Health Trends and Interactions Across Age Groups

BRFSS 2015 Analysis

Ethan Dang, Maggie Feng, Bao Hu, Jianheng Luo

Fushing Hsieh

STA 160

May 12, 2024

I. Introduction

The intricate relationship between health and disease has long attracted the attention of the medical and scientific communities. The World Health Organization defines health as “a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity”. This definition emphasizes the multifaceted nature of health, extending beyond the absence of disease to encompassing broader well-being. Rather, a disease is a condition that impairs normal life functions and is characterized by specific symptoms and signs. The dynamic interplay between these two states shapes individual lives and affects broader public health outcomes. In addition, through the Centers for Disease Control and Prevention, it said that age is an important factor about disease and aging increases the risk of chronic diseases.

The research aims to explore the complex interdependencies between health and disease, looking at how various factors such as body indicators, lifestyle and social determinants influence heart disease, storks and diabetes. and explore whether there are effects between these health factors.

In conducting this investigation, the research will review existing literature and incorporate empirical data analysis to assess the impact of three diseases on health indicators. In addition, the study will also explore the impact of age groups on diseases and the impact of various health factors in different age groups. By integrating these disparate elements, the study aims to provide valuable insights into ongoing discussions about health promotion and disease prevention, ultimately helping to develop more effective public health policies and interventions.

II. Data Description

The dataset originates from the Behavioral Risk Factor Surveillance System (BRFSS) 2015, a comprehensive health-related telephone survey conducted annually by the Centers for Disease Control and Prevention (CDC) since 1984. The 2015 survey gathered data from 441,455 Americans, covering a broad spectrum of health-related risk behaviors, chronic health conditions, and demographic information. The data consists of 253,680 observations with 22 variables. Below is a short description of key variables based on the datacard provided by the author:

- **HeartDiseaseorAttack**: Indicator of coronary heart disease or myocardial infarction
- **HighBP**: Indicator of high blood pressure diagnosis by a health professional
- **HighChol**: Indicator of high cholesterol levels diagnosed by a health professional
- **BMI**: Body Mass Index, a measure of body fat based on height and weight
- **Stroke**: Indicates a history of stroke
- **Diabetes**: Indicates diabetes status–0: no diabetes, 1: pre-diabetes, 2: has diabetes
- **GenHlth**: Self-rated general health from 1 (excellent) to 5 (poor).

- **Menthlth**: Number of days with poor mental health in the past 30 days.
- **PhysHlth**: Number of days with poor physical health in the past 30 days.
- **DiffWalk**: Difficulty in walking or climbing stairs, where 0 is having difficulty and 1 is not having difficulty.
- **Sex**: Gender, where 0 is female and 1 is male.
- **Age**: Age group ranging from 1 (18-24 years) to 13 (80 years or older).

To explore the impact of these variables on heart health, we formulated a multivariate response variable Y with eight potential combinations from (0,0,0) to (1,1,1) capturing the presence or absence of heart disease, stroke, and diabetes. Originally, the diabetes variable had three categories. For simplicity and effectiveness in our analysis, we combined the categories for pre-diabetes and diagnosed diabetes into a single category. We also simplify the analysis of age-related trends by categorizing the original 13 age groups into 4 distinct categories:

- **Young Adult**: 18-29 years
- **Middle Age**: 30-44 years
- **Elders**: 45-64 years
- **Seniors**: 65+ years

Furthermore, we categorized BMI to 4 groups to better understand its impact on heart health:

- **Underweight**: Less than 18.5
- **Healthy weight**: 18.5-24.9
- **Overweight**: 25-29.9
- **Obesity**: 30.0 and above

Our primary goal is to investigate how these age groups interact with the risk factors encapsulated in Y, allowing us to identify age-specific trends and potential preventative measures against heart disease.

III. Methodology

We categorized individuals into distinct age groups as previously mentioned. These age groups were structured to capture different stages of life and implemented into an analysis of age-related patterns within our dataset. Then we applied stacked histograms of our response variables and age groups across different categories such as DiffWalk, HighBP, GenHlth, and BMI.

Thereafter, we calculated the entropy, joint entropy, and conditional entropy to get mutual information of each independent variable from each age group in our dataset. By conducting an entropy-based analysis, we can understand the relationships within our dataset. Entropy measures uncertainty, whereas joint entropy assesses interdependence between variables. To calculate entropy we used the following equation:

$$Entropy = -\sum_{i=1}^{n} p_i log(p_i)$$

where p_i is the probability of each unique value of variable and n is the total number of unique values in this variable. Information gain was evaluated between age groups and x values to get the mutual information of the top five variables of each age bracket. To calculate mutual information we employed the following equation:

$$I[X,Y] = CE[Y] - CE[Y|X]$$

where $CE[Y]$ is the conditional entropy of Y and $CE[Y|X]$ is the conditional entropy of Y given X. To evaluate the presence of interaction effects, we utilized the calculator of conditional

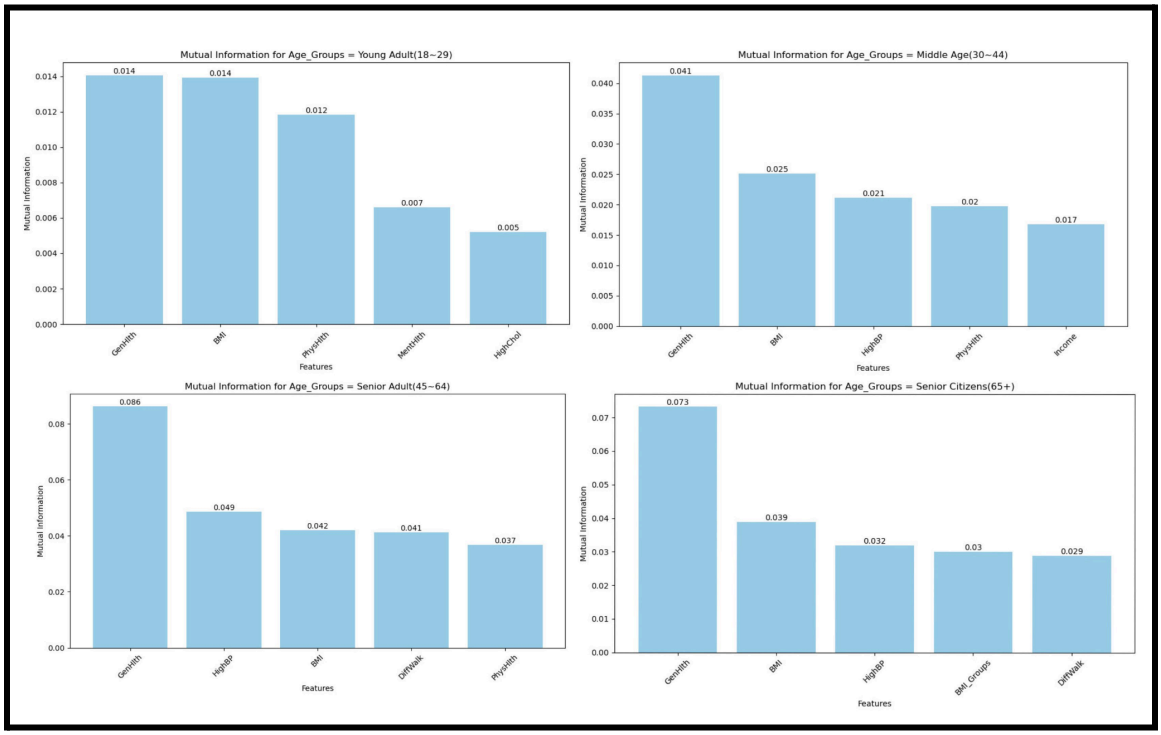
entropy. We followed the global perspective considering a fused covariate variable $X = (X_1, X_2)$ to compare whether:

$$CE[Y] - CE[Y|X| > CE[Y] - CE[Y|X_1] + CE[Y] - CE[Y|X_2]$$

and determine if the interactions are present. Subsequently, we acquired the top five interactive pairs of each age group.

IV. Results and Discussion

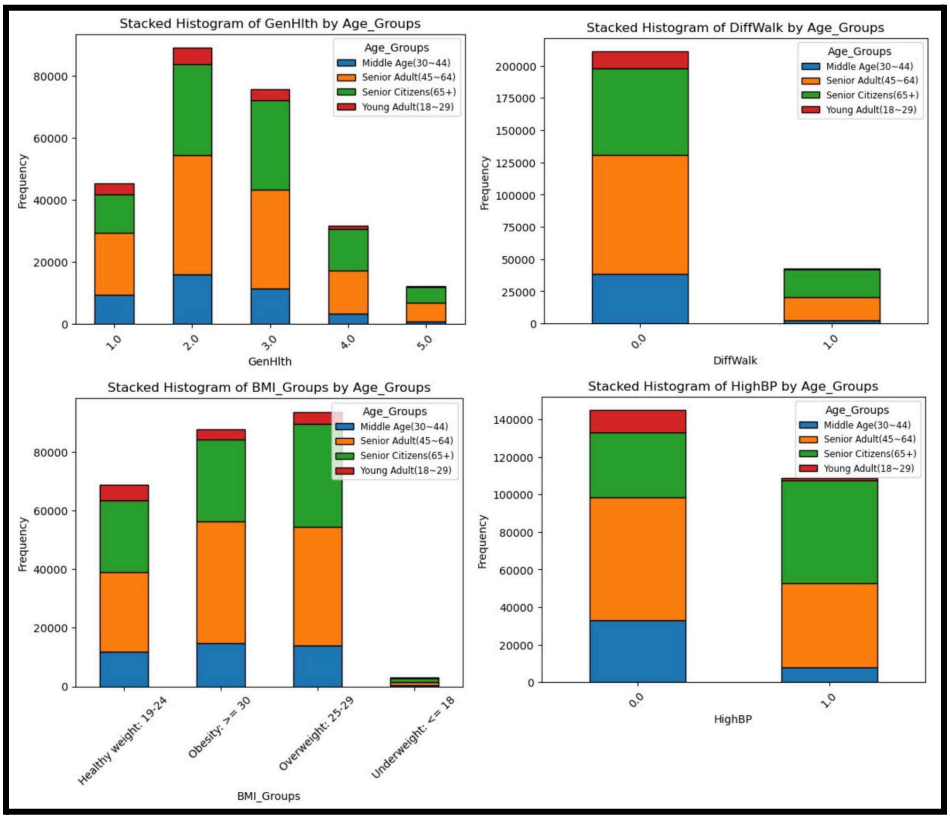
IV.I Entropy



To find out the indicators which affect the disease most for each group, we made a graph about the top 5 mutual information for each age group. Based on these four graphs, we can see that for each age group, general health has the highest mutual information value. Especially for senior adults and senior citizens, the mutual information of them is much higher than that of young adults and middle age. As a result, we can conclude that general health is much more important for seniors, adults and senior citizens than young adults and middle ages. In other words, the value of general health has the highest impact on stroke, heart disease, and diabetes. In addition, except general health, the mutual information of other features for senior adults and citizens is generally higher than that for young adults and middle age. Based on this finding, we can see that for older people, almost all health indicators have a greater impact on disease than for younger people.

Based on the findings, we observe that GenHealth, DiffWalk, BMI, and HighBP occur the most frequently across different age groups. Therefore, we chose to focus on those factors in creating our stacked histograms in the next section.

IV.II Stacked Histograms



Top left. In the stacked histogram of general health by age groups, we observe the majority of our age demographics have very good general health. Additionally, young adults mostly have excellent to fair general health. In the 5.0 category of poor health, we see its majority consists of senior adults and senior citizens.

Top right. In the histogram showcasing difficulty walking by age groups, the prevalence of young and middle-aged adults is low in the difficulty walking category. While most of our demographic indicates no difficulty walking, we observe senior citizens and senior adults consist of the majority of those who do have difficulty walking.

Bottom left. In the histogram of BMI groups by age group, we observe only a small amount of our dataset is underweight. The majority of our data set falls into some unhealthy weight (obesity, overweight, and underweight), which consists of mostly senior adults and senior citizens.

Bottom right. In the histogram of high blood pressure by age groups, we observe minimal occurrence among young and middle-aged groups of people with high blood pressure. Those with high blood pressure mostly consist of senior adults and senior citizens in our dataset.

Overall, our analysis of stacked histograms across different age groups shows insights into our focused health-related factors. We observed a trend where young adults occur most in the positive health related categories, whereas seniors are distributed across both positive and

negative health related categories. These findings emphasize such health factors that span across different age groups in our dataset.

IV.III Interactions

Interaction effects are essential as they help us understand how combinations of different variables influence health outcomes–specifically heart disease, stroke, and diabetes–in ways that are not evident when these variables are considered independently. Our analysis revealed the following five pairs of features with the highest differences for each age group, indicating strong interactions effects. From our results, we observe consistent strong interactions across all age groups between BMI and Mental Health, BMI and Physical Health, and BMI and Income.

```
Age Group: Middle Age(30~44)
     Feature1   Feature2 Interaction Effect  Difference
56        BMI   MentHlth              Exist    0.024039
57        BMI   PhysHlth              Exist    0.023220
61        BMI     Income              Exist    0.012620
138  MentHlth   PhysHlth              Exist    0.008475
60        BMI  Education              Exist    0.007094

Age Group: Senior Adult(45~64)
    Feature1   Feature2 Interaction Effect  Difference
56       BMI   MentHlth              Exist    0.018467
57       BMI   PhysHlth              Exist    0.016787
61       BMI     Income              Exist    0.004409
60       BMI  Education              Exist    0.002758
151      Sex     Income              Exist    0.002700

Age Group: Senior Citizens(65+)
     Feature1  Feature2 Interaction Effect  Difference
57        BMI  PhysHlth              Exist    0.022568
56        BMI  MentHlth              Exist    0.020625
61        BMI    Income              Exist    0.011017
138  MentHlth  PhysHlth              Exist    0.008838
151       Sex    Income              Exist    0.007031

Age Group: Young Adult(18~29)
     Feature1  Feature2 Interaction Effect  Difference
56        BMI  MentHlth              Exist    0.023908
57        BMI  PhysHlth              Exist    0.022509
61        BMI    Income              Exist    0.021050
138  MentHlth  PhysHlth              Exist    0.011846
142  MentHlth    Income              Exist    0.011810
```

The interaction between BMI and mental health suggests that mental well-being can influence how body weight impacts heart health. People with poor mental health may face different heart risks at the same BMI level compared to those with better mental health. Similarly, the interaction between BMI and physical health indicates that existing health conditions can affect the heart risks associated with one's BMI. For instance, a person with a high BMI and health issues may have more severe heart problems than someone without such issues. The interaction between BMI and income shows how one's financial status can affect the health implications of their BMI. For instance, individuals with lower income might experience different heart health challenges at the same BMI compared to those with higher income, possibly due to differences in access to healthcare.

V. Conclusion

Our analysis, beginning with the mutual information values across different age groups, has revealed significant insights. General health consistently emerged as the most influential indicator affecting diseases such as stroke, heart disease, and diabetes, with its impact being particularly pronounced in senior adults and senior citizens. This finding underscores the increasing importance of general health as individuals age, with older groups showing higher mutual information values for nearly all health indicators compared to younger groups.

In addition, the stacked histogram analysis further enhances our comprehension of age-related trends in health conditions. While younger adults typically display healthier profiles, older age groups exhibit increased prevalence of challenges like difficulty walking and high blood pressure. This trend underscores the need for early interventions and ongoing health monitoring to prevent the deterioration of health as individuals age.

Furthermore, exploring the interaction effects between various features such as BMI, mental health, physical health, and income deepens our understanding of the complex dynamics at play. These interactions highlight how factors like socioeconomic status and physical well-being intricately combine to influence individual health outcomes.

Nonetheless, by integrating these findings we gain a comprehensive view of the factors influencing heart health across different life stages. This nuanced understanding can guide more targeted and effective prevention and intervention strategies, tailored to address the specific health needs and conditions of each age group. Such tailored approaches are crucial for improving overall health outcomes and effectively managing the risk of heart disease, stroke, and diabetes.

VI. Reference

Centers for Disease Control and Prevention. (2022, June 3). *About adult BMI*. Centers for Disease Control and Prevention. https://www.cdc.gov/healthyweight/assessing/bmi/adult_bmi/index.html

Stages of life: Health for every age. Integris Health. (2015, October 9). https://integrishealth.org/resources/on-your-health/2015/october/stages-of-life-health-for-every-age

Centers for Disease Control and Prevention. (2022b, September 8). *Promoting health for older adults*. Centers for Disease Control and Prevention. https://www.cdc.gov/chronicdisease/resources/publications/factsheets/promoting-health-for-older-adults.htm#:~:text=Aging%20increases%20the%20risk%20of%20chronic%20diseases,with%20age%2C%20and%20these%20conditions%20are%20most

VII. Appendix

Link to the code: https://github.com/hubao666/Data-Analysis-for-BRFSS2015-Data