Examining the interplay between physical heath and socio-economic factors with special emphasis on prevalence of Obesity

Samriddhi Soni*

* MS in Applied Statistics, Teachers College, ss6557@tc.columbia.edu

Abstract

The impact of Obesity in today's world is a very global public health concern, with its impact reaching and extending beyond individual health. This paper tries to explore the relationship between the diagnosis of obesity among individuals and its connection with socio-economic factors including limited economic opportunities and demographic factors. This piece of research tries to explore the relationship between limited economic opportunities such as housing crisis, low family income and lack of health insurance on the diagnosis of obesity. Apart from analyzing the possible connection, an emphasis is laid on the physical health of individuals. The interplay of diagnosis of other diseases such as diabetes, common heart diseases and individual's physical health habits can be instrumental in causing more prevalence of obesity. The relationship of demographic factors and physical health have been explored in past literature; thus, this study aims at exploring and establishing the impact of demographic factors on the prevalence of obesity.

Keywords

Physical health, economic health, chronic diseases, diabetes, obesity, socio-economic issues, housing crisis, demographic information

1. INTRODUCTION

Obesity is a global health concern, that has garnered significant attention due to its association with a multitude of chronic health conditions. This paper presents a systematic investigation into the relationship between the diagnosis of obesity, physical health and socio-economic factors. To analyze the interplay of these chronic health conditions a comprehensive literature review was conducted, evidence of which is emphasized in TM Powell-Wiley 2021. The findings from TM Powell-Wiley 2021 elaborates on a connection between Obesity and other chronic disease and states that obesity contributes directly to

incident cardiovascular risk factors, including dyslipidemia, type 2 diabetes, hypertension, and sleep disorders. Similar findings have been presented in Chobot, A., Górowska-Kowolik, K., Sokołowska, M., & Jarosz-Chobot, P. (2018) that there is increasing scientific evidence regarding the role of obesity and overweight in type 1 diabetes. Present literature has evidently highlighted strong relationship between these diseases; therefore, this paper aims to assess the presence of relationship between diagnosis of obesity with diabetes and heart diseases.

Drawing on an extensive literature review, there is existing evidence on the associations between obesity and socio-economic factors, as well as their combined influence on diabetes and heart diseases. The findings reveal a multifaceted relationship between obesity and socio-economic factors. The housing crisis for example, often linked to overcrowded or bad living condition is associated with higher risks of CVD. Gu, K. D., Faulkner, K. C., & Thorndike, A. N. (2023) talks about the association of housing instability and physical health, their findings indicate adverse associations between housing instability and cardiometabolic health; including higher prevalence of overweight/obesity, hypertension, diabetes, and cardiovascular disease; worse hypertension and diabetes control. Similarly, individuals lacking health insurance face barriers to accessing healthcare, potentially resulting in delayed diagnoses. Low economic opportunity and financial insecurity are also implicated in the obesity epidemic, influencing dietary choices, physical activity, and overall health behaviors.

Some other demographic factors such as factors household income of a family influence eating habits and life choices. Following framework from Nowrouzi B, et al. (2015), there are various pathways in which income relates to obesity and vice versa: with reference to the perspective of causation, income does not only restrict one's access to (healthy) food, but is also linked to higher health literacy which, in

turn, is positively related to health-promoting behaviors (i.e., healthy nutrition, physical activity). Thus, highlighting the findings of Kim, T. J., & von dem Knesebeck, O. (2018) relates back to the finding that lower income can lead to higher chances of diagnosis of obesity.

The major factors that will be analyzed for the study are the physical health conditions, presence of chronic health diseases, day to day habits and demographic information. Therefore, by examining these two research questions, this study will provide insight on the major contributing factors of diagnosis of obesity.

1. To what extent does the physical health and chronic health conditions have similar pattern with the prevalence of obesity?
2. Do economic and demographic factors such as housing conditions, metropolitan area, lack of health insurance and low household income relate to prevalence of obesity?

2. RESEARCH METHODOLOGY

2.1 Sample and Data Source

The current research sample is based on the data presented in the study Apenyo, T., Vera-Urbina, A. E., Ahmad, K., Taveira, T. H., & Wu, W. C. (2022). The data set includes data from 3126 US counties using publicly available data relating poor housing conditions with COVID-19 outcomes. Counties with missing data for poor housing condition (n = 6) were excluded yielding a sample size of 3126 US counties. Since there is no individual identifying information, the data were in aggregate by county, and publicly available from the Centers for Disease Control (CDC), US census Bureau and John Hopkins Coronavirus Resource Center, the protocol received exemption from the Providence Veterans Affairs Medical Center Institutional Review Board students.

The data set has variables that indicate the proportion of diagnosis of chronic health diseases and demographic information. The dataset contains 11 variables and 3126 data points (number of counties). The missing value rows have been removed, for analysis. The primary variable for the study is "Obesity_diag", which is the percentage of people diagnosed with obesity in a county.

Variables selected to detect the information about the geographical location, presence of chronic health conditions and information about physical health are "County" which is the county from where the data is collected, "State" which is the state in which the respective county belongs, "Diabetes_diag" which indicates the percentage of people diagnosed with diabetes, "Heart_disease" which indicates the percentage of cardiac dysrhythmia hospitalization, "Smokers" which indicates the percentage of adults who are current smokers and "Metrocity" which is a factor variable indicating the metropolitan status of the county; 1 is large central metro, 2 is large fringe metro , 3 is medium small metro and 4 is nonmetro.

The variables selected to factor the socio-economic issues and demographic information are "Housing_crisis" which indicates the percentage of households living with Severe Housing Problems, "Without_healthinsurance" which indicates the percentage of people without health insurance and "Median_income" which indicates the Median Household Income (in thousands of $) and "Median_age" which is the Median age of individuals in the county.

2.2 Methods

2.2.1. Descriptive Statistics

To obtain the information about the distribution of the variables across all data points and its average and spread, the analysis of descriptive statistics of the variables is beneficial. Table 1

represents the information about all the numerical variables, the categorical variables such as metrocity, county and state have not been included in the table.

Table 1

Descriptive Statistics of Variables

Variable Name	Minimum	Maximum	Mean	Standard Deviation
Median_age	22.4	62.7	40.364	5.0201976
Median_income	22	134.6	49.372	12.8117419
Diabetes_diag	3.6	18.3	9.8494	2.2486204
Obesity_diag	13.5	49.7	31.973	4.715183
Heart_disease	16.6	136.9	58.653	16.8881612
Smokers	6.0	41.0	17.439	3.551772
Housing_crisis	2.7	39.1	14.305	4.337582
Without_healthinsurance	2.1	31.1	11.049	4.8488057

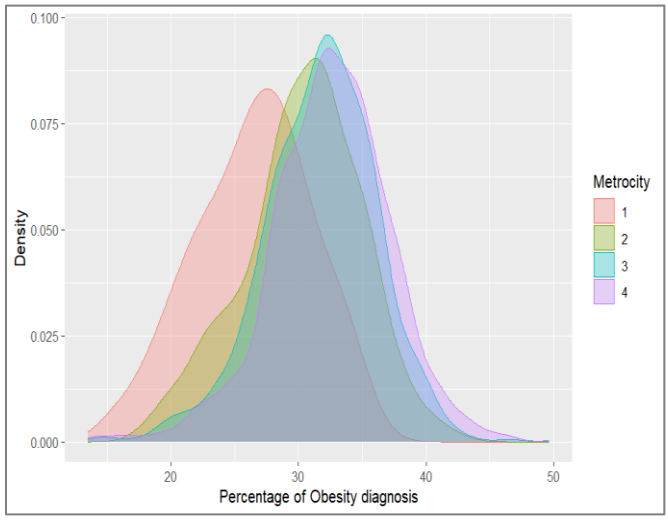
Note. The table does not have null values because all the rows with missing values has already been removed in the data preprocessing step

2.2.2 Exploratory data visualization

Exploratory data visualizations allow users to explore data interactively. It can be used to manipulate variables, zoom in on specific aspects, and gain a deeper understanding of the data through exploration and analysis of variables. It is very important to accurately represent data in a visualization, avoid distracting elements and highlight the important information as mentioned in Schwabish (2021). Gestalt principles are very useful set of rules that help in achieving effective and useful visualizations. The laws of connection, similarity and proximity are followed in the visualizations to achieve the purpose of communicating the results from the data. The figures depict the percentage of diagnosis of obesity with the geographical indicators such as metropolitan region, demographic information such as median income and presence of other chronic disease such as diagnosis of diabetes.

Figure 1

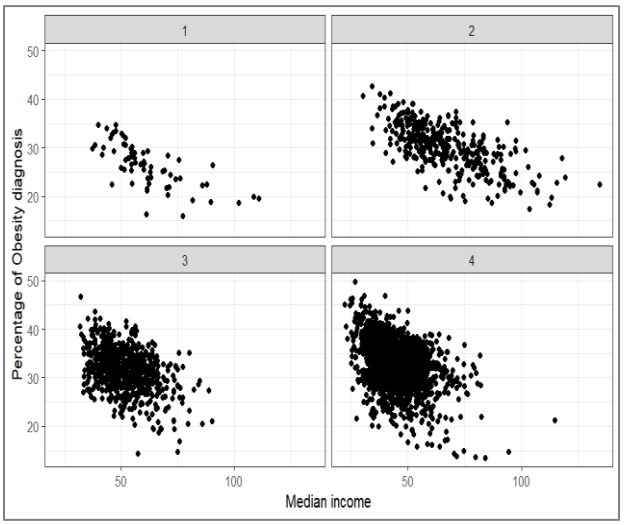
Density graph of obesity diagnosis based on metropolitan status



Note. 1 is large central metro, 2 is large fringe metro, 3 is medium small metro and 4 is nonmetro

Figure 2

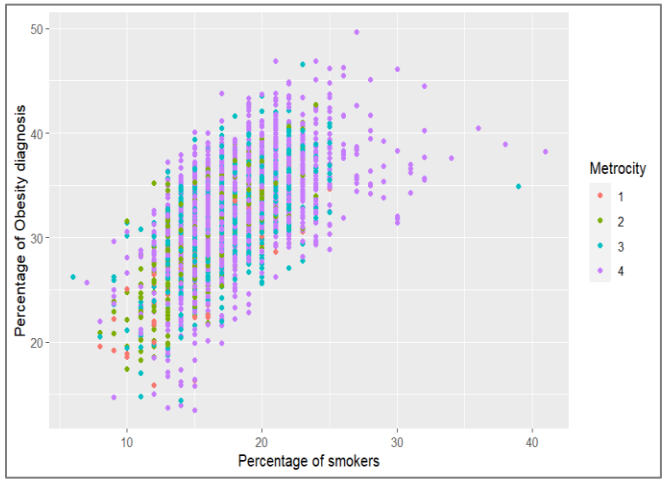
Relationship between income and obesity



Note. 1 is large central metro, 2 is large fringe metro, 3 is medium small metro and 4 is nonmetro

Figure 3

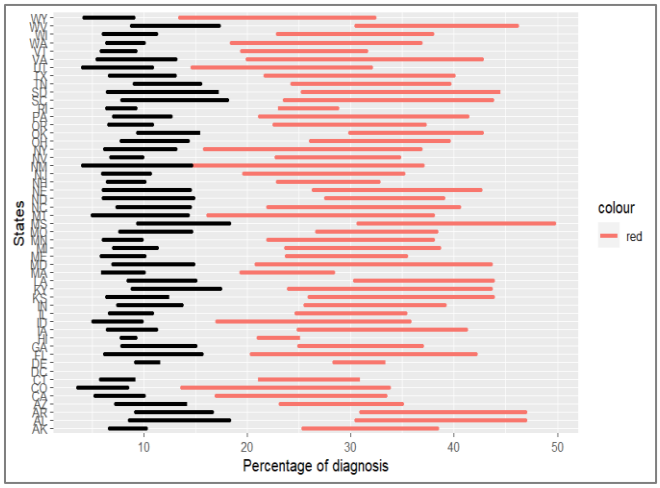
Grouped Scatter plot for obesity diagnosis vs smokers



Note. 1 is large central metro, 2 is large fringe metro, 3 is medium small metro and 4 is nonmetro

Figure 4

Line Chart for percentage of obesity and diabetes diagnosis



Note. Red lines denote the percentage of obesity diagnosis and black lines denote the percentage of diabetes diagnosis

2.2.3 HCA Heatmap

In this study HCA heatmap are used as a data visualization technique to display the results of hierarchical clustering. For modeling the relationship between the presence of chronic health diseases, demographic factors and physical health conditions this author chose cluster analysis heatmaps after following recommendations for HCA heatmaps from Bowers (2010) using uncentered correlation as the distance metric and average linkage as the unsupervised hierarchical clustering algorithm. This heatmap provide a graphical representation of the relationships and similarities between covariates by arranging theme in similar clusters.

The clustering distance metric used is uncentered correlation (Equation 1) as the distance metric, $r(x_i, y_i)$, which is preferred over the similar Pearson correlation as it assumes the mean is zero for each vector x and y after recommendation from Bowers (2010). Thus, the uncentered correlation for any two vectors x_i and y_i of sample size *n* is:

$$r(x_i, y_i) = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{x_i}{\sigma_x^{(0)}}\right)\left(\frac{y_i}{\sigma_y^{(0)}}\right) \qquad \textbf{Equation 1}$$

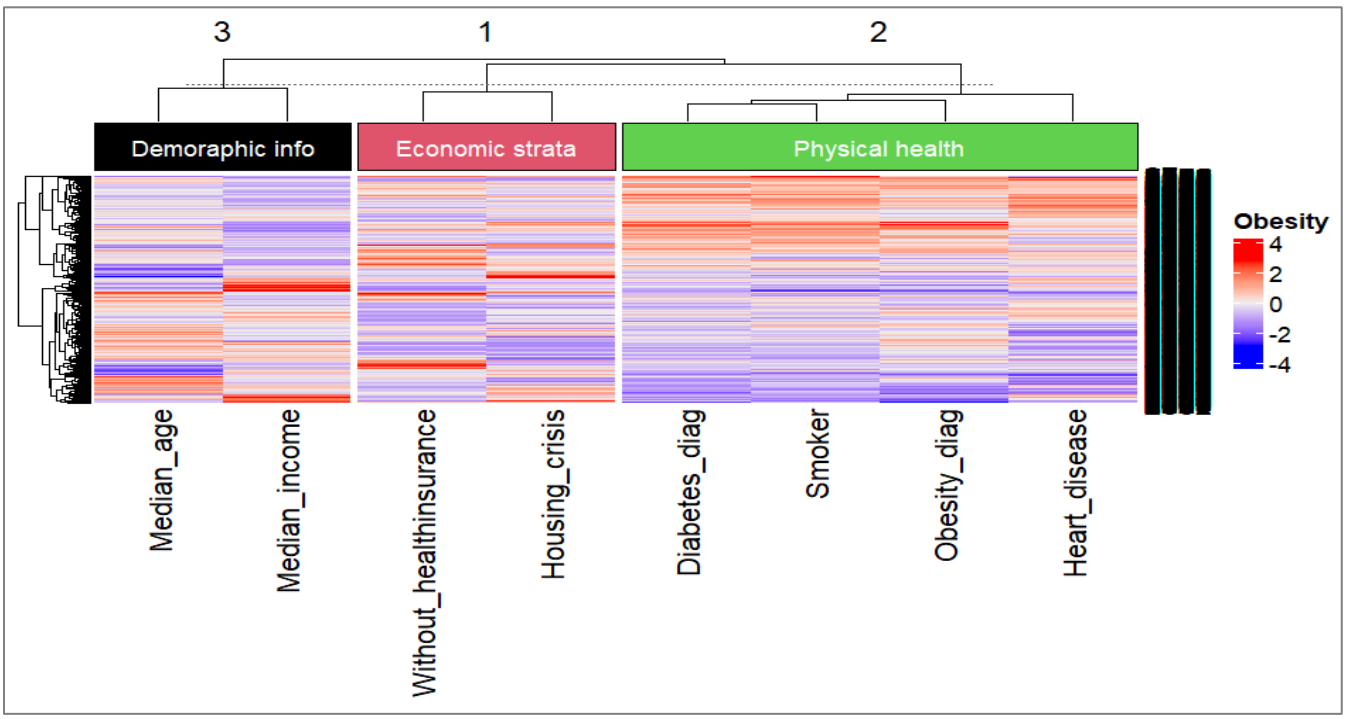
where:

$$\sigma_x^{(0)} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i)^2}$$

$$\sigma_y^{(0)} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i)^2}$$

Figure 5

HCA Heatmap of demographic information, economic factors and physical health conditions



3. RESULT AND FINDINGS

Figure 1 displays the percentage of diagnosis of obesity based on the metropolitan status of the counties, the counties which are smaller metro areas have the maximum density for higher percentage of diagnosis of obesity. Figure 2 displays that as the median household income increases the percentage of diagnosis of obesity decreases, and this is constant across all metropolitan areas. One interesting observation here is that the prevalence of obesity is less in larger metropolitan areas and the with the increase in median income the prevalence of obesity decreases which is similar to the finding in Kim, T. J., & von dem Knesebeck, O. (2018). Figure 3 displays that with increase in the percentage of smokers in the population the prevalence of obesity also increases, the higher percentage of smokers and people diagnosed with obesity belong to the counties that belong small metropolitan areas. This finding can relate

to the lack of economic opportunities and awareness available in these small metropolitan areas, this result is consistent with the discussion about the effect of higher income on the less prevalence of obesity in Kim, T. J., & von dem Knesebeck, O. (2018).

Figure 4 displays the percentage of diagnosis of diabetes and obesity for various states the counties belong to, the prevalence of diabetes affects the prevalence of obesity as the trend of the proportion of diagnosis of these diseases are very similar. Chobot, A., Górowska-Kowolik, K., Sokołowska, M., & Jarosz-Chobot, P. (2018) highlights similar relationship that there is increasing scientific evidence regarding the role of obesity and overweight in type 1 diabetes.

The HCA heatmap (Figure 5) displays three major categories in which all the covariates have been divided, economic strata, physical health and demographic information. The percentage of diagnosis of obesity has been clustered under the physical health cluster. This cluster contains other variables such as the percentage of individuals diagnosed with diabetes, percentage of individuals with heart disease and percentage of current smokers. These four variables have almost identical pattern; thus, it indicates a clear relationship that the presence of chronic diseases such as diabetes and heart diseases and unhealthy habit such as smoking increases the prevalence of obesity. Present literature provides strong evidence in favor of the interplay of the diagnosis of diabetes, heart diseases and unhealthy lifestyle in increasing the prevalence of obesity.

The second cluster groups the covariates based on the economic strata, this cluster includes housing crisis and lack of health insurance. This cluster showcases similar pattern as indicated by the physical health cluster, thus indicating that adverse living conditions and economic instability increases

the risk and prevalence of overweight/obesity. Third cluster which includes the demographic variables such as median household income has an opposite pattern to the physical health cluster which indicates that with the increase in household income the prevalence of obesity decreases, thus higher income provides access to nutritious diet options and healthy lifestyle opportunity with awareness about this subject as highlighted in present literature Nowrouzi B, et al. (2015).

1. *To what extent does the physical health and chronic health conditions have similar pattern with the prevalence of obesity?*

The findings from visualizations and clustered heatmaps have indicated a strong relationship between the presence of chronic health conditions and physical health such as percentage of diabetes, heart diseases, smoking habits and the prevalence of obesity

2. *Do economic and demographic factors such as housing conditions, metropolitan area, lack of health insurance and low household income relate to prevalence of obesity?*

The heatmap visualizes the relationship between economic factors and individual's physical health, individuals lacking health insurance face barriers to accessing healthcare, potentially resulting in delayed diagnoses. Low economic opportunity and financial insecurity are also implicated in the obesity epidemic, influencing dietary choices, physical activity, and overall health behaviors. Therefore, housing insecurity and low-income security increases the prevalence of obesity. The prevalence of diabetes is higher in small metropolitan areas due to the lack of economic opportunities and accessibility.

DISCUSSION AND CONCLUSION

This study aims as analyzing the relationship between physical health, presence of chronic diseases, economic and demographic factors in assessing the prevalence of obesity among individuals. The visualizations and HCA heatmap indicate a strong relationship between the presence of chronic disease (diabetes) and heart disease and the prevalence of obesity. The economic and demographic factors such as low household income, housing insecurity and lack of health insurance also acts as an indicator towards low economic security and has a strong relationship in the increasing the prevalence of obesity.

The scope of study is limited based on the datasets that we have drawn from Apenyo, T., Vera-Urbina, A. E., Ahmad, K., Taveira, T. H., & Wu, W. C. (2022) and the data is extracted from a cross-sectional study therefore the long-term relationship of the variables cannot be assessed through this study.

REFERENCES

1. Powell-Wiley, T. M., Poirier, P., Burke, L. E., Després, J. P., Gordon-Larsen, P., Lavie, C. J., ... & American Heart Association Council on Lifestyle and Cardiometabolic Health; Council on Cardiovascular and Stroke Nursing; Council on Clinical Cardiology; Council on Epidemiology and Prevention; and Stroke Council. (2021). Obesity and cardiovascular disease: a scientific statement from the American Heart Association. *Circulation*, *143*(21), e984-e1010.
2. Chobot, A., Górowska-Kowolik, K., Sokołowska, M., & Jarosz-Chobot, P. (2018). Obesity and diabetes—Not only a simple link between two epidemics. *Diabetes/metabolism research and reviews*, *34*(7), e3042.
3. Gu, K. D., Faulkner, K. C., & Thorndike, A. N. (2023). Housing instability and cardiometabolic health in the United States: a narrative review of the literature. *BMC Public Health*, *23*(1), 1-27.

4. Apenyo, T., Vera-Urbina, A. E., Ahmad, K., Taveira, T. H., & Wu, W. C. (2022). Association between median household income, state Medicaid expansion status, and COVID-19 outcomes across US counties. *Plos one*, *17*(8), e0272497
5. Nowrouzi, B., McDougall, A., Gohar, B., Nowrouz-Kia, B., Casole, J., & Ali, F. (2015). Weight bias in the workplace: A literature review. *Occup Med Health Aff*, *3*(206)
6. Bowers, A. J. (2010). Analyzing the longitudinal K-12 grading histories of entire cohorts of students: Grades, data driven decision making, dropping out and hierarchical cluster analysis. *Practical Assessment, Research, and Evaluation*, *15*(1)
7. Schwabish, J. (2021). The practice of visual data communication: what works. *Psychological Science in the Public Interest*, *22*(3), 97-1