Physical activity status and cardiovascular health in U.S. using quantile regression analysis NHANES 2017-2020

Tianze Tang*

School of Life Science and Technology, Xi'an Jiaotong University, Xi'an, China

*Corresponding author: a1291346840@stu.xjtu.edu.cn

Keywords: Physical activity, cardiovascular disease, quantile regression, metabolic equivalent.

Abstract: Modern life has led to a decline in physical activity, which may be linked to the risk of cardiovascular disease. The study was designed to find out whether there is an obvious trend between physical activity and cardiovascular health, and whether such relationship varied by gender or race. Triglycerides, high sensitivity C-reactive protein, and HDL have been used as indicators of cardiovascular health. Physical activity was assessed using metabolic equivalent (MET) calculated from the amount of exercise and sedentary time. All data was collected from NHANES '2017-2020 survey. Quantile regression was used in this study.

1. Introduction

There is a large body of research showing that physical activity can significantly improve an individual's health and mortality [1]. Physical inactivity is a major risk factor for cardiovascular disease [2][3]. Regular exercise has been shown to be a good way to control and prevent obesity and diabetes [4], which can further lead to cardiovascular disease [5]. In addition to the exercise itself, too much sedentary time can damage cardiovascular health [6].

National Health and Nutrition Examination Survey (NHANES) is a cross-sectional study conducted by the CDC to assess the health of Americans. The study examined the effects of physical activity and sedentary time on markers of cardiovascular disease in US adults based on 2017-2020 data from NHANES. Since the data distribution is not normal, quantile regression would obtain more convincing results [7].

Studies have shown that high-density lipoprotein (HDL) content is negatively correlated with cardiovascular disease [8-9]. They have a strong anti-atherosclerosis effect [10-12] . Triglycerides have been shown to be a cardiovascular health risk factor independent of HDL [13]. Similar to triglycerides, C-reactive protein has been shown to be a marker that predicts cardiovascular disease [14]. Processes such as atherosclerosis are also thought to be inflammatory. So there have been studies suggesting that C-reactive protein is a predictor of cardiovascular events that seems to be stronger than LDL [14].

Based on the above information, HDL, triglyceride and C-reactive protein were used as markers of cardiovascular events in this study, and quantile regression and Bootstrap test were adopted to test the model. Series of valuable results were finally obtained as illustrated in the result and discussion sections.

2. Methods

This study is based on NHANES pre-epidemic data from 2017 to 2020. Detailed information about the program can be found on the website: https://www.cdc.gov/nchs/nhanes/index.htm. The NHANES program suspended field operations in March 2020 due to the coronavirus disease 2019 (COVID-19) pandemic. As a result, data collection for the NHANES 2019-2020 cycle was not completed and the collected data are not nationally representative. Therefore, data collected from 2019 to March 2020 were combined with data from the NHANES 2017-2018 cycle to form a nationally representative sample of NHANES 2017-March 2020 pre-pandemic data.

2.1 Study sample

The initial sample size was 15,560. After removing some samples lacking the necessary variables for the study and adjusting the covariates, the final study sample size was 3,200. There were 1,554 males and 1,646 females. The age range from 18 to 80 years presents a fairly uniform distribution. There is also ample data for different ethnic groups.

2.2 Data outcomes

Triglyceride data were measured by photometric method using Cobas 6000 Chemistry Analyzer. HDL protein was also measured by Cobas 6000 Chemistry Analyzer and total cholesterol was obtained by Roche Cobas 6000. High sensitivity C-reactive protein was measured using immunoturbidimetry using Roche Cobas 6000 (C501 Module). Related laboratory manuals can be found on the NHANES website.

The data used to measure physical activity were obtained from questionnaires. The questionnaire was asked, in the home, by trained interviewers using Computer-Assisted Personal Interview (CAPI) system. A proxy provided information for survey participants who could not answer the questions themselves. The respondent selected the language of interview (English or Spanish) or requested that an interpreter be used. Hand cards, showing response choices or information that survey participants needed to answer the questions, were used for some questions. The hand cards were printed in English, Spanish, Mandarin Chinese (both traditional and simplified), Korean, and Vietnamese. The interviewer directed the respondent to the appropriate hand card during the interview. When necessary, the interviewer further assisted the respondent by reading the response choices listed on the hand cards. The corresponding questionnaire can also be found on the NHANES website.

In this study, these data were calculated as the metabolic equivalent (MET) recommended by NHANES, and the sedentary time obtained from the questionnaire was used as the activity index of the samples.

All outcomes are treated as continuous variables.

2.3 Covariates

Covariates include race, ethnicity, age and sex. Racial and ethnic data were obtained through questionnaires for Mexican Americans, other Hispanics, non-Hispanic whites and non-Hispanic blacks. There were 1,067 non-Hispanic whites, 729 non-Hispanic blacks, 302 Hispanics and 418 Mexicans. Age and sex were also obtained through questionnaires. A pseudo - R - square analysis was performed for age.

2.4 Statistical Analysis

After a simple preliminary analysis, HDL, triglyceride and high-sensitivity C-reactive protein were selected as dependent variables, while sedentary time and exercise metabolic equivalent calculated according to the official recommendations of NHANES were used as independent variables. The overall sample and subsamples by race, age, and sex were analyzed.

Because the sample data does not conform to normal distribution (even after attempting mathematical transformation), and also contains some outliers. Therefore, quantile regression, which is insensitive to extreme data and does not need normal distribution, was selected to process the data. In addition, interior point method is selected for fitting, which is more suitable for large sample $n \geq 50$[7]. During the processing, five different loci of 0.1, 0.25, 0.5, 0.75 and 0.9 were selected, and bootstrap test was performed on the obtained model, which also does not rely on normal distribution. The standard for significance was $P < 0.05$.

3. Results

After analyzing the overall sample, HDL levels decreased with sedentary time ($P<0.001$), triglycerides increased with sedentary time and decreased with MET ($P<0.001$), and C-reactive protein and triglycerides showed similar results, as shown in Table 1. This is logically consistent with existing research showing that inactivity, an unhealthy lifestyle, leads to a decrease in beneficial HDL and an increase in harmful triglycerides and C-reactive protein.

Analysis of different ethnic groups showed that the trend was almost completely insignificant among Hispanics and Mexicans, somewhat significant among non-Hispanic blacks, and quite significant among non-Hispanic whites ($P<0.001$). So the next control for covariates is going to be directly for non-Hispanic whites.

The gender difference was not significant, and the conclusion was basically consistent with the overall analysis. It should be noted, however, that men and women had the same distribution of sedentary time, but men had significantly more MET overall than women (For men, the mean MET is 1,380 and for women, 761.8).

Analysis of age showed that the association between cardiovascular health markers and physical activity was not significant in younger adults, while a stronger association was observed in older adults (≥40 years). Note that there is a relationship between cardiovascular health and age itself, with aging leading to a decline in a range of cardiovascular physiological functions [15]. This raises the question of whether the rise in markers of cardiovascular health is simply due to age and whether the observed significance come from the fact that physical activity and sedentary time also change with age? Therefore, the pseudo R square index is calculated in this study to characterize the interpretation rate of the model. The addition of MET and sedentary time, or MET and sedentary time alone, yielded higher pseudo-R-squared outcomes than age alone. There are also studies showing that increased physical activity can extend life expectancy and reduce the risk of cardiovascular disease [16]. This suggests that age is not the only determinant, and that exercise still plays a role.

Table 1. Effects of MET and sedentary time on HDL, triglyceride, and CRP.

Subgroups	**Indicators**	**p-value Quantile0.1**		**p-value Quantile0.25**		**p-value Quantile0.5**		**p-value Quantile0.75**		**p-value Quantile0.9**	
		MET	**Sedentary Time**	**MET**	**Sedentary Time**	**MET**	**Sedentary Time**	**MET**	**Sedentary Time**	**MET**	**Sedentary Time**
Overall(n=3067)	HDL	0.534	0.005***(-)	0.454	0.004***(-)	0.164	0.010**(-)	0.002***(-)	0.000***(-)	0.000***(-)	0.002**(-)
	TRI	0.000***(-)	0.458	0.001***(-)	0.003**(+)	0.000***(-)	0.021*(+)	0.000***(-)	0.008***(+)	0.026*(-)	0.544
	CRP	0.247	0.499	0.016*(-)	0.007**(+)	0.266	0.000***(+)	0.147	0.001***(+)	0.037*(-)	0.002**(+)
Non-Hispanic whites(n=1067)	HDL	0.403	0.392	0.041*(-)	0.135	0.013*(-)	0.012*(-)	0.011*(-)	0.032*(-)	0.000***(-)	0.001***(-)
	TRI	0.447	0.592	0.002**(-)	0.005**(+)	0.000***(-)	0.089	0.005**(-)	0.014*(+)	0.032*(-)	0.450
	CRP	0.398	0.728	0.317	0.273	0.356	0.001***(+)	0.077	0.028*(+)	0.013*(-)	0.060
Non-Hispanic black(n=729)	HDL	0.212	0.111	0.363	0.035*(-)	0.625	0.116	0.186	0.001***(-)	0.008**(-)	0.141
	TRI	0.015*(-)	0.391	0.007**(-)	0.040*(+)	0.284	0.118	0.311	0.401	0.766	0.176
	CRP	0.135	0.013*(+)	0.055	0.035*(+)	0.441	0.006**(+)	0.060	0.099	0.753	0.421
Hispanic(n=302)	HDL	0.731	0.574	0.332	0.580	0.452	0.666	0.639	0.610	0.801	0.326
	TRI	0.210	0.250	0.250	0.936	0.122	0.302	0.603	0.335	0.777	0.452
	CRP	0.192	0.299	0.022	0.743	0.150	0.871	0.291	0.442	0.867	0.615
Mexican(n=418)	HDL	0.015**（+）	0.610	0.177	0.365	0.708	0.934	0.712	0.103	0.122	0.440
	TRI	0.083	0.179	0.264	0.688	0.001**(-)	0.980	0.031*(-)	0.560	0.279	0.419
	CRP	0.233	0.715	0.857	0.406	0.025*(-)	0.364	0.156	0.546	0.134	0.007**(+)
Over 40 years of age in significant races(n=757)	HDL	0.943	0.211	0.071	0.403	0.024*(-)	0.114	0.034*(-)	0.501	0.001**(-)	0.008**(-)
	TRI	0.067	0.380	0.018*(-)	0.039*(+)	0.452	0.140	0.692	0.112	0.114	0.796
	CRP	0.457	0.138	0.905	0.234	0.405	0.001**(+)	0.078	0.036*(+)	0.023*(-)	0.033*(+)
Less than 40 years of age in significant races(n=310)	HDL	0.184	0.680	0.449	0.684	0.555	0.162	0.225	0.250	0.005**(-)	0.002**(-)
	TRI	0.273	0.227	0.399	0.227	0.251	0.116	0.071	0.174	0.042*(-)	0.139
	CRP	0.063	0.323	0.620	0.878	0.440	0.432	0.607	0.369	0.213	0.764

Males of significant races (n=546)	HDL	0.356	0.306	0.579	0.223	0.472	0.083	0.718	0.006**(-)	0.505	0.156
	TRI	0.585	0.111	0.019*(-)	0.030*(+)	0.196	0.006**(+)	0.072	0.360	0.004**(-)	0.867
	CRP	0.917	0.714	0.877	0.187	0.972	0.052	0.814	0.042*(+)	0.468	0.012*(+)
Females of significant races (n=521)	HDL	0.284	0.375	0.220	0.571	0.792	0.834	0.099	0.099	0.071	0.009**(-)
	TRI	0.336	0.193	0.003**(-)	0.662	0.000***(-)	0.836	0.000***(-)	0.959	0.396	0.983
	CRP	0.420	0.997	0.880	0.826	0.899	0.049*(+)	0.218	0.323	0.268	0.151

"Significant races" mean Non-Hispanic whites. *P＜0.05; **P<0.01; ***P<0.001

4. Discussion

The results showed that increasing physical activity and reducing sedentary time improved levels of markers of cardiovascular status in non-Hispanic whites and blacks. Among HDL, triglyceride and C-reactive protein, the change of triglyceride was the most significant. The correlation was not significant among young people. In the absence of follow-up studies for each sample, it remains unclear whether exercise habits in youth have a similar effect on cardiovascular health in old age. Exercise habits in youth have no effect on health in old age, a study suggests [17]. However, since there is evidence that cardiovascular disease in the United States is decreasing at a younger age, health advise is thus given as people should also develop good exercise habits at young age [18].

There are still some problems with the study. First, using MET to assess exercise is a feasible but not entirely accurate option. The effect of exercise depends on the type and frequency of exercise, and the MET can only provide a rough estimate. In HDL analysis, higher MET led to a decrease in HDL. Second, there was a negative correlation between MET and sedentary time. If only one marker is negatively correlated with sedentary time, it does not necessarily mean there is no positive association with MET. Due to the large sample size, MET and sedentary time were not modeled separately.

5. Conclusions

This study provides a general conclusion that more sedentary time and less MET have an adverse linear association with markers of cardiovascular disease, and that the association is more significant in non-Hispanic whites and older adults. The association was not significant among non-Hispanic blacks, Mexicans, and Hispanics. And the correlation held true for both men and women, even though their physical activity levels were different. Since the use of MET as an indicator of physical activity may not be comprehensive and does not control for more covariates, a larger P-value does not necessarily prove that there is no correlation. In addition, the intrinsic correlation between MET and sedentary behavior may cause the two to mask each other's effects in the model. So the fact that some of the items in this study were associated with sedentary behavior and MET doesn't necessarily mean that one of them wasn't associated with the marker. Because MET together with sedentary are logically simple and complementary. Since the sample size is large, the two were not modeled separately, but the above issues still need to be noted. To obtain more accurate correlation model or individual analysis, it is necessary to further refine the measurement method of exercise amount, control more covariables and carry out follow-up surveys.

References

[1] Kokkinos P. Physical activity, health benefits, and mortality risk [J]. International Scholarly Research Notices, 2012, 2012.

[2] Luepker R V, Johnson S B, Breslow L, et al. Physical activity and cardiovascular health[J]. Journal of the American Medical Association, 1996, 276(3): 241-246.

[3] Lavie C J, Ozemek C, Carbone S, et al. Sedentary behavior, exercise, and cardiovascular health[J]. Circulation research, 2019, 124(5): 799-815.

[4] Klein S, Sheard N F, Pi-Sunyer X, et al. Weight management through lifestyle modification for the prevention and management of type 2 diabetes: rationale and strategies. A statement of the American Diabetes Association, the North American Association for the Study of Obesity, and the American Society for Clinical Nutrition[J]. The American journal of clinical nutrition, 2004, 80(2): 257-263.

[5] Boyle J, Thompson T, Gregg E, et al. D., 2010. Projection of the year 2050 burden of diabetes in the us adult population: Dynamic modeling of incidence, mortality, and prediabetes prevalence [J]. Population Health Metrics, 8: 29.

[6] Healy G N, Matthews C E, Dunstan D W, et al. Sedentary time and cardio-metabolic biomarkers in US adults: NHANES 2003–06[J]. European heart journal, 2011, 32(5): 590-597.

[7] Koenker R, Hallock K F. Quantile regression [J]. Journal of economic perspectives, 2001, 15(4): 143-156.

[8] Lewis G F, Rader D J. New insights into the regulation of HDL metabolism and reverse cholesterol transport [J]. Circulation research, 2005, 96(12): 1221-1232.

[9] Santos-Gallego, C.G.; Badimon, J.J.; Rosenson, R.S. Beginning to understand high-density lipoproteins. Endocrinol. Metab. Clin. N. Am. 2014, 43, 913–947.

[10] Kontush A, Chapman M J. Functionally defective high-density lipoprotein: a new therapeutic target at the crossroads of dyslipidemia, inflammation, and atherosclerosis [J]. Pharmacological reviews, 2006, 58(3): 342-374.

[11] Rosenson R S, Brewer H B, Ansell B J, et al. Dysfunctional HDL and atherosclerotic cardiovascular disease[J]. Nature reviews cardiology, 2016, 13(1): 48-60.

[12] Rader D J. Spotlight on HDL biology: new insights in metabolism, function, and translation [J]. Cardiovascular Research, 2014, 103(3): 337-340.

[13] Toth P P. Triglyceride-rich lipoproteins as a causal factor for cardiovascular disease [J]. Vascular health and risk management, 2016, 12: 171.

[14] Ridker P M. Clinical application of C-reactive protein for cardiovascular disease detection and prevention [J]. Circulation, 2003, 107(3): 363-369.

[15] Lakatta E G. Age-associated cardiovascular changes in health: impact on cardiovascular disease in older persons [J]. Heart failure reviews, 2002, 7(1): 29-49.

[16] Franco O H, de Laet C, Peeters A, et al. Effects of physical activity on life expectancy with cardiovascular disease [J]. Archives of internal medicine, 2005, 165(20): 2355-2360.

[17] Brill P A, Burkhalter H E, Kohl H W, et al. The impact of previous athleticism on exercise habits, physical fitness, and coronary heart disease risk factors in middle-aged men [J]. Research quarterly for exercise and sport, 1989, 60(3): 209-215.

[18] Page R L, II V G, Nair K. A call to action: Responding to the future forecasting of cardiovascular disease in America [J]. American Health & Drug Benefits, 2011, 4(5): 280.