

Exploring association between social determinants of health and physical inactivity among people with diabetes in the US: Analysis of BRFSS 2024 data

Fauzan Budi Prasetya¹

¹UCLA MDSH

I. INTRODUCTION

Regular physical activity has been known to bring benefits to individuals with diabetes, such as managing weight, managing blood pressure, lowering LDL cholesterol, and even lowering risk to get complications.¹ Exercises also help people with diabetes by lowering blood glucose levels and boosting body's sensitivity to insulin, countering insulin resistance.² However, there might be variables that might play a role to make people with diabetes physically inactive, such as social determinants of health.³ This study aims to explore the association between social determinants of health and physical inactivity among people with diabetes in the US by using Behavioral Risk Factor Surveillance System 2024 data, which can be used for further analysis in the future.

II. METHODS & MATERIALS

A. Data Set

This study utilizes a cross-sectional annual survey data from Centers for Disease Control and Prevention, called the Behavioral Risk Factor Surveillance System and collected in 2024. This dataset contains 345 variables and 457,670 individual records from 49 states of the United States of America. The datasets, which consist of landline and cell phone questionnaires, collect health-related risk behaviors, chronic health conditions, and use of preventive services. The inclusion criteria of this study is that they answered Yes to a question if they had been told that they had diabetes and have answered Yes or No to a question of participating in any physical activities or exercises such as running, calisthenics, golf, gardening, or walking for exercise, other than their regular job in the past month. The main outcome of this study is named "is physical inactive", and coded 1 for Yes and 0 for No, which is a reverse version of the original value. Focusing on social determinants, this study takes 19 predictor variables, which are respondent's basic information (sex, race, BMI

group), their socio economic status (education level group, income group, employment group, marital status, veteran, home ownership), respondent's feeling about their mental condition (life satisfaction, emotional supported, loneliness), and respondent's condition of basic necessity in the past 12 months (employment security, food security, receiving food stamps, mortgage, rent or utility bills security, lost/threatened losing utility service of electric, gas, oil, or water, neighborhood safety from crime, and transportation security for getting daily things needed). All the variables are in categorical format.

B. Study Population

The study population is split nearly evenly between Male (51.7%) and Female (48.3%). They are mainly categorized as Obese (53.4%) and Overweight (32.0%), followed by Normal (14.1%), and Underweight (0.6%). In terms of education, the highest proportion are College graduates or higher (35.3%) and those who Attended college or technical school (30.6%), followed by High school graduates (27.5%), and finally those who Did not graduate high school (6.7%). For employment, the largest proportion is the group of Not in the Labor Force (63.4%) that consist of homemaker, student, retired, and those who are unable to work, followed by Employed (33.4%) that consist of employed for wages and self-employed, and Unemployed (3.3%). Regarding home ownership, the majority Own House (75.5%), while 21.3% Rent House, and 3.3% are categorized as Other. The income groups are distributed with the largest proportion being Medium Income (45.8%) which are , then Low Income (34.3%), and High Income (19.9%). The marital status groups show that the majority are Married (51.1%), followed by those Formerly Married (34.6%), and then Unmarried (14.3%). The most common race group is White, Non-Hispanic (70.7%), followed by Black, Non-Hispanic (12.3%), Hispanic (8.9%), Other race, Non-Hispanic (3.4%), American Indian/Alaskan Native, Non-Hispanic (2.7%), and Asian, Non-Hispanic (2.0%). A portion

of them are Veteran (16.9%). Finally, the groups with the smallest proportions are: Bills Insecure (11.2%), Emotionally Unsupported (9.9%), Employment Insecure (7.3%), Food Insecure (15.9%), received Foodstamp (13.0%), Life Dissatisfied (7.3%), Lonely (32.3%), Neighborhood Unsafe (5.5%), Transportation Insecure (8.4%), and Utilities Threatened (8.1%).

C. Statistical Methods

This study focuses on performing exploratory analyses to determine which variables in the dataset can serve as covariates in predicting physical inactivity among people with diabetes in the US. Since this study utilizes a cross-sectional survey study, this study is not intended for causal analysis, but rather exploratory. To optimize the analysis, data preparation consists of collapsing small groups to a bigger group, such as the Employment group, which combines Self employment and employed for wages, to a group called Employed.⁴ Data cleaning consists of excluding records that are missing and refusing to answer. The whole analysis process consists of descriptive statistics, bivariate statistics by chi square to choose potential predictors, and multivariate statistics of logistic regression to measure the Odds Ratio. Alpha 0.05 is used. To supplement, percent of physical inactivity of the study population by state will be visualized by a hex map. The data is managed and analyzed using R software R version 4.5.2. The main library used is tidyverse. Hex map is drawn by utilizing library usmap. Since the format of the data is in the XPT file, a haven package is used.

Table 1: Baseline Characteristics and Bivariate Results (Chi-squared)

Variables	Total (N=22,775)	p-value
Physical inactive		
0 - Physical Inactive	14512 (64%)	
1 - Physical Active	8263 (36%)	
Sex		<0.01
1 - Male (ref)	11772 (51.7%)	
2 - Female	11003 (48.3%)	

BMI group		<0.01
1 - Underweight	133 (0.6%)	
2 - Normal (ref)	3204 (14.1%)	
3 - Overweight	7286 (32.0%)	
4 - Obese	12152 (53.4%)	
Education group		<0.01
1 - Did not graduate high school (ref)	1519 (6.7%)	
2 - High school graduate	6253 (27.5%)	
3 - Attended college or technical school	6964 (30.6%)	
4 - College graduate or higher	8039 (35.3%)	
Employment Group		<0.01
1 - Employed (ref)	7598 (33.4%)	
2 - Unemployed	747 (3.3%)	
3 - Not in Labor Force	14430 (63.4%)	
Home Ownership		<0.01
1 - Own House (ref)	17184 (75.5%)	
2 - Rent House	4840 (21.3%)	
3 - Other	751 (3.3%)	
Income Group		<0.01
1 - Low Income	7805 (34.3%)	
2 - Medium Income (ref)	10432 (45.8%)	
3 - High Income	4538 (19.9%)	
Marital Status		<0.01
1 - Married (ref)	11642 (51.1%)	
2 - Formerly Married	7876 (34.6%)	
3 - Unmarried	3257 (14.3%)	
Race group		<0.01
1 - White, Non-Hispanic (ref)	16107 (70.7%)	
2 - Black, Non-Hispanic	2804 (12.3%)	
3 - Asian, Non-Hispanic	456 (2.0%)	
4 - American Indian/Alaskan Native, Non-Hispanic	619 (2.7%)	
5 - Hispanic	2019 (8.9%)	
6 - Other race, Non-Hispanic	770 (3.4%)	
Is veteran	3854 (16.9%)	0.09

Is foodstamp	2960 (13.0%)	<0.01
Is utilities threatened	1836 (8.1%)	<0.01
Is bills insecure	2550 (11.2%)	<0.01
Is employment insecure	1670 (7.3%)	0.61
Is transportation insecure	1913 (8.4%)	<0.01
Is neighborhood unsafe	1245 (5.5%)	<0.01
Is food insecure	3620 (15.9%)	<0.01
Is lonely	7357 (32.3%)	<0.01
Is emotional unsupported	2263 (9.9%)	<0.01
Is life dissatisfied	1652 (7.3%)	<0.01

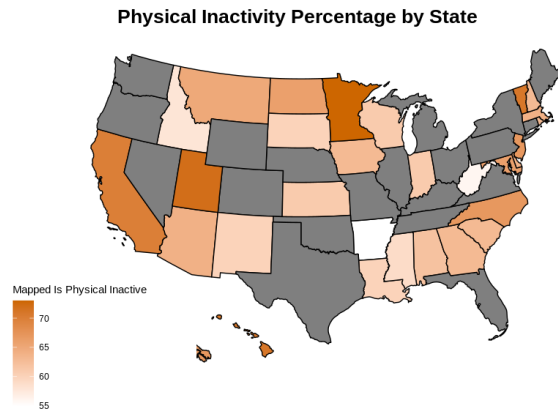


Figure 1: Physical Inactivity Percentage by State

III. RESULTS

By showing a hex map of percent physical inactivity each state, it found that there are some states missing. It happened because this study only includes records that have no missing values for all the variables and the data is collected by each state government. The highest percentage of physical inactivity are found in the state of Minnesota, Ohio, California, and Vermont.

After running the chi-square analysis for each independent variable and physical inactivity, almost all variables come out significant with p-value < 0.05, except Is employment insecure (p-value = 0.61) and is veteran (p-value = 0.09). Therefore, all the variables except for those two are included in the logistic regression.

Table 2: Multivariate Results (Logistic Regression)

Variables	Odds Ratio	p-value
Sex		
1 - Male (ref)		
2 - Female	1.39	<0.01
BMI group		
1 - Underweight	1.62	<0.01
2 - Normal (ref)		
3 - Overweight	1.07	0.142
4 - Obese	1.55	<0.01
Education group		
1 - Did not graduate high school (ref)		
2 - High school graduate	0.84	<0.01
3 - Attended college or technical school	0.70	<0.01
4 - College graduate or higher	0.52	<0.01
Employment Group		
1 - Employed (ref)		
2 - Unemployed	1.05	0.566
3 - Not in Labor Force	1.29	<0.01
Home Ownership		
1 - Own House (ref)		
2 - Rent House	0.96	0.307
3 - Other	1.10	0.218
Income Group		
1 - Low Income	1.25	<0.01
2 - Medium Income (ref)		
3 - High Income	0.71	<0.01

Marital Status

1 - Married (ref)		
2 - Formerly Married	1.03	0.408
3 - Unmarried	0.91	0.048

Race group

1 - White, Non-Hispanic (ref)		
2 - Black, Non-Hispanic	0.85	<0.01
3 - Asian, Non-Hispanic	0.82	0.085
4 - American Indian/Alaskan Native, Non-Hispanic	0.94	0.482
5 - Hispanic	0.95	0.328
6 - Other race, Non-Hispanic	0.74	<0.01

Is foodstamp 1.06 0.263

Is utilities threatened 0.95 0.394

Is bills insecure 1.01 0.829

Is transportation insecure 1.14 0.021

Is neighborhood unsafe 1.03 0.673

Is food insecure 1.04 0.398

Is lonely 1.19 **<0.01**

Is emotional unsupported 1.32 **<0.01**

Is life dissatisfied 1.77 **<0.01**

After running the logistic regression, top groups of variables have a positive correlation with physical inactivity are Is life dissatisfied (OR = 1.77), BMI group of underweight and obese (OR 1.62 and 1.55), sex Female (OR = 1.39), and Emotionally unsupported (OR = 1.32). Top variables with negative correlation are the education groups that graduated from college and attended college (OR 0.52 and 0.70) and a high income group (0.71).

For better visual understanding of these variables correlating with physical inactivity, a stacked bar chart is drawn. It can be seen that physical inactivity rate differs between groups, and will be higher as

education level and income decrease, and is the lowest in the normal BMI group.

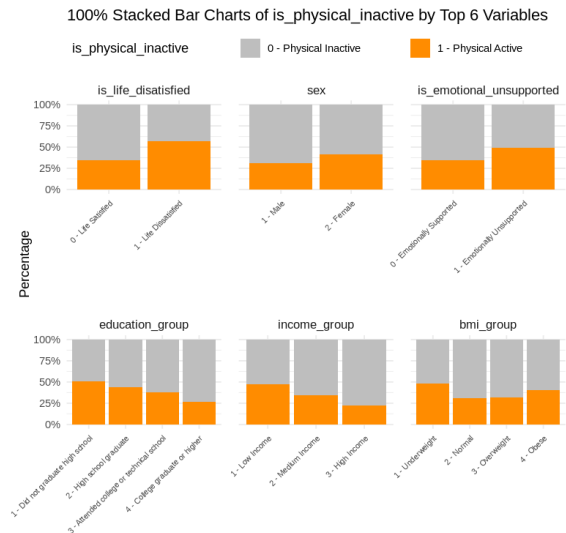


Figure 2: Stacked Bar Charts of % Physical Inactive for 6 Top Predictors

IV. CONCLUSIONS

Aside from BMI and Sex, social determinants, especially life satisfaction, emotional supports, education and income level are important to explain the role of physical activity among people with diabetes in the US. Individuals having life dissatisfaction will be more likely to have physical inactivity 1.77 times compared to those who feel satisfied in their life, and individuals who graduated from college will be less likely to be physically inactive by 52% compared to those who did not graduate from high school.

V. DISCUSSION

Limitations of the study is the data collected by a self answered questionnaire, which can produce reporting bias.⁵ The study also did not involve the methods for confounding and interaction variables, which could happen in observational studies.⁶

Similar to other studies, social support plays a role for physical activities in people with diabetes.⁷ If not maintained, high levels of distress can impact medication-taking behaviors and are linked to poorer dietary and exercise behaviors and higher A1C.⁸

ACKNOWLEDGEMENTS

Thanks to Dr. Grace Hyun J. Kim and TA's support and suggestions, as well as the students of BIOSTAT 203A for their suggestions.

REFERENCES

1. Harvard Health. The importance of exercise when you have diabetes [Internet]. Harvard Health. 2023. Available from: <https://www.health.harvard.edu/staying-healthy/the-importance-of-exercise-when-you-have-diabetes>
2. Get active [Internet]. Diabetes. 2024. Available from: <https://www.cdc.gov/diabetes/living-with/physical-activity.html>
3. O'Donoghue G, Kennedy A, Puggina A, Aleksovska K, Buck C, Burns C, Cardon G, Carlin A, Ciarapica D, Colotto M, Condello G, Coppinger T, Cortis C, D'Haese S, De Craemer M, Di Blasio A, Hansen S, Iacoviello L, Issartel J, Izzicupo P, Jaeschke L, Kanning M, Ling F, Luzak A, Napolitano G, Nazare JA, Perchoux C, Pesce C, Pischon T, Polito A, Sannella A, Schulz H, Simon C, Soehnle R, Steinbrecher A, Schlicht W, MacDonncha C, Capranica L, Boccia S. Socio-economic determinants of physical activity across the life course: A "DEterminants of Diet and Physical Activity" (DEDIPAC) umbrella literature review. PLoS One. 2018 Jan 19;13(1):e0190737. doi: 10.1371/journal.pone.0190737. PMID: 29351286; PMCID: PMC5774703. <https://pmc.ncbi.nlm.nih.gov/articles/PMC5774703/>
4. Lund B, Raimi S. Collapsing Levels of Predictor Variables for Logistic Regression and Weight of Evidence Coding. In: MWSUG 2012 Proceedings: Midwest SAS Users Group Conference; 2012. Paper SA-03. Available from: <https://www.mwsug.org/proceedings/2012/SA/MWSUG-2012-SA03.pdf>
5. Rosenman R, Tennekoon V, Hill LG. Measuring bias in self-reported data. Int J Behav Health Res. 2011 Oct;2(4):320-332. doi: 10.1504/IJBHR.2011.043414. PMID: 25383095; PMCID: PMC4224297.

- <https://pmc.ncbi.nlm.nih.gov/articles/PMC4224297/>
6. Fumo-Dos-Santos C, Ferreira JC. Dealing with confounding in observational studies. J Bras Pneumol. 2023 Sep 18;49(4):e20230281. doi: 10.36416/1806-3756/e20230281. PMID: 37729251; PMCID: PMC10578949. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10578949/>
7. Morowatisharifabad MA, Abdolkarimi M, Asadpour M, Fathollahi MS, Balaei P. Study On Social Support for Exercise And Its Impact on the Level of Physical Activity of Patients with Type 2 Diabetes. Open Access Maced J Med Sci. 2019 Jan 4;7(1):143-147. doi: 10.3889/oamjms.2019.016. PMID: 30740179; PMCID: PMC6352472. <https://pmc.ncbi.nlm.nih.gov/articles/PMC6352472/>
8. Young-Hyman D, de Groot M, Hill-Briggs F, Gonzalez JS, Hood K, Peyrot M. Psychosocial Care for People With Diabetes: A Position Statement of the American Diabetes Association. Diabetes Care. 2016 Dec;39(12):2126-2140. doi: 10.2337/dc16-2053. Erratum in: Diabetes Care. 2017 Feb;40(2):287. doi: 10.2337/dc17-er02. Erratum in: Diabetes Care. 2017 May;40(5):726. doi: 10.2337/dc17-er05. PMID: 27879358; PMCID: PMC5127231. <https://pmc.ncbi.nlm.nih.gov/articles/PMC5127231/>

Appendix

I. Source Data File

Dataset can be found here on CDC Website that are publicly available https://www.cdc.gov/brfss/annual_data/annual_data.htm

II. Data set

Upload the dataset that is generated from an R code in the bruin learn. If there are multiple data set, you can provide the details of data set with data schema.

It is recommended reporting the output of the file that describes the variables.

III. R Code

Include R Codes that are associated with generating Tables and Figures. It is expected that outcomes of R program will match the numbers that are reported in th Table 1 and Table 2, as well as in the Figures.