes)	−0.09	-	-	0.05	-	-
-	(0.09)	-	-	(0.12)	-	-
Diabetes (Yes)	−0.29**	−0.26**	−0.29***	0.01	−0.02	0.00
-	(0.09)	(0.09)	(0.09)	(0.10)	(0.11)	(0.10)
SmokeCigs (Former)	−0.03	−0.03	−0.04	0.06	0.06	0.06
-	(0.06)	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)
SmokeCigs (Current)	−0.26***	−0.26***	−0.26***	0.58***	0.59***	0.59***
-	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
BMI (Underweight)	−0.06	−0.06	−0.06	0.02	0.01	0.02
-	(0.22)	(0.22)	(0.22)	(0.22)	(0.22)	(0.22)
BMI (Overweight)	0.05	0.05	0.05	0.02	0.02	0.02
-	(0.06)	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)
BMI (Obese)	−0.23***	−0.22**	−0.22***	0.30***	0.30***	0.30***
-	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Age	0.02***	0.02***	0.02***	−0.04***	−0.04***	−0.04***
-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Gender (Male)	0.44***	0.45***	0.44***	−0.33***	−0.34***	−0.34***
-	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)
Race (Black)	0.04	0.04	0.04	0.36***	0.36***	0.36***
-	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Education (College or above)	0.63***	0.62***	0.63***	−0.36***	−0.36***	−0.36***
CHF (Yes)	(0.06)	(0.06) −0.32*	(0.06)	(0.08)	(0.08) 0.41*	(0.08)
-	-	(0.14)	-	-	(0.17)	-
CHD (Yes)	-	-	0.07	-	-	0.15
-	-	-	(0.12)	-	-	(0.17)
AIC	9253.59	9249.48	9254.32	8223.05	8217.89	8222.50
BIC	9338.09	9333.98	9338.82	8307.55	8302.39	8307.00
Log Likelihood	-4614.80	-4612.74	-4615.16	-4099.53	-4096.94	-4099.25
Deviance	9229.59	9225.48	9230.32	8199.05	8193.89	8198.50
Num. obs.	8448	8448	8448	8448	8448	8448

Note: Models 1, 2, and 3 are run for the high physical activity group with PA time at the 75th percentile and up, whereas Models 4, 5, and 6 are run for low physical activity with PA time at the 25th percentile or less.
***$p < 0.001$, **$p < 0.01$, *$p < 0.05$

3.3. Mortality

This study next identifies the association between objectively measured physical activity intensity and several prevalent diseases and mortality separately. As shown in Table **3**, a set of logistic regressions are estimated for the 10-year mortality rate. The specific research question that is addressed in this section is, "who is more likely to be deceased after ten years, based on the physical activity level?" To answer this question, nine different models are identified to individually isolate cancer, chronic heart failure, and chronic heart disease, controlling the PA levels separately in each cohort. Although cancer, CHF, and CHD are all significantly and positively associated with mortality, the mortality rate decreases with high PA levels.

The results document that participants who reported a high level of physical activity are significantly less likely to die within the ten-year follow-up period for all types of models controlling cancer and heart diseases (both for CHD and CHF). For example, after controlling the cancer patients in models 1 through 3, the low physically active group has a 55% greater relative mortality risk within ten years (*i.e.*, the odds ratio is exp (0.44) =1.55. In comparison, the mortality risk of high physical active individuals is decreased by almost 20% (*i.e.*, the odds ratio is exp (−0.23) = 0.7945). Additionally, similar magnitudes for heart disease are discovered.

Comparing samples who smoke currently and formerly, the models show a higher and more significant mortality rate for currently smoking individuals. In model 1, the odds value of mortality for current smokers is 2.07 (*i.e.*, exp (0.73) times higher than nonsmokers, while the odds value of mortality for the former smoker is 1.22 (*i.e.*, exp (0.20) times higher than

nonsmokers. Although it is good to quit smoking, it is better not to smoke at all to reduce smoke-related mortality. Although the difference is insignificant for cancer and CHF groups, for the CHF group, the result showed that those categorized under the obese groups are more likely to be alive during the follow-up period (10 years) if they have high PA. Obesity and the risk of heart failure are documented in the literature [17]. Therefore, it is unsurprising that obese individuals with high physical activity levels reduce their risk of congestive heart failure (*i.e.*, CHF).

The results also indicate the association between gender and mortality. Although it is documented that the mortality risk of males is higher than women in each disease category, males with diabetes and cancer with a high level of physical activity are around 3.7% less likely to die compared to women (*i.e.*, exp (0.62) −exp (0.6)=0.0369) with similar health problems [18], confirming the benefit of the high physical activities.

Table 3. Logistic Regression with dependent variables of deceased as of 2015.

-	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
(Intercept)	−7.64***	−7.45***	−7.44***	−7.60***	−7.41***	−7.40***	−7.64***	−7.47***	−7.44***
-	(0.22)	(0.22)	(0.21)	(0.22)	(0.22)	(0.21)	(0.22)	(0.22)	(0.21)
Cancer (Yes)	0.49***	0.49***	0.49***	-	-	-	-	-	-
-	(0.10)	(0.10)	(0.10)	-	-	-	-	-	-
Diabetes (Yes)	0.62***	0.62***	0.60***	0.50***	0.50***	0.49***	0.58***	0.58***	0.56***
-	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Smoke Cigs (Former)	0.20*	0.20*	0.20*	0.21*	0.20*	0.20*	0.20*	0.20*	0.20*
-	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Smoke Cigs (Current)	0.73***	0.76***	0.75***	0.71***	0.74***	0.73***	0.72***	0.76***	0.75***
-	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
BMI (Underweight)	0.80**	0.79**	0.80**	0.84**	0.84**	0.84**	0.84**	0.84**	0.84**
-	(0.31)	(0.31)	(0.31)	(0.30)	(0.30)	(0.30)	(0.30)	(0.30)	(0.30)
BMI (Overweight)	−0.18*	−0.19*	−0.18*	−0.21*	−0.21*	−0.21*	−0.20*	−0.20*	−0.20*
-	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
BMI (Obese)	−0.23*	−0.22*	−0.23*	−0.29**	−0.28**	−0.29**	−0.26**	−0.24**	−0.25**
-	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Age	0.09***	0.09***	0.09***	0.09***	0.09***	0.09***	0.10***	0.09***	0.09***
-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Gender (Male)	0.47***	0.45***	0.48***	0.44***	0.42***	0.44***	0.43***	0.41***	0.44***
-	(0.07)	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.07)	(0.08)
Race (Black)	−0.02	−0.01	−0.01	−0.04	−0.02	−0.03	−0.02	−0.00	−0.01
-	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Education College or above	−0.70***	−0.71***	−0.68***	−0.63***	−0.65***	−0.62***	−0.65***	−0.67***	−0.64***
-	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)
PA Low	0.44***	-	-	0.42***	-	-	0.43***	-	-
-	(0.10)	-	-	(0.10)	-	-	(0.10)	-	-
PA Medium	-	−0.02	-	-	−0.02	-	-	−0.01	-
-	-	(0.07)	-	-	(0.07)	-	-	(0.07)	-
PA High	-	-	−0.23**	-	-	−0.22**	-	-	−0.24**
-	-	-	(0.08)	-	-	(0.08)	-	-	(0.08)
CHF (Yes)	-	-	-	1.18***	1.18***	1.18***	-	-	-
-	-	-	-	(0.15)	(0.15)	(0.15)	-	-	-
CHD (Yes)	-	-	-	-	-	-	0.53***	0.53***	0.54***
-	-	-	-	-	-	-	(0.12)	(0.12)	(0.12)
AIC	5067.05	5085.25	5076.7	5024.62	5041.21	5033.6	5074.44	5092.06	5082.75
BIC	5158.6	5176.8	5168.24	5116.16	5132.75	5125.14	5165.98	5183.61	5174.29
Log Likelihood	-2520.53	-2529.63	-2525.35	-2499.31	-2507.61	-2503.8	-2524.22	-2533.03	-2528.37
Deviance	5041.05	5059.25	5050.7	4998.62	5015.21	5007.6	5048.44	5066.06	5056.75
Num. obs.	8448	8448	8448	8448	8448	8448	8448	8448	8448

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Models 1-3 only include cancer patients, Models 4-6 with CHF, and Models 7-9 with CHD. Within each model, we only control PA low, medium, and high levels separately.

4. DISCUSSION

The aims of the study are first to present the association between physical activity and disease and, secondly, to research the impact of physical activity on mortality after controlling cancer and heart diseases using a comprehensive sample of NHANES. A summary of the main findings is presented as follows. First, it is proven by the regression analysis that high physically active individuals significantly decrease diabetes, chronic heart failure, and obesity. Second, individuals currently smoking are found to be associated with fewer PA times, mainly since they are 23% times (*i.e.*, the odds ratio is exp (−0.26) = 0.77, see Table **2**) less likely to do any high physical activity compared to nonsmokers. Third, the mortality rates of highly physically active individuals are almost 20% less compared to non-physically active individuals.

5. STRENGTHS AND LIMITATIONS

Due to data availability, this paper adds to the growing literature documenting the impact of physical activity on mortality and associations between physical activity and disease. Particularly, this study included socioeconomic information of individuals into the models to provide a more thorough picture of the PA impact. There has been a growing number of studies indicating that individuals with higher socioeconomic status are associated with more physical activity compared to the lower socioeconomic status group [19]. If this association is significant, this raises the endogeneity concern between the level of physical activity and the occurrence of disease or mortality rates, which might produce biased outputs. One of the statistical solutions that can address endogeneity in this context is to control the socioeconomic status of the subjects (*i.e.*, survey participants) in the model. Through this approach, this study presents the unbiased impact of physical activity on health outcomes.

However, this study still has some limitations. Even though the data used in this paper is a comprehensive and representative sample of individuals with demographic, disease, and physical activity, it does not solve all covariates that are interplayed with the results captured in the study. For example, previous studies document that physical activity is associated with improved eating habits [20 - 22]. Therefore, nutrition, healthy eating habits, and physical activity may negatively contribute to mortality, including its association with diseases, which the current data does not cover to test. This is a future research agenda.

CONCLUSION

The leading causes of mortality in the United States are cardiovascular diseases (CHF, CHD) at 23% and cancer at 21%, accounting for 44% of total deaths. These diseases cost the healthcare system $715 billion and cause a loss of productivity on the job [1, 23]. However, increasing awareness of physical activity on some diseases and mortality can reduce the high cost. This study shed some light on the impact of physical activity and mortality while documenting its association with diseases and found that (1) physical activity is positively associated with education level and negatively associated with obesity, diabetes, chronic heart failure, and smoking. (2) physical activity is associated with a lower mortality rate after controlling the current health status and smoking habits.

In this regard, this study contributes to the public health literature on physical activities by incorporating the individual characteristics of the national health survey participants into the analyses and linking those characteristics to the relationship between physical activity and different types of major diseases, which was less addressed in previous studies.

LIST OF ABBREVIATIONS

CDC	=	The Centers for Disease Control and Prevention
PA	=	Physical Activity
NHANES	=	National Health and Nutrition Examination Survey
ERB	=	Ethics Review Board
CHD	=	Coronary Heart Disease
CHF	=	Congestive Heart Failure

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The data source of this study (the NHANES Survey) was approved by the U.S. CDC as well as the National Center for Health Statistics (NCHS) Ethics Review Board (ERB). The informed consent for the survey is also publicly available on the NHANES website: https://www.cdc.gov/nchs/nhanes/about_nhanes.htm.

HUMAN AND ANIMAL RIGHTS

No animals were used in this research. All procedures performed in studies involving human participants were in accordance with the ethical standards of institutional and/or research committees and with the 1975 Declaration of Helsinki, as revised in 2013.

CONSENT FOR PUBLICATION

The informed consent for the survey is also publicly available on the NHANES website: https://www.cdc.gov/nchs/nhanes/about_nhanes.html.

AVAILABILITY OF DATA AND MATERIALS

Datasets analyzed during the current study are publicly available at https://github.com/andrew-leroux/rnhanesdata.

FUNDING

None.

CONFLICT OF INTEREST

I declare no conflict of interest.

ACKNOWLEDGEMENTS

I would like to thank my mentor Dr. Ciprian Crainiceanu for his guidance and comments, and Dr. Eun Kyu Lee for his feedback.

REFERENCES

[1] Health and Economic Costs of Chronic Diseases 2023. Available From: https://www.cdc.gov/chronicdisease/about/costs/index.htm

[2] Benefits of Physical Activity 2023. Available From: https://www.cdc.gov/physicalactivity/basics/pa-health/index.htm

[3] Physical Activity Guidelines & Recommendations 2023. Available

From:
https://www.cdc.gov/physicalactivity/resources/recommendations.html

[4] Di J, Leroux J, Urbanek J, *et al.* Patterns of sedentary and active time accumulation are associated with mortality in US adults: The NHANES study. BioRxiv 2017. Preprint

[5] Schmid D, Ricci C, Baumeister S, Leitzmann MF. Replacing sedentary time with physical activity in relation to mortality. Med Sci Sports Exerc 2016; 48(7): 1312-9.
[http://dx.doi.org/10.1249/MSS.0000000000000913] [PMID: 26918559]

[6] Dibben GO, Gandhi MM, Taylor RS, *et al.* Physical activity assessment by accelerometry in people with heart failure. BMC Sports Sci Med Rehabil 2020; 12(1): 47.
[http://dx.doi.org/10.1186/s13102-020-00196-7] [PMID: 32817798]

[7] Lloyd-Jones DM. Cardiovascular risk prediction: Basic concepts, current status, and future directions. Circulation 2010; 121(15): 1768-77.
[http://dx.doi.org/10.1161/CIRCULATIONAHA.109.849166] [PMID: 20404268]

[8] Loprinzi PD, Addoh O. The effects of free-living physical activity on mortality after coronary artery disease diagnosis. Clin Cardiol 2016; 39(3): 165-9.
[http://dx.doi.org/10.1002/clc.22508] [PMID: 26748944]

[9] Cancer Causes and Prevention – Obesity. 2015. Available From: https://www.cancer.gov/about-cancer/causes-prevention/risk/obesity

[10] Thune I, Furberg AS. Physical activity and cancer risk: Dose-response and cancer, all sites and site-specific. Med Sci Sports Exerc 2001; 33(6)(Suppl.): S530-50.
[http://dx.doi.org/10.1097/00005768-200106001-00025] [PMID: 11427781]

[11] Sigal RJ, Armstrong MJ, Bacon SL, *et al.* Physical Activity and Diabetes. Can J Diabetes 2018; 42(1)(Suppl. 1): S54-63.
[http://dx.doi.org/10.1016/j.jcjd.2017.10.008] [PMID: 29650112]

[12] Leroux A, Di J, Smirnova E, *et al.* Organizing and analyzing the activity data in NHANES. Stat Biosci 2019; 11(2): 262-87.
[http://dx.doi.org/10.1007/s12561-018-09229-9] [PMID: 32047572]

[13] Schuna JM Jr, Johnson WD, Tudor-Locke C. Adult self-reported and objectively monitored physical activity and sedentary behavior: NHANES 2005–2006. Int J Behav Nutr Phys Act 2013; 10(1): 126.
[http://dx.doi.org/10.1186/1479-5868-10-126] [PMID: 24215625]

[14] Smirnova E, Leroux A, Cao Q, *et al.* The predictive performance of objective measures of physical activity derived from accelerometry data for 5-year all-cause mortality in older adults: National Health and Nutritional Examination Survey 2003–2006. J Gerontol A Biol Sci Med Sci 2020; 75(9): 1779-85.
[http://dx.doi.org/10.1093/gerona/glz193] [PMID: 31504213]

[15] R-package: rnhanesdata. 2020. Available From: https://github.com/andrew-leroux/rnhanesdata

[16] National Death Index 2021. Available From: https://www.cdc.gov/nchs/ndi/index.htm

[17] Kenchaiah S, Evans JC, Levy D, *et al.* Obesity and the risk of heart failure. N Engl J Med 2002; 347(5): 305-13.
[http://dx.doi.org/10.1056/NEJMoa020245] [PMID: 12151467]

[18] Jonker JT, De Laet C, Franco OH, Peeters A, Mackenbach J, Nusselder WJ. Physical activity and life expectancy with and without diabetes: Life table analysis of the Framingham Heart Study. Diabetes Care 2006; 29(1): 38-43.
[http://dx.doi.org/10.2337/diacare.29.01.06.dc05-0985] [PMID: 16373893]

[19] Stalsberg R, Pedersen A. Are differences in physical activity across socioeconomic groups associated with choice of physical activity variables to report? Int J Environ Res Public Health 2018; 15(5): 922-44.
[http://dx.doi.org/10.3390/ijerph15050922] [PMID: 29734745]

[20] Agyemang K, Banstola A, Pokhrel S, Anokye N. Determinants of physical activity and dietary habits among adults in Ghana: A cross-sectional study. Int J Environ Res Public Health 2022; 19(8): 4671-82.
[http://dx.doi.org/10.3390/ijerph19084671] [PMID: 35457539]

[21] Christofaro DGD, Werneck AO, Tebar WR, *et al.* Physical activity is associated with improved eating habits during the COVID-19 pandemic. Front Psychol 2021; 12: 664568.
[http://dx.doi.org/10.3389/fpsyg.2021.664568] [PMID: 33912120]

[22] Sjöblom L, Bonn SE, Alexandrou C, Dahlgren A, Eke H, Trolle Lagerros Y. Dietary habits after a physical activity mHealth intervention: A randomized controlled trial. BMC Nutr 2023; 9(1): 23.
[http://dx.doi.org/10.1186/s40795-023-00682-4] [PMID: 36732788]

[23] Kochanek KD, Xu J, Arias E. Mortality in the United States, 2019. NCHS Data Brief 2020; (395): 1-8.
[PMID: 33395387]