American Journal of Humanities and Social Sciences Research (AJHSSR)
e-ISSN :2378-703X
Volume-07, Issue-02, pp-53-60
www.ajhssr.com
Research Paper Open Access

Examination of the incidence of heart disease in the US. A multivariate logistic regression approach.

Olawande Olayinka Jaiyeola

ABSTRACT: Heart disease is a condition that affects the human heart and blood vessels. Heart disease affects about half of American adults, and it also played a role in the high death rate in the rest of the world. The data extracted from National Center for Health Statistics (NCHS) span from December 2019 to December 2021. The only goal of this study is to look at the risk factors that affect the incidence of heart disease. After that, it will estimate a Youden index to find the best cut-off point and measure how well the multivariate logistic regression model's diagnostic test performed, adding to the body of knowledge. The application of logistic regression yielded the finding that socioeconomic and health risk variables strongly influence the incidence of heart disease. According to the Youden index, the ideal cutoff value is around 52%. Consequently, it is crucial for American adults to monitor their lifestyle, have their BMI, blood pressure, diabetes, and other risk factors for heart disease diagnosed, and then make sure they are receiving adequate treatment to prevent the tendency to develop heart disease, which in turn will lower the death rate brought on by heart disease.

KEYWORDS: *Heart disease, Multivariate logistic regression, Youden index, Health risk factors, socioeconomic factors.*

I. INTRODUCTION

Heart disease develops when plaque accumulates in the arteries and blood vessels that provide blood to the heart, blocking important nutrients and oxygen from reaching the heart and causing it to malfunction (Fryar CD, Chen T-C, and Li X, 2019). Heart diseases are conditions that affect both your heart and your blood arteries; in the United States, about half of all adults have some form of heart disease (Cleveland, 2022). Heart disease is reportedly the leading cause of death both internationally and in the United States, as stated by Cleveland (2022). Beginning with chest pain that is accompanied by shortness of breath, numbness and chilly sensations in the extremities, and in some cases, neck and jaw pain, among other symptoms, are the first signs of heart disease to be seen. Interestingly, about three-quarters of all cardiovascular deaths attributed to coronary heart disease occur in low-income nations all over the world (Wang SY, Tan ASL, and Claggett B, et al 2019). The people of developing countries, particularly those in poorer countries such as Nigeria, may not always reap the benefits of integrated primary health care programs that attempt to identify and treat people who are exposed to risk factors at an early stage, as do the people of affluent countries according to the world health organization (WHO, 2021). However, despite the large number of studies that have dealt with heart disease physiologically, it has been demonstrated that there are no significant works that have extensively dealt with the detection of factors affecting the incidence of heart disease statistically, using a robust logistic regression model with detailed post estimation of the model (Mohammed & Raheem, 2020). In this research study, we will look at the risk factors for heart disease in greater depth. Diabetes, high blood pressure, smoking, family history of heart disease, sexual activity, consumption of alcoholic beverages, BMI and marital status are all considered to be major risk factors that trigger heart disease. It is crucial to come up with the prevention of heart disease and therefore, the primary objective of this study is to model the risk factors that influence heart disease incidence and then estimate a Youden index to determine the optimum cut-off point as well as the measure of the performance of a diagnostic test of the multivariate logistic regression model which will also contribute to the body of knowledge.

II. LITERATURE REVIEW

This section will be presenting the theoretical review and the empirical review.

2.1 Theoretical review

A technique used in healthcare research to develop hypotheses about the experience of sickness is called grounded theory (GT) (Glaser, 1967). The interpretive tradition, one of many to come from the idea of symbolic interactionism, served as the foundation for GT (Glaser, 1967; Chemnitz, 1986). According to some interpretations, GTS provide knowledge about how life is structured and lived; this knowledge has been referred

to as practical knowledge (Chemnitz, 1986). Two fundamental assumptions underpin GT research. First, rather than imposing theory on facts, the study should focus more on how theory emerges from data (Glaser, 1992). The second is that the focus of the study should be on the people's daily activities as they go through crucial life transitions, as well as the internal and external impacts of these activities (Benolie, 1996). Good GT studies so demonstrate how social relationships and surroundings affect how each people interprets their experiences. The analytical strategy used looked at how patients interpreted having a heart problem in the context of their lifestyles. According to the GT principles, an individual's behavior is influenced by the meaning they give to their experiences. Consequently, there were two analytical goals. The first step was to determine the theory behind the meaning patients assigned to their experiences with heart disease. The second step was to provide a framework to explain how the theory of experience affected health behavior. This required determining the theory's tenets, which were applied to explain disease behavior (Figure. 1).

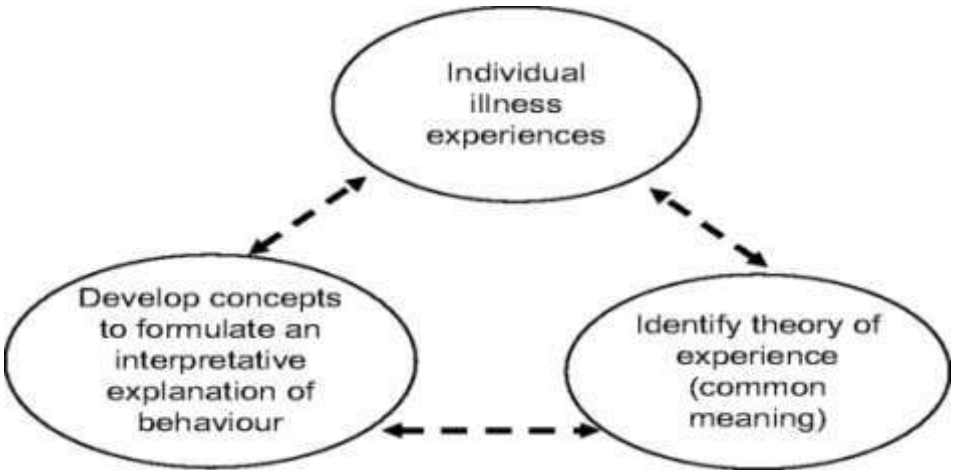

Figure 1 shows how the concepts of GT and symbolic interactionism are used to illustrate the iterative process of recognizing and interpreting the meaning of experience and its impact on disease behaviour.

To help reduce the "researcher effect," two researchers (C.B. and A.W.M.) independently examined the transcripts. A.M. also reviewed several transcripts and the evaluation of the outcomes. Each of the three reviewers has published work that has gone through peer review and has undertaken qualitative research in the past. C.B. has had a big impact on how qualitative primary care research has developed. Themes were examined and discussed to come to a consensus on interpretation and meaning. Maxwell's (1992) framework was also used to enhance the outcomes' validity and the data analysis procedures' dependability. For instance, most patients expressed some amount of anxiety about significant discoveries. The point at which it was thought that the main concepts had been properly probed within the constraints of the study's time and resources were known as "data saturation" (Morse, 2000).

2.2 Empirical review

By using logistic regression, Amir W et al. (2014) determined that the chance factors of age, body mass index (BMI), and systole were connected to high blood pressure. According to Abedin (2016), the results of the logistic regression regarding the fictitious facts that factors like age and smoking reputation relate to one's CVD status. Smokers, however, have a higher risk of developing CVD than non-smokers do at the same age. To obtain an accurate model for analyzing coronary illness, Mythili T et al. (2013) suggested a standard-based model evaluate the correctness of applying rules to the character results of logistic regression and support vector machines in the Heart Disease Database at Cleveland. Rahman et al. (2015) applied logistic regression to examine the sociodemographic factors and risk factors for stroke or myocardial infarction among Bangladeshi hypertension patients.Cardiovascular disease (CVD), which accounts for 73.4% of all fatalities globally, is the leading cause of death. Ischemic heart disease (IHD) is the main reason why people die from cardiovascular illness, according to the Global Burden of Disease (GBD) report from 2017 (Global, 2018). 2.6 million Deaths, or around 35% of all fatalities, are caused by NCDs in Sub-Saharan Africa (SSA), second only to a combination of infectious, maternal, neonatal, and nutritional disorders (CMNNDs). CMNNDs are still fairly common, and since their contributions to overall mortality peaked in the late 1990s and early 2000s, they have been progressively declining in frequency, incidence, and death rates (WHO 2018). Currently, NCDs and communicable diseases carry a twofold burden of disease. A more positive prognosis scenario where CMNND mortality keeps down predicts that NCDs would cause more than half of all fatalities in SSA by 2030 (Mathers &Loncar, 2016). 90% of all deaths in high-income countries (HIC) in Western Europe and North America, where the epidemiologic shift occurred in the early and middle of the 20th century, are caused by cardiovascular

illnesses (CVDs), the most prevalent NCDs (WHO 2018). Obesity, physical inactivity, smoking, hypertension, diabetes, and hyperlipidemia—the usual modifiable atherothrombotic CVD risk factors seen in HIC populations—are now clearly manifesting in SSA (Price, Crampin, Amberbir, Kayuni-Chihana, Music &Tafatatha et al, 2018). However, because most countries either lack data or have inadequate data collection systems that are unreliable enough to allow mounting a proportionate health system response, the actual impact of these risk factors and the issues they generate in SSA are yet unknown (Atun et al, 2017). The number of deaths in Africa attributable to tobacco use increased by over 70% between 1990 and 2016. (Magenta, 2018). It is very alarming that the percentages for diagnosing, treating, and controlling hypertension are still 40%, 35%, and 10–20%, respectively (Ataklte&Erqou, 2015). Comparing estimates for Western Europe, which are 60-70%, 50%, and 30-40% respectively, with those for North America, where >80% of patients are aware of their diagnosis, >70% are receiving treatment, and >55% have reasonable blood pressure control, these results are quite alarming (Niklas et al, 2018). Less than half of Americans with diabetes who take medication have well-controlled diabetes, and more than 40% of people with the disease are not aware of their diagnosis. As a result, SSA still has a sizable portion of the population with undiagnosed diabetes (Price et al, 2018). The SSA, in contrast to the HIC, has only recently become aware of the obesity pandemic, yet the prevalence there is rapidly increasing. Anecdotal and scientific evidence indicates that overweight and obese are idealized body types in several SSA nations and are still regarded as indicators of good health, material success, happiness, and well-being, especially among women. This issue may be made worse by these claims (Belue, 2009). Overall, it is critical to acknowledge the great contributions made by different research on risk factors in numerous nations, which resulted in substantial advancements in United States of America.

III. DATA AND METHODOLOGY

3.1 Data

This study will adopt a quantitative survey research design which was extracted from the National Centerfor health statistics (NCHS) for daily interactive summary health statistics (https://www.cdc.gov/nchs/fastats/heart-disease.htm) from December 2019 to December 2021 and was used for the analysis of the study. The period was determined based on the availability of the dataset.

3.2 Methodology

The method of analysis for this study are summary statistics (mean and standard deviation) and the multiple logistic regression model, which will be used to investigate the most important factors that influence the incidence of heart disease. Everything happens in pairs, according to the dynamics of the global globe, whether it is infected or not, whether it is yes or no, and so on. When determining a dichotomous dependent variable, this tends to follow the logic of the logistic regression model according to the work of Mohammed & Raheem (2020). A multiple logistic regression model is a non-parametric regression used to predict a dichotomous dependent variable (that is, variables that fall into two categories) when there are numerous independent variables, which can be either continuous or categorical. Although the multiple logistic regression model is a nonparametric regression which does not follow the assumptions of the parametric regression model (OLS) post estimation statistics such as goodness of fit, sensitivity and specificity, receiver operating characteristic curve (ROC) and Youden index should be checked to have a robust and reliable model. The improvement in this research is to apply the ROC curve and Youden index to determine the best possible threshold for the incidence of heart disease because several works have been done on the incidence of heart this without the application of the ROC curve and Youden Index for selection of an optimum cut-off point in this regard. Calculating the Youden index is not too difficult. Add together the sensitivity and specificity of a diagnostic test, and then take 100 off the result. When the Youden index is less than 50%, it means the diagnostic test will lack the ability to detect the possibility of having heart disease or not (Heidel, 2022).

3.3 Model specification

The multiple regression model adopted for this study can be adequately specified as:

$$Iny = Log\ (P/1\text{-}P) = \beta_o + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \beta_5x_5 + \beta_6x_6 + \beta_7x_7 + \varepsilon$$

Where β_1 to β_7 are the coefficient estimates of the factors that affect the incidence of heart disease, ε is the stochastic random error and the other components of the model are described below:

Iny: Patient has a heart disease (1 = yes, 0 = No)

x_1: Number of Adult examined daily (in 000')

x_2: Diabetes level(in Millimoles per liter)

x_3: Family history (1 = those Infected, 0 = there is no one infected)

x_4: Daily Alcohol consumption (in percentage)

x_5: Blood pressure level (mmHg)

x_6: Body mass index (BMI) measured in kg/m^2

x_7: Age (in years)

The odd ratio = Exp (B) = e^B which will help us determine the major one/ most critical out of the seven factors that influence the incidence of heart disease.

3.4 Goodness of fit diagnostic for Logistic regression

Hosmer &Lemeshow test is a diagnostic test for goodness of fit for logistic regression model using Chi-square statistic.

Hypothesis for goodness of fit

H_0: The fit is good.

H_a: The fit is not good

Decision rule: Reject the null hypothesis if P-value is less than the significant level and do not reject if otherwise. The generally acceptable significant level in practice are 1% and 5% respectively (Adebanjo et al, 2022).

3.5 Youden Index

Youden's *J* statistic is also called the Youden index and can be written as:

J = Sensitivity + specificity-1 with the two right-hand quantities being sensitivity and specificity. Hence, we can expand the formula as

J = true positives/ (true positives + false negatives) + true negatives/ (true negatives + false positives)

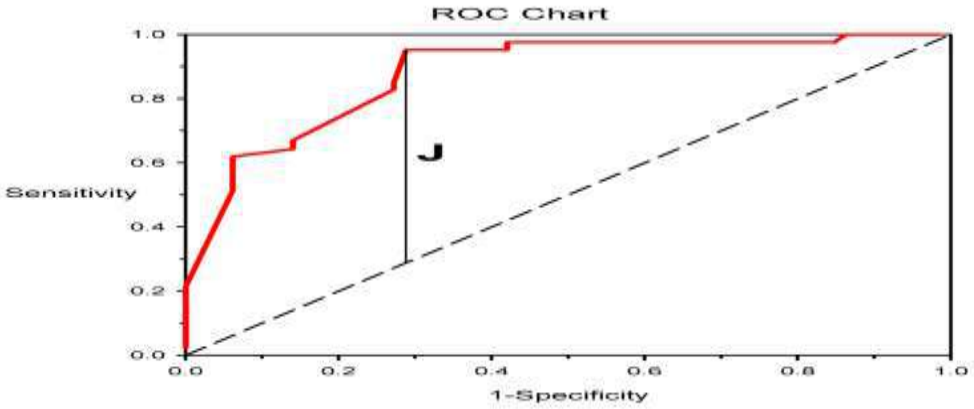

Figure 2: ROC Chart (Youden, 1950).

Solid red: ROC curve; Dashed line: Chance level; Vertical line (J) maximum value of Youden's index for the ROC curve.

IV. RESULT AND DISCUSSION

This section presents the results of the analysis and discussion of the crucial findings of this study.

Table 1: Summary statistics

	N	Mean	Std. Deviation
Adult	768	120.89	31.973
Blood Pressure	768	69.11	19.356
Alcohol	768	20.54	15.952
BMI	768	31.993	7.8842
Diabetes	768	4.719	3.3132
Age	768	33.24	11.760
Frequency	%		
Those with heart disease	34.90		
Those without heart disease	65.10		
Family history	%		
Those diagnosed	49.22		
Those that are not affected by family history	50.78		

Source: Author's computation using Stata Software

Table 1 shows that the average number of adults examined daily for coronary heart disease is about 121 thousand with a variability of about 32 thousand during the period under review. The blood pressure on average

is about 69 mmHg with a variability of about 19 mmHg. The average daily alcohol consumption is about 21% with a variability of about 16%. The average body mass index (BMI) is about 32 kg/m2 with a variability of about 8 kg/m2. The intermediate diabetes level is about 4.7 Millimoles per litre with a variability of about 3.3 Millimoles per litre while the average age of the adults under study is about 33 years with a variability of about 12 years. More so, about 35% of the adults in the US are diagnosed with heart disease while about 65% do not have heart disease. Additionally, about 49% of the patient were traced to family history while about 51% was not.

Table 2: Logistic regression model

Logistic regression
Number of observations = 768
Accuracy of the logistic regression classifier = 0.78
LR chi2(7) = 254.80
Prob > chi2 = 0.0000
Log likelihood = -369.34386
Pseudo R2 = 0.2565

Patient has a heart disease (Yes or No)	Odds ratio	Standard error	Test Statistic	P-value
Adult	1.034	0.004	9.69	0.000
Blood Pressure	0.988	0.005	-2.29	0.022
Alcohol	1.003	0.007	0.34	0.735
Family history	0.710	0.163	-1.49	0.136
BMI	1.089	0.016	5.74	0.000
Diabetes	2.403	0.710	2.97	0.003
Age	1.032	0.008	3.86	0.000
Constant	0.000	0.000	-11.62	0.000

Pearson or Hosmer-Lemeshow goodness of fit (Prob > chi2) = 0.0190

Source: Author's computation using Stata Software

Table 2 shows the results of the estimated multivariate logistic regression. We can see that the overall model Prob > chi2 = 0.0000 which means that the model is statistically significant at a 1% significant level, and this suggests that the incidence of heart disease is significantly associated with health risk factors (such as the blood pressure, alcohol consumption, BMI, and diabetes) and socioeconomic factors like several adults, family history and age. Meanwhile, the accuracy of the logistic regression classifier is 0.78. The model goodness of fit using Hosmer-Lemeshow shows that P>0.01, indicating that the fitted logistic regression model is a good fit for the data. Diabetes has an odd ratio of about 2.4 which is higher than the other risk factors considered under the review. This tells us that Diabetes contributes majorly to the incidence of heart disease than other risk factors. Besides, looking at the contributions of the risk factors associated with heart disease, we can see that number of adults, BMI, Diabetes, and age are statistically significant at a 1% level. In comparison, blood pressure is statistically significant at a 5% level. This implies that several adults, BMI, Blood pressure, Diabetes and age contribute significantly to the incidence of heart disease. This is very consistent with the works of Mohammed & Raheem (2020), Niklas et al (2018) and Price et al (2018).

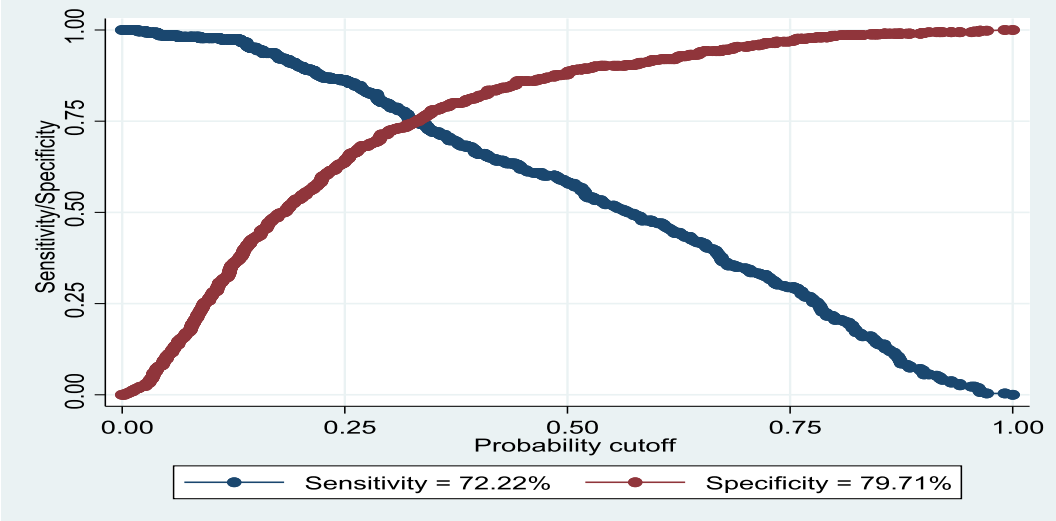

Figure 3: Sensitivity and Specificity curve

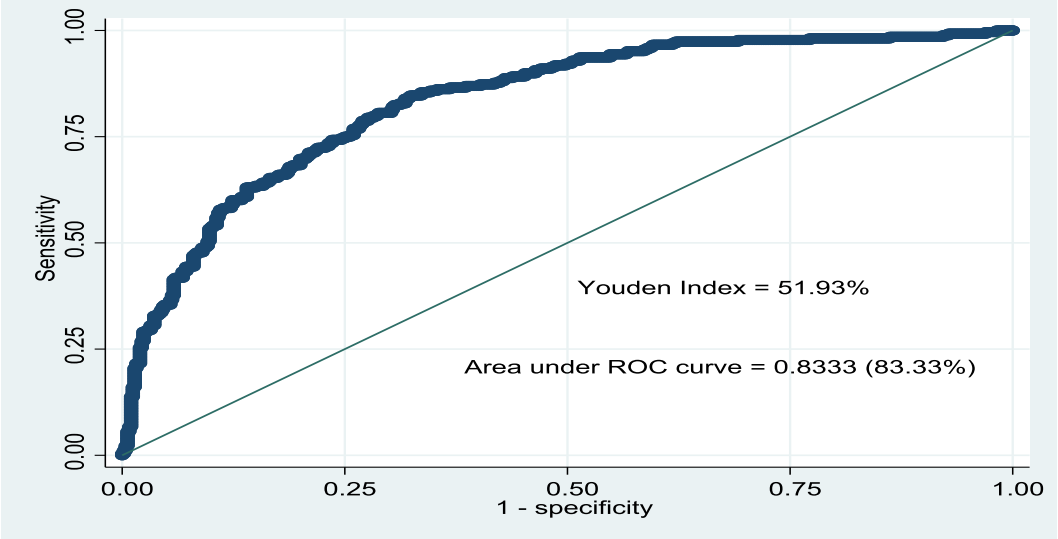

Figure 4: ROC curve showing Youden index and area under the ROC curve.

Figure 3 shows the sensitivity is 72.22%. In comparison, the specificity is 79.71% and figure 4 shows that the area under the ROC curve is 0.83 accounting for about 83% while the Youden index is about 52% which is more than the 50% threshold value. This indicates that the diagnostic test has the potency to detect whether the patients/adults under review possess heart disease or not. The diagnostic test reveals with about 78% accuracy that those not diagnosed with heart disease are more than those with heart disease.

4.1 Discussion of findings

The discussion of notable findings from the result of the analysis of this study is discussed as follows.

According to Table 1, approximately 121 thousand persons were assessed daily for coronary heart disease during the review period, with a variation of roughly 32 thousand. The average blood pressure is roughly 69 mmHg, with a fluctuation of 19 mmHg. An estimated 21% of people drink alcohol on a daily average, with a variance of 16%. The body mass index (BMI) has a range of roughly 8 kg/m2, with the average value being around 32 kg/m2. The average age of the persons being studied is approximately 33 years, with a variability of approximately 12 years, and the intermediate diabetes level is approximately 4.7 millimoles per litre. Furthermore, while 65% of adults in the US do not have cardiac disease, approximately 35% of them have had their condition identified. Furthermore, 51% of the patients had no family history, compared to 49%.

Table 2 displays the estimated multivariate logistic regression findings. We can observe that the entire model (Prob > chi2 = 0.0000), which indicates that the model is statistically significant at a 1% significant level, reveals that the incidence of heart disease is significantly associated with socioeconomic characteristics like multiple adults, family history, and age. The logistic regression classifier's accuracy, meanwhile, is 0.78. P>0.01 in the Hosmer-Lemeshow test for model goodness of fit indicates that the fitted logistic regression model is a good fit for the data. Diabetes has a larger odd ratio than the other risk variables reviewed in the review, which is roughly 2.4. This indicates that compared to other risk factors, diabetes has a greater impact on the incidence of heart disease. Additionally, when we examine the contributions of the risk factors for heart disease, we can find that age, BMI, and the number of adults is statistically significant at a 1% level. In contrast, blood pressure statistically differs at a level of 5%. This suggests that several adult factors, including age, blood pressure, diabetes, and BMI, greatly affect the prevalence of heart disease. This is highly compatible with the works of Niklas et al. (2018), Price et al. (2018), and Mohammed and Raheem (2020). (2018).

Figure 3 reveals that the sensitivity is, in fact, 72.22%. In contrast, the specificity is 79.71 per cent, and figure 4 illustrates that the area under the ROC curve is 0.83, accounting for almost 83% of the data, while the Youden index is approximately 52 per cent, which is higher than the threshold value of 50%. This shows that the diagnostic procedure is capable of determining whether the patients or adults under examination have cardiac disease or not. The diagnostic procedure identifies those who are more likely to have heart disease than those who have it with about 78% accuracy.

V. CONCLUSION AND POLICY IMPLICATION.

Heart disease is an illness that is connected with the heart and bloodstream of humans. Nearly half of the adults in America are affected by this disease and it has also contributed to high mortality in the global world. The sole purpose of this study is to examine the risk factors that influence heart disease incidence and then estimate a Youden index to determine the optimum cut-off point as well as the measure of the performance of a diagnostic test of the multivariate logistic regression model which will also contribute to the body of knowledge. The result

of the applied logistic regression shows that the incidence of heart disease is significantly associated with health risk factors (such as blood pressure, alcohol consumption, BMI, and diabetes) and socioeconomic factors like several adults, family history, and age. In addition, no of adults, BMI, Blood pressure, Diabetes, and age contribute significantly to the incidence of heart disease. Therefore, American adult needs to watch their lifestyle, get diagnosed with BMI, blood pressure, diabetes, and other risk factors that are critically responsible for heart disease, and then ensure adequate treatment is received to prevent the tendency of having heart disease which in turn will reduce death rate that is triggered by hearth disease.

REFERENCES

[1]. Abedin, T., Chowdhury, Z., Afzal, A.R., Yeasmin, F., Turin, T.C. (2016). Application of Binary Logistic Regression in Clinical Research, JNHFB, 5, pp 8-11.

[2]. Amir, W., Mamat, M., & Ali,Z.(2014). Association of hypertension with risk factors using logistic regression, Applied Mathematical Sciences, 8: 2563 – 2572.

[3]. Akpunonu, S. (2018). Reducing burden of heart disease in Nigeria. https://guardian.ng/features/health/reducing-burden-of-heart-disease-in-nigeria/.

[4]. Ataklte F, Erqou S, Kaptoge S, Taye B, Echouffo-Tcheugui JB, Kengne AP. (2015). Burden of undiagnosed hypertension in sub-saharan Africa: A systematic review and meta-analysis. Hypertension (Dallas, Texas: 1979). 65(2): 291–8.

[5]. Atun R, Davies JI, Gale EAM, Barnighausen T, Beran D, Kengne AP, et al. (2017). Diabetes in sub-Saharan Africa: From clinical care to health policy. The Lancet Diabetes & Endocrinology. 5(8): 622–67.

[6]. BeLue R, Okoror TA, Iwelunmor J, Taylor KD, Degboe AN, Agyemang C, et al. (2009). An overview of cardiovascular risk factor burden in sub-Saharan African countries: A socio-cultural perspective. Globalization and Health.

[7]. Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, Chang AR, Cheng S, et al. (2018). Heart Disease and Stroke Statistics – 2018 Update: A Report from the American Heart Association. Circulation. 137(12): e67–e492.

[8]. Benoliel JQ. (1996). Grounded theory and nursing knowledge. Qual Health Res. 1996; 6: 406–428.

[9]. Chenitz CW, Swanson JM. (1986). From Practice to Grounded Theory: Qualitative Research in Nursing. Menlo Park, California: Addison-Wesley.

[10]. Cleveland. (2022). Symptoms, causes and treatment of heart diseases. Available at: https://my.clevelandclinic.org/health/diseases/21493-cardiovascular-disease.

[11]. Fryar CD, Chen T-C, Li X. (2019). Prevalence of uncontrolled risk factors for cardiovascular disease: United States. NCHS data brief, no. 103. Hyattsville, MD: National Center for Health Statistics.

[12]. Glaser B, Strauss A. (1967). The Discovery of Grounded Theory. Chicago: Aldine, 1967.

[13]. Glaser B. (1992). Emergence vs forcing: Basics of Grounded Theory Analysis. Mill Valley, CA: The Sociology Press.

[14]. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: A systematic analysis for the Global Burden of Disease Study 2017. Lancet (London, England). 2018; 392(10159): 1736–88.

[15]. Heidel, E. (2022). Determine the overall benefit of using a diagnostic test. Available at: https://www.scalestatistics.com/youden-index.html.

[16]. Mathers CD, Loncar D. (2016). Projections of global mortality and burden of disease from 2002 to 2030. PLoS medicine. 3(11): 442.

[17]. Magitta NF. (2018). Epidemiology of tobacco use and dependence in Sub-Saharan Africa: A systematic review. J Pulmonol Clin Res. 2(1): 9–15.

[18]. Maxwell JA. (1992). Understanding and validity in qualitative research. Har Educ Rev. 62: 279–299.

[19]. Morse JM. (2000). Determining sample size. Qual Health Res. 10(1), 3–5.

[20]. Mythili, T., Dev Mukherji, M., Padalia, N., & Naidu, A. (2013). A heart disease prediction model using SVM-decision trees-logistic regression (SDL), International Journal of Computer Applications, 68, pp 11-14.

[21]. Mohammed, M.A., Raheem, S.H. (2020). Determine of the most important factors that affect the incidence of heart disease using logistic regression model. Research gate.

[22]. Niklas A, Flotynska A, Puch-Walczak A, Polakowska M, Topor-Madry R, Polak M, et al. (2018) Prevalence, awareness, treatment and control of hypertension in the adult Polish population – Multi-center National Population Health Examination Surveys – WOBASZ studies. Archives of Medical Science: AMS.14 (5): 951–61.

[23]. Price AJ, Crampin AC, Amberbir A, Kayuni-Chihana N, Musicha C, Tafatatha T, et al. (2018). Prevalence of obesity, hypertension, and diabetes, and cascade of care in sub-Saharan Africa: A cross-

sectional, population-based study in rural and urban Malawi. The Lancet Diabetes & Endocrinology. 6(3): 208–22.

[24]. Rahman, M., Nakamura, K., Seino, K., Kizuki, M. (2015). Sociodemographic Factors and the Risk of Developing Cardiovascular Disease in Bangladesh, Am J Prev Med, 48(4): 451-461.

[25]. Wang SY, Tan ASL, Claggett B, et al. (2019). Longitudinal associations between income changes and incident cardiovascular disease: The Atherosclerotic Risk in Communities Study.

[26]. World Health Organization (2018). Global Health Estimates 2016: Deaths by Cause, Age, Sex, by Country and by Region, Geneva: World Health Organization. Available at: https://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.html.