Research on the Influencing Factors that May Lead to Heart Attack

Xichen Wang*

Collage of Liberal Art, Jiangsu, University of Minnesota, 211400, China

*Corresponding author: wan00714@umn.edu

Abstract. Coronary artery disease stands as the primary factor leading to most heart attacks, a prevalent type of heart condition. This condition arises when the coronary arteries, which play a crucial role in delivering oxygenated blood to the heart muscle, face obstruction. The obstruction is due to the accumulation of plaque, a waxy substance that narrows the arteries over time. This narrowing can lead to angina, a condition where blood flow to parts of the heart muscle is restricted. Over time, the plaque within the artery has the potential to rupture, leading to the formation of a blood clot on its surface. Findings indicate that when this clot reaches a significant size, it has the capacity to obstruct the normal blood flow to the heart. Without proper intervention, it can result in damage to a specific portion of the heart muscle. It's crucial to note that not all heart attacks stem from these blockages. In instances where alternative heart and blood vessel conditions instigate a heart attack without the presence of obstructive coronary artery disease (referred to as MINOCA), the findings reveal that it is labeled as a myocardial infarction.

Keywords: Heart disease; MINOCA; myocardial infarction.

1. Introduction

Heart disease has been a major concern of the medical community as one of the long-intractable human diseases. Heart disease is the leading cause of death in the United States for men, women, and most racial and ethnic groups. According to data given by the Centers for Disease Control and Prevention, one person dies of cardiovascular disease every 33 seconds in the U.S. In 2021, about 695,000 people in the U.S. will die of heart disease, or one in five deaths [1]. Between 2018 and 2019, heart disease is costing the cost of heart disease to the United States is approximately $239.9 billion annually between 2018 and 2019 [2]. This encompasses expenses related to healthcare services, pharmaceuticals, and the economic impact resulting from fatalities [3]. Heart disease stands as the predominant cause of mortality in the United States across various racial and ethnic groups, encompassing African Americans, American Indians, Alaska Natives, Hispanics, and whites.

Heart disease is second only to cancer among Pacific Island women, Asian Americans, Alaska Natives, American Indians, and Hispanic women. Many factors influence heart attacks [4]. An unhealthy diet, inclusive of consuming excessive saturated fat or sodium, insufficient engagement in regular physical activity, tobacco use, elevated blood cholesterol levels, and other diseases are all important factors that can later influence the appearance of a heart attack or heart attack, as well as factors such as age, family history, and bacterial infections. Therefore, the researcher must deal with various data that have been mentioned or not mentioned, by searching for data and relevant professional backgrounds. It is widely believed that research on heart disease and related scientific studies can be continued from the perspective of medical and statistical knowledge to construct statistical models for prediction and analysis [5]. With the advent of the information age, more sophisticated computer technology and more high-tech instruments have enabled physicians to obtain more accurate and sensitive data and facilitate the advancement of medicine [6]. The NHLBI categorizes factors that contribute to heart disease into two components, lifestyle habits, and other medical conditions, and defines diseases that fit into three or more of these two components and increase the risk of developing heart disease as metabolic syndrome [7].

Moreover, on the website of the Centers for Disease Control and Prevention, high blood pressure, elevated cholesterol levels, and tobacco use are identified as the three primary risk factors associated

with the development of heart disease [8]. This paper remained skeptical about these three key risk factors, so I found some data sets and wanted to analyze them to determine if high blood pressure, high cholesterol, and smoking ranked in the top three out of all the factors that influence them. Since it is influenced by many factors, multiple linear regression problems and ARIMA models can be considered to deal with it.

2. Methods

2.1. Data Source

The data for this literature comes from the Kaggle website, which were collected and organized in collaboration with Dileep and Naveen, published and updated in 2020 with a design of 4,238 people.

2.2. Variable Selection

The data used in this paper counts a total of 4238 people, including those who suffered from heart disease over a ten-year period, of which 1819 were male and 2419 were female (Table 1). The ages of the patients ranged from 32 to 70 years. The data contains 15 variables which are listed in table 1.

Table 1. Logogram of the 15 factors

Elements	Logogram	Qualitative/Quantitative
male	X1	Qualitative
age	X2	Quantitative
education	X3	Quantitative
currentSmoker	X4	Qualitative
cigsPerDay	X5	Quantitative
BPMeds	X6	Qualitative
prevalentStroke	X7	Qualitative
prevalentHyp	X8	Qualitative
diabetes	X9	Qualitative
totChol	X10	Quantitative
sysBP	X11	Quantitative
diaBP	X12	Quantitative
BMI	X13	Quantitative
heartRate	X14	Quantitative
glucose	X15	Quantitative

2.3. Research Protocol

In this paper, a logistic regression model is used, with whether or not one has suffered from heart disease within ten years as the dependent variable (Y) and the 15 factors as the independent variables (X), where 0 means no and 1 means yes. Next, the relationship between the effect of X on Y, that is, the effect of 15 factors on lung cancer, is analyzed [9, 10].

3. Result and Discussion

3.1. Model Results

Figure 1 shows that the order of ranking of the features most closely related to the output variables is first systolic blood pressure, second blood sugar, third age, fourth cholesterol, fifth smoking per day, sixth diastolic blood pressure, seventh hypertension, eighth diabetes mellitus, and ninth blood pressure medication. So we conclude that lipids, blood glucose and age are the first three important features of the data.

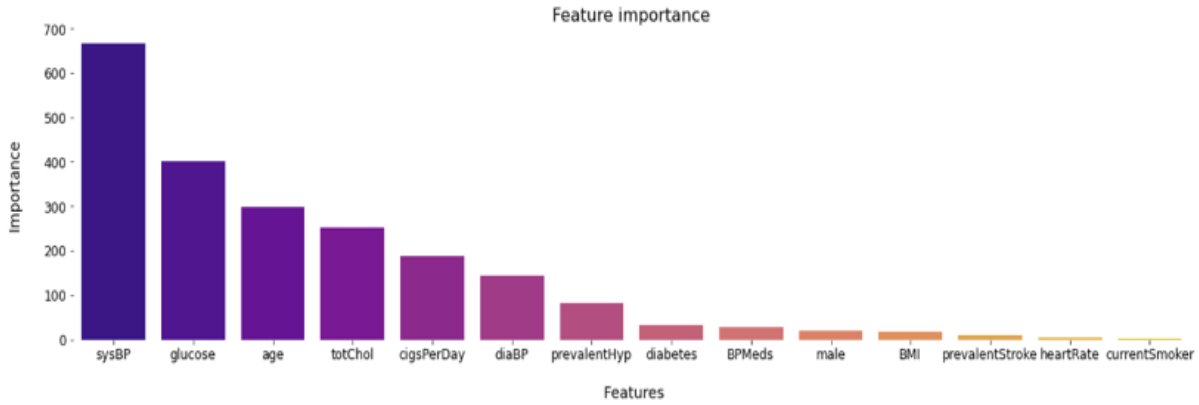

Fig. 1 Arrangement of factors affecting the occurrence of heart disease

As shown in Figure 2, the study introduced various factors that may be associated with heart disease into the model, including gender, age, currentSmoker, cigsPerDay and other elements listed in table 1. By calculation, the final linear regression equation was derived in this study:

$$P = -8.322 - 0.555X1 + 0.063X2 - 0.047X3 - 0.071X4 \ldots\ldots + 0.007X15 \quad (1)$$

Where p represents the probability of heart disease being 1.

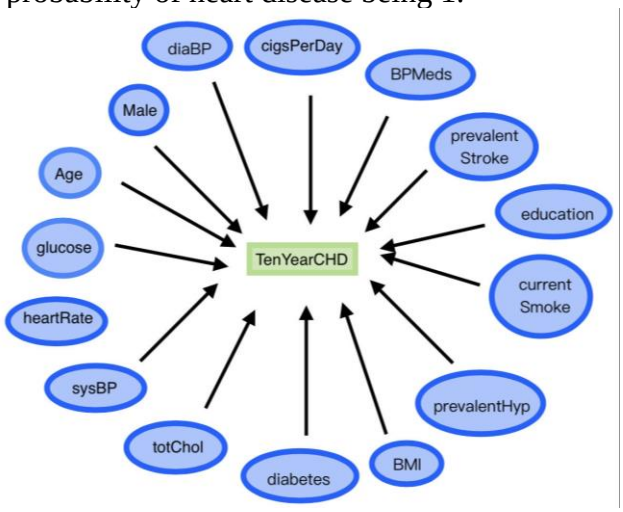

Fig. 2 Variable-related schematic

Table 2. Model results

Variable	coefficient	standard error	P	OR value 95% confidence interval upper limit	lower limit
constant	-8.322	0.715	0.000***	0	0.001
age	0.063	0.007	0.000***	1.052	1.08
education	-0.047	0.049	0.336	0.866	1.051
cigsPerDay	0.018	0.006	0.004***	1.006	1.031
totChol	0.002	0.001	0.039**	1	1.005
sysBP	0.015	0.004	0.000***	1.008	1.023
diaBP	-0.004	0.006	0.521	0.983	1.009
BMI	0.007	0.013	0.605	0.982	1.032
heartRate	-0.003	0.004	0.440	0.989	1.005
glucose	0.007	0.002	0.001***	1.003	1.012
male_1.0	0.555	0.109	0.000***	1.407	2.157
currentSmoker_1.0	0.071	0.157	0.651	0.79	1.46
BPMeds_1.0	0.162	0.234	0.489	0.743	1.862
prevalentStroke_1.0	0.694	0.49	0.157	0.766	5.223
prevalentHyp_1.0	0.235	0.138	0.089*	0.965	1.657
diabetes_1.0	0.039	0.315	0.900	0.561	1.931

As can be seen in Table 2, the modeling yields the conclusion that by determining whether the p-value is greater than 0.05, this study can further infer which factors may be associated with heart disease. On the one hand, diaBP, education, BMI, heartRate, currentSmoker, BPMeds, prevalentStroke, prevalentHyp, and diabetes all have p-values greater than 0.05, meaning that they do not affect heart disease. The p-values for the others, male, age, cigsPerDay, totChol, sysBP, and glucose, were all less than 0.05, i.e., they affect heart disease.

3.2. Discussion

Speaking of effect relationships, the regression coefficient values combined with the marginal effect values show the specific effect relationships: totChol, glucose, sysBP, cigsPerDay, age, and male, for each unit, an increase in them, the probability of having a TenYearCHD of 1.0 is higher than the probability of having a TenYearCHD of 0.0 by 0.233%, 0.715%, 1.552%, 1.809%, 6.551% and 74.211%. The top three factors influencing the occurrence of heart disease over a 10-year period are gender, age, and the average number of cigarettes a person smokes in a day.

It is undeniable that these factors have a significant impact on the likelihood of developing heart disease, as these figures are relatively large, which means that they are of some reference and research value. In contrast to previous studies, they tend to use single-factor specific analyses that are limited to known possible factors for developing heart disease, such as family genetics, smoking, and dietary styles. In contrast, the factors involved in this study are more extensive and comprehensive, which not only effectively avoids errors caused by not controlling for a single variable, but also broadens the idea of future research on heart disease, helping medical personnel to identify more directions for treatment, and earlier detection of possible phenomena of heart disease, to obtain timely treatment.

4. Conclusion

In this study, the author focused on identifying key factors associated with the development of heart disease by analyzing various data. The analysis concluded that the likelihood of developing heart disease is linked to factors such as gender (MALE), age, daily cigarette consumption (cigsPerDay), total cholesterol (totChol), systolic blood pressure (sysBP), and glucose levels.

The findings of this paper demonstrate the significant value of the study and exhibit a variety of strengths. Unlike many previous studies that used one-way analyses, this experiment used logistic regression to consider potential interactions between influences, making the study more comprehensive.

First, a graphical approach was used to visually assess whether exposure to multiple influences increases the risk of heart disease in a population. This approach enhances the clarity and visualization of the results. Second, the study goes beyond known factors such as high blood pressure, cholesterol, and smoking to positively impact heart disease treatment. It emphasizes the importance of considering other potential factors, such as the climate of the living environment and psychological stress caused by patient communication. Further identification and study of these factors could be extremely helpful in understanding the occurrence of heart disease and guiding future research.

The paper concludes that while the model has limitations due to a limited dataset that could introduce errors, it also recognizes that the sample did not cover all age groups and ethnicities. These limitations could lead to discrepancies that could affect the accuracy of the results. Nonetheless, the insights gained from this study set the stage for future research, emphasizing the need for a more comprehensive understanding of the factors that influence heart disease to improve prevention and treatment strategies.

References

[1] Alexander Egbe, et al. Temporal Variation of Birth Prevalence of Congenital Heart Disease in the United States. Congenit Heart Dis, 2015.

[2] Dunlay S M, Weston S A, Jacobsen S J, Roger V L. Risk factors for heart Failure: A Population-Based Case-Control Study. The American Journal of Medicine, 2009, 122(11): 1023–1028.

[3] Van Lennep J E R, et al. Risk factors for coronary heart disease: implications of gender. Cardiovascular Research, 2002, 53(3): 538–549.

[4] Khot U N, et al. Prevalence of conventional risk factors in patients with coronary heart disease. JAMA, 2003, 290(7): 898.

[5] Canto J G, et al. Number of coronary heart disease risk factors and mortality in patients with first myocardial infarction. JAMA, 2011.

[6] Yan Huan, Cao Yongwen, Ye Jiaxin, et al. Factors influencing congenital heart disease in children aged 0-6 years old based on case-control and dose-response relationships. Chinese Journal of Disease Control, 2021, 25(5): 7.

[7] Xia Yuxian, Fu Lijuan, Luo Wenyi, et al. The current situation and influencing factors of disease management difficulties in postoperative caregivers of complex congenital heart disease. Journal of Nursing of the People's Liberation Army, 2022, 039.

[8] Lin Lixiao, Chen Shushu, Chen Sisi. Perioperative nutritional risk assessment and related influencing factors analysis of young children with congenital heart disease. China Maternal and Child Health, 2022, 37(12): 4.

[9] Cao Shouyan, Xu Ying, Li Jing. Analysis of self perceived burden and influencing factors in patients with congenital heart disease. Maternal and Child Care, 2022, 14: 3292-3294.

[10] Zhu Yingjie, Wan Lixin, Zhang Wenyan, et al. A case-control study on the influencing factors of congenital heart disease in children in Jilin region. Chinese Journal of Child Health, 2022, 30(11): 4.