The Impact of Physical Activity Levels on the Prevalence of Diabetes in US Adults: Findings from NHANES 2017–2018 Survey data.

Okiemute Omoru, Hridya Divakaran, Nithya Reddy, Bhavya Karamudi

ABSTRACT

Introduction: Diabetes is a metabolic illness that involves hyperglycemia because of a lack of insulin production or insulin activity. It is strongly influenced by lifestyle-related factors such as physical activity levels. We aimed to study the effect of physical activity levels on diabetes status.
Methods: Demographic characteristics were compared between diabetic and nondiabetic participants. Multiple logistic regression models were used to assess the link between different physical activity levels and diabetes.
Results: Participants with moderate to high physical activity levels had a higher prevalence of diabetes. Also, the multivariate logistic regression result showed that compared to those with moderate to a high level of physical activity, subjects with no reported physical activity levels had higher odds of being diabetic.
Conclusions: A physically active lifestyle may decrease the risk of diabetes. Further studies with larger sample size, also involving other countries, are needed to verify and strengthen the association between physical activity and the risk of diabetes. It is critical to improve the awareness of physical activity to prevent diabetes.

Keywords: Association, Diabetes, Physical activity, NHANES

INTRODUCTION

Diabetes is a metabolic illness that involves hyperglycemia because of a lack of insulin production or insulin activity. Many factors contribute to the development of diabetes. Insulin shortage is caused by autoimmune death of pancreatic Beta islet cells or abnormalities that increase glycemic control. Difficulties in the digestion of carbohydrates, lipids, and proteins indicate diabetes. Human insulin deficit is caused by either a lack of hormone insulin or a loss in the body's ability to respond to insulin. In many patients, both insulin release and insulin action abnormalities coincide, making it difficult to determine the primary factor of diabetes[1].
Diabetes is one of the most prevalent metabolic illnesses, and it develops because of interactions between genetic, environmental, and lifestyle variables. Although diabetes can be hereditary, it is also significantly impacted by lifestyle variables like nutrition and physical exercise. Diabetes prevention by suitable lifestyle adjustments is consequently a critical public health concern. The rising diabetes incidence emphasizes the need for research on lifestyle guidelines and better therapies for diabetes risk factors impacting those with a high incidence of diabetes [2].
One of the most important ways to avoid diabetes-related medical conditions is to engage in regular physical activity. To reap the long-term advantages of cardiovascular health and improved insulin signaling, it is necessary to engage in regular physical activity of various sorts. Individuals with type 2 diabetes who intend to live a long and healthy life must increase their physical activity via participation in an exercise program or other ways [3]. Exercise is safe for most people with type 2 diabetes if they follow a few safety considerations.

This paper aims to study the effect of physical activity levels on diabetes status. To demonstrate this, we performed a comprehensive calculation of physical activity of different intensities and durations and analyzed the relationship between physical activity levels and the prevalence of diabetes. The null hypothesis for this paper is that there is no relationship between physical activity level and the prevalence of diabetes. The alternate hypothesis is that there is a relationship between physical activity level and the prevalence of diabetes.

METHODOLOGY

Study Population and data collection

We performed a cross-sectional study, using data from the 2017-2018 National Health and Nutrition Survey data (NHANES). NHANES is an annual survey conducted by the National Center for Health Statistics (NCHS) since 1960. It combines physical examination, laboratory, and interview questionnaire information related to demographic, socioeconomic, dietary, and health-related from civilians and non-institutionalized US adults and children, to assess their health and nutritional status. All participants in the survey are recruited voluntarily. The NHANES sampling is a modified random sampling method that involves a complex, multistage, probability sampling design that oversamples certain subgroups such as African Americans and Hispanics to increase the certainty of the health status estimates of the subgroups. Each participant is given a sample weight, which signifies the number of people in the population represented by a study participant. Using sample weight is recommended while analyzing the survey data to avoid bias and overstated p values. More on the Survey design is found on the CDC NHANES website [10]. A total of 16,211 individuals participated in the 2017-2018 NHANES survey. We performed the analysis on study participants aged 20 years and above, with no missing information in our variables of interest. 4861 participants were included in our analysis.

Outcome / Dependent variable

Diabetes was defined using NHANES laboratory data and the American Diabetes Association guidelines. Glycated HbA1c was the diagnostic criterion; participants with HbA1c values < 6.5 % and $\geq$ 6.5 % were categorized as non-diabetics and diabetics, respectively.

Exposure / Independent variables

The primary exposure variable in our analyses was physical activity level. The NHANES participants completed the Global Physical Activity Questionnaire (GPAQ) [10]. The questionnaire gathered information on work and recreational based activities physical activities. Physical activity was categorized into five types: walking or cycling, moderate work-related activity, vigorous work-related activity, moderate recreational activity, and vigorous recreational activity. The duration and intensity of physical activity were estimated by multiplying the number of days a specific physical activity was performed per week (days/week) by the duration in minutes (minutes/day) to produce the Metabolic Equivalent Score (MET). The MET score was multiplied by the CDC suggested factors 8, 4, 4, 8, and 4 for vigorous recreational activity, moderate recreational activity, moderate work-related activity, vigorous work-related activity, and walking or bicycling. By doing this, we effectively quantified the physical activity levels of each participant based on the various intensities and duration of performance per week.

The following demographic characteristics: age, gender, educational level, marital status, and BMI as the covariates of our analysis. Studies have shown that these factors affect diabetes status[4]. Age was divided into four groups: $\geq$80 years, 65 – 79 years, 45-64 years, and 20 -44 years. The five race groups were White, Blacks, Hispanics, Asians, and Others/Multiracial. We categorized marital status as married, unmarried (never married), and other (widowed, separated, and divorced). BMI was calculated as body weight in kilograms divided by height in meters squared and categorized as underweight (BMI < 18.5 kg/m$^{2)}$, normal weight (18.5 – 24.99 kg/m$^{2)}$, overweight (25-29.9 kg/m^2), and obese (> 30 kg/m^2). Finally, gender was categorized as male and female.

Statistical analysis

All statistical analysis performed in this study was conducted using R version 4.1.2. We examined the population variables using t-test for continuous variables and Pearson's chi-square test for nominal categorical variables. Multivariate logistic regression was conducted to estimate the association between exposure (physical activity levels) and outcome (diabetes status). In the first model, the unadjusted association between exposure and outcome was examined. The second model was adjusted for age, gender, and race, while the third model was adjusted for age, race, gender, marital status, educational level, BMI, and marital status. We performed a subgroup analysis of the relationship between physical activity level and diabetes status. The data were stratified by gender before performing a multiple logistic regression to identify the impact of gender differences on the association between physical activity and diabetes. Odds ratio, 95% confidence intervals, and AUC were calculated to demonstrate the strength of the associations and ability of each model to accurately classify the positive and negative outcomes of all the possible cutoffs[5].
The probability value of < 0.05 was considered statistically significant.

RESULT

Demographic Characteristics of study participants

The population characteristics of the study participants are presented in Table 1. The inclusion criteria resulted in 4861 subjects from NHANES 2017–2018 being enrolled in this study, of which 47.97% were male. The percentage of subjects with diabetes was 14.15%, and the mean physical activity levels, defined as MET score in hours, in the diabetic and nondiabetic participants were 68.28 METs and 97.47 METs, respectively. Table 1 indicates that the gender distribution differed slightly between the subjects with and without diabetes. 51.8% of participants with diabetes were male. Older subjects and Blacks and Hispanics had a higher prevalence of diabetes. The diabetic participants were more likely to have a lower level of education (less than 9th grade). Married subjects had a higher prevalence of diabetes. The prevalence rates of having higher BMI and low physical activity level per week were higher among the diabetic participants.

Table 1. Population characteristics by diabetes status in NHANES 2017-2018(N=4861)

Stratified by Diabetes status

	Diabetic (%)	Nondiabetic (%)	P-value
Participants (n)	722 (14.85)	4139 (85.15)	
Age (Years)			<0.001
>= 80	69 (9.6)	267 (6.5)	
20-44	60 (8.3)	1711 (41.3)	
45-64	343 (47.5)	1465 (35.4)	
65-79	250 (34.6)	696 (16.8)	
Gender Male	374 (51.8)	1958 (47.3)	0.029
BMI Category (Kg/m^2)			<0.001
< 18.5	1 (0.1)	68 (1.6)	
>= 30	411 (56.9)	1644 (39.7)	
18.5 - 24.9	91 (12.6)	1075 (26.0)	
25 - 29.9	219 (30.3)	1352 (32.7)	
Marital Status			<0.001
Married	415 (57.5)	2068 (50.0)	
Other	238 (33.0)	1322 (31.9)	
Unmarried	69 (9.6)	749 (18.1)	
Education			<0.001
9-11th grade	90 (12.5)	457 (11.0)	
College graduate or higher	527 (73.0)	3373 (81.5)	
Less than 9th grade	105 (14.5)	309 (7.5)	
Race			0.002
Hispanic	185 (25.6)	926 (22.4)	
Other/Multiracial	38 (5.3)	206 (5.0)	
Asian	108 (15.0)	585 (14.1)	
Black	187 (25.9)	926 (22.4)	
White	204 (28.3)	1496 (36.1)	
METs score category (Physical activity levels)			<0.001
No Physical activity	255 (35.3)	1004 (24.3)	
Moderate Physical activity level	462 (64.0)	3121 (75.4)	
High Physical activity level	5 (0.7)	14 (0.3)	
Combined MET Scores in hours (mean (SD))	68.28 (294.74)	97.47 (350.26)	0.035

MET metabolic equivalents

Multivariate Logistic Regression Analysis of Physical Activity and Diabetes

The results obtained from the multiple logistic regression models are presented in Table 2. The primary outcome variable was diabetes, and the level of physical activity was the primary exposure variable after adjusting for age, race, educational level, marital status, and BMI. Among all the participants, the odds of being diabetic were significantly lower in the high physical activity level group than in the low and moderate groups, before and after adjusting for the covariates.

Table 2. ORs and 95% CI for the relationship between diabetes status and physical activity in NHANES 2017 – 2018.

Physical activity		**Study participants**			
		OR	95% CI	p-value	AUC
Model 1	No physical activity	3.93	3.43 - 4.53	< 2e-16 ***	0.5573
	Moderate Physical activity level	1.72	1.45 - 2.02	3.5e-10 ***	
	High Physical activity level	0.71	0.27 - 2.22	0.517	
Model 2	No physical activity	2.41	1.72 - 3.40	4.16e-07 ***	0.7228
	Mod. Physical activity level	1.36	1.14 - 1.63	0.000649 ***	
	High Physical activity level	0.68	0.24 - 2.20	0.484416	
Model 3	No physical activity	2.41	1.72 - 3.40	0.00303 **	0.7609
	Moderate Physical activity level	1.36	1.14 - 1.63	0.00307 **	
	High Physical activity level	0.68	0.24 - 2.20	0.39166	

OR odds ratio, CI confidence interval, NHANES National Health and Nutrition Examination Survey, AUC Area Under the Curve.
Model 1 only compares diabetes to physical activity
Model 2 adjusted for baseline age, gender, and race
Model 3 adjusted for baseline age, race, gender, education level, marital status, and BMI

Subgroup Analysis

The results obtained from the multiple logistic regression models, stratified by gender, are presented in Table 3. Both groups had higher odds of being diabetic without physical activity; however, female participants with no physical activity levels had higher odds of being diabetic than male participants, especially after adjusting for the covariates.

Table 3. ORs and 95% CI for the relationship between diabetes status and physical activity stratified by gender.

		Male participants					Female participants			
		OR	95% CI	p-value	AUC		OR	95% CI	p-value	AUC
Model 1	No physical activity	3.3043478	2.6921112 - 4.087514	< 2e-16 ***	0.5595		4.4571429	3.7233636 - 5.373603	< 2e-16 ***	0.5592
	Moderate Physical activity level	0.8583470	11.4483247 - 2.374071	8.69e-07 ***			1.6870488	1.3334127 - 2.129606	1.18e-05 ***	
	High Physical activity level	0.6052632	0.1570922 - 2.903094	0.483			0.8974359	0.2219567 - 5.988737	0.892	
Model 2	No physical activity	1.8143723	1.1475199 - 2.911414	0.011980 *	0.732		2.7106098	1.6852016 - 4.461988	5.73e-05 ***	0.7209
	Mod. Physical activity level	1.3822225	1.0622607 - 1.791120	0.015083 *			1.3518244	1.0557347 - 1.726774	0.0163 *	
	High Physical activity level	0.5053011	0.1188712 - 2.548143	0.364441			1.0790502	0.2434876 - 7.502776	0.9269	
Model 3	No physical activity	4.700778e+06	5.295678e-03 - 1.788754e+59	0.97048	0.7655		10.1396494	1.853586456 - 190.2491391	0.03037 *	0.7655
	Moderate Physical activity level	1.290674e+00	9.850183e-01 - 1.683959e+00	0.06197			1.3292289	1.032208354 - 1.7075710	0.02658 *	
	High Physical activity level	3.653550e-01	8.420537e-02 - 1.887279e+00	0.18872			1.2139816	0.269279394 - 8.6217458	0.818	

OR odds ratio, CI confidence interval, NHANES National Health and Nutrition Examination Survey, AUC Area Under the Curve.
Model 1 only compares diabetes to physical activity
Model 2 adjusted for baseline age, gender, and race
Model 3 adjusted for baseline age, race, gender, education level, marital status, and BMI

DISCUSSION

The result of our analysis indicates that people with diabetes are more likely to be less educated compared to nondiabetics. This finding is consistent with those of other developed countries[6]. The high prevalence of diabetes among this cohort may be because people with a higher level of education are more knowledgeable and perform health promoting behaviors more than the less educated. We also observed that the prevalence of diabetes was higher in older participants (> 45 years). The overall cost of diabetes management in the elderly has become a significant burden in the United States. In a report by the American Diabetes Association, 59% of healthcare expenditure for diabetes in the US in 2012 was spent on elderly patients[7]. Insulin production or sensitivity reduces with age, leading to impaired glucose tolerance, and effects can be worsened by a lack of an active lifestyle [24].

Hispanics and Blacks were more likely to have diabetes compared to other races. Gender distribution differed slightly between the subjects with diabetes, and there was a somewhat higher prevalence of diabetes in males.

The results obtained from the multivariate logistic regression models showed that among all the study participants, the prevalence of diabetes was significantly lower in the Moderate-physical-activity level group than in the no physical-activity level group before and after adjusting for covariates; however, there was no significant association with High physical activity levels, though the odds ratio was low. This is likely due to the low number of participants in the category.

We further explored the relationship between physical activity and diabetes and performed subgroup analyses stratified by gender while adjusting for confounding factors (e.g., age and race). We observed a similar relationship between physical activity and the prevalence of diabetes. The results indicated that physical activity was significantly and inversely related to an increased prevalence of diabetes.

In the subgroup analysis, significant relationships were observed only among the female subjects compared to males after adjusting for age, race, education level, marital status, and BMI. Indicating that females in the survey who did not practice any form of physical activity had a higher odds of being diabetic. This suggests that a moderate increase in physical activity in females is more beneficial for reducing the prevalence of diabetes.

Several biologic mechanisms may explain the association between physical activity and diabetic status. Physical activity increases the energy balance and maintains a healthy BMI. Another mechanism that might explain the impact of physical activity on the prevalence of diabetes regardless of obesity is that physical activity can directly reduce the blood glucose level and increase insulin sensitivity. Also, long-term physical activity leads to adaptive changes in the skeletal muscle. These changes include increased mitochondrial content and activity, changes in the muscle fiber type, and increased expression of the GLUT4 protein, which may reduce diabetes prevalence[8].

This study successfully analyzed the association between physical activity levels and diabetic status, using a large representative survey dataset, which increased the statistical power of the findings. After adjusting for several covariates, we performed a multiple logistic regression analysis. Still, we observed the positive effect of physical activity on diabetes, with increasing AUC values, indicating the stability and reliability of our results. We conducted a detailed subgroup analysis, which produced essential findings.

There were several limitations of this study. Firstly, we only considered the glycated HbA1c level as the diagnostic criteria for diabetes because of many missing values for fasting plasma glucose (FPG) and oral glucose tolerance test (OGTT) data. Because of this, we might have missed some subjects with prediabetes, though the HbA1c test is more sensitive than FPG and OGTT in the diagnosis of diabetes[9].
Although our models adjusted for the effects of important confounders, direct relationships could not be established because both the exposure and outcome variables were collected synchronously; hence, the association of physical activity level and diabetic status still needs to be analyzed in a larger-scale study.
Our analyses' information on physical activity was obtained by self-report, which might be subjected to recall bias. Self-report is not as accurate as standard biometric examination. Genetic factors that predispose to diabetes were not factored into this study, which might also cause bias. Data survey weights were not considered in our analysis, so this study does not represent the whole of the United States; rather, it represents only the survey's study participants. Finally, the study findings might not apply to other countries.

CONCLUSION

In conclusion, our present findings indicate that a physically active lifestyle may decrease the risk of diabetes. Further studies with larger sample size, also involving other countries, are needed to verify and strengthen the association between physical activity and the risk of diabetes. It is critical to improve the awareness of physical activity to prevent diabetes.

AUTHORS CONTRIBUTIONS

All authors contributed to the study. O.O downloaded the dataset performed exploratory data analysis and created the regression models. H.D, N.R, and B.K performed the correlation analysis and visualization. All authors contributed the manuscript writing.

ACKNOWLEDGEMENTS

The authors wish to acknowledge Dr. Meeta Pradhan and T.A Hardik Patel for their guidance throughout this project.

DISCLAIMER

The findings and conclusions in this study are those of the authors.

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES

1. American Diabetes, A., *Diagnosis, and classification of diabetes mellitus.* Diabetes care, 2012. **35 Suppl 1**(Suppl 1): p. S64-S71.
2. Zhao, F., et al., *Physical Activity Levels and Diabetes Prevalence in US Adults: Findings from NHANES 2015-2016.* Diabetes therapy : research, treatment and education of diabetes and related disorders, 2020. **11**(6): p. 1303-1316.
3. Colberg, S.R., et al., *Exercise and type 2 diabetes: the American College of Sports Medicine and the American Diabetes Association: joint position statement.* Diabetes care, 2010. **33**(12): p. e147-e167.
4. *Screening for type 2 diabetes.* Diabetes Care, 2003. **26 Suppl 1**: p. S21-4.
5. Cho, H., G.J. Matthews, and O. Harel, *Confidence Intervals for the Area Under the Receiver Operating Characteristic Curve in the Presence of Ignorable Missing Data.* Int Stat Rev, 2019. **87**(1): p. 152-177.
6. Blomster, J.I., et al., *The impact of level of education on vascular events and mortality in patients with type 2 diabetes mellitus: Results from the ADVANCE study.* Diabetes Res Clin Pract, 2017. **127**: p. 212-217.
7. Sircar, M., A. Bhatia, and M. Munshi, *Review of Hypoglycemia in the Older Adult: Clinical Implications and Management.* Can J Diabetes, 2016. **40**(1): p. 66-72.
8. Röckl, K.S., C.A. Witczak, and L.J. Goodyear, *Signaling mechanisms in skeletal muscle: acute responses and chronic adaptations to exercise.* IUBMB Life, 2008. **60**(3): p. 145-53.
9. *(2) Classification and diagnosis of diabetes.* Diabetes Care, 2015. **38 Suppl**: p. S8-s16.
10. Centers for Disease Control and Prevention. "National Health and Nutrition Examination Survey".
Available online: https: //www.cdc.gov/nchs/nhanes/about_nhanes.html (accessed on 22 November 2020).