UMYU Scientifica, Vol. 4 NO. 1, March 2025, Pp 379 – 386
https://doi.org/10.56919/usci.2541.038
ISSN: 2955 – 1145 (print); 2955 – 1153 (online)
A periodical of the Faculty of Natural and Applied Sciences, UMYU, Katsina

ORIGINAL RESEARCH ARTICLE

Binary Logistic Regression Modeling for Characterisation of Hypertension and Some Non-Hereditary Risk Factors

Abiodun D. Obaromi[1], Oga Ode[1], Mutah Wadai[2] and Omoha Joy[1]
[1]Department of Biostatistics, Federal University of Health Sciences, Otukpo, Nigeria
[2]Department of Mathematics, Federal University of Health Sciences, Otukpo, Nigeria

ABSTRACT
High blood pressure is a serious concern for public health, which is a crucial area of study as it is a key transmitting agent for coronary artery diseases and other complications. The population of hypertensive individuals is on the increase due to a number of factors, and the rate of prevalence of this morbidity and its terminal effect on humanity is alarming. This study aimed to identify some preventable and controllable risk factors of hypertension. Secondary data (n = 310) were collected from the departments of endocrinology and cardiology of the Teaching Hospital at the University of Ilorin (UITH), Kwara State. A binary logistic regression model was fitted on factors like age, gender, and smoking. Results showed that age, gender, alcohol consumption, height, BMI, and hours of daily rest are positive predictive factors for hypertension, where gender, working status, and BMI are statistically significant response variables (p-value < 0.05). Also, the odds of developing hypertension with respect to gender, working status, and BMI are 4.25, 0.55, and 7.09, respectively, when other predictor variables are held constant. The receiver operating characteristics, ROC, which measures the sensitivity and specificity of the model (AUC = 0.7141), indicated the probability that the model is more likely to assign a higher probability to a positive case (hypertension) compared to a negative case. In conclusion, the developed model, being one of the recent studies to examine the predictive power of a model, can be adopted for better precisions of the explanatory risk factor variables for hypertension and ultimately help to reduce its prevalence.

ARTICLE HISTORY
Received December 14, 2024
Accepted March 27, 2025
Published March 31, 2025

KEYWORDS
Hypertension, diabetes, logistic, Wald test, Odds

INTRODUCTION

Hypertension or high blood presuer is a vital public and community health concern and an essential study area because of its significant frequent occurrence and importance as a key transmitting factor for heart diseases and associated effects. A number of reasons are contributing to the rise in the large number of individuals who suffer from high blood pressure. Furthermore, it is concerning how common this illness is and how it may ultimately affect humanity (Forouzanfar et al. 2016). Finding risk variables that affect this illness's development, maintenance, and prognosis becomes a top priority for researchers. Globally, hypertension is a major contributor to early death and heart disease (Campbell, N. R. et al. 2014). The average blood pressure (BP) across has either stayed the same or decreased marginally throughout the past four decades, based on a widespread use of anti-hypertensive medications and therapies. Compared to other countries, the prevalence of hypertension has risen, particularly among LMICs. About 1.39 billion individuals worldwide, or 31.1% of the entire population, lived with hypertension in 2010, according to statistics. Adult high blood pressure was prevalent when compared to high-income nations (28.5%, 349 million people) with countries with low or middle incomes (31.5%, 1.04 billion people), according to Katherine et al. (2020).

Studies showed that the majority of grownups worldwide suffer from hypertension, one of the prevalent non-communicable disorders. According to standard medical guidelines, hypertension or high blood pressure is when blood pressure exceeds 140 over 90 mmHg (millimeters of mercury). This indicates that the diastolic measurement of the the heart's pressure when it rests and fills with blood exceeds 90 mmHg, and/or the systolic measurement, which monitors the pressure of the heart when it circulates blood throughout the body system, is over 140 mmHg. The force exerted on the blood vessel walls by the blood is known as blood pressure, and its magnitude is measured by both the cardiac output and blood vessel resistance. The American Heart Association, AHA, has set the ranges for blood pressure as follows: a hypertensive crisis, which

Correspondence: Abiodun D. Obaromi. Department of Biostatistics, Federal University of Health Sciences, Otukpo, Nigeria. ✉ obaromi.davies@fuhso.edu.ng.

How to cite: Obaromi, A. D., Ode, O., Wadai, M., & Omoha, J. (2025). Binary Logistic Regression Modeling for Characterisation of Hypertension and Some Non-Hereditary Risk Factors. *UMYU Scientifica*, 4(1), 379 – 386. https://doi.org/10.56919/usci.2541.038

is a medical emergency, occurs when blood pressure surpasses 180 systolic or 110 diastolic. Normal blood pressure is less than 120 systolic and less than 80 diastolic; pre-hypertension is 120-139 systolic or 80-89 diastolic; stage 1 high blood pressure (high blood pressure) is 140-159 systolic or 90-99 diastolic; and stage 2 high blood pressure (hypertension) is 100 or higher diastolic or 160 systolic.

As one of the leading public health and medical complications in treating and preventing hypertension or high blood pressure, healthcare workers need to acquire more knowledge of the significance of systolic blood pressure (Boutayeb et al., 2005). An estimated 4.5% of the world's disease burden is attributed to hypertension, which is as common in many developing nations as it is in rich ones. Globally, hypertension has been linked to seven million preventable deaths (Syer et al., 2010). It has been more evident in the last ten years that stroke, coronary heart disease, and kidney damage are all caused due to high blood pressure. All willing residents of two villages in Penang, Malaysia, who were at least eighteen years old, participated in a cross-sectional survey. Additionally, it has been reported that one billion individuals suffer from hypertension and that nine million people globally die from high blood pressure (Manandhar et al., 2016). In Nepal, 36% of adults aged 25 and older had high blood pressure, compared to 26.2% of women. The greatest number of patients had stage I hypertension (diastolic: 33.3%, systolic: 39.1%). Using logistic regression analysis, variables such as age, intake of alcohol, smoking, and a family record of hypertension were all considered informative at the five percent significance level. By lowering blood pressure, lifestyle changes can either prevent or postpone the onset of hypertension. Thus, it is possible to avoid and regulate high blood pressure.

Hypertension that is not triggered by an alternative disease or condition is known as primary hypertension or vital hypertension. This is more prevalent compared to secondary high blood pressure, which has a known cause. The system of renin and angiotensin, which is a hormone regulator of blood pressure and volume, and blood plasma volume, are two of the many elements that often contribute to main high blood pressure, this likely has more than one source. Environmental features that involve the previously mentioned lifestyle-related problems also have an impact on primary hypertension. Because it is secondary to another issue, secondary hypertension has distinct causes. Primary aldosteronism, a hormone illness, is one example that causes an imbalance between potassium and sodium levels, which is now believed to be one of the most common causes of treatment-resistant hypertension.

According to World Health Organization report (WHO 2002), the worldwide prevalence rate in 2000 was about 26%, affecting around 1 billion people, and by 2025, it is projected that 1.56 billion individuals globally will suffer from hypertension. The global epidemic of cardiovascular diseases, including hypertension, is on an upward scale mostly as a result of the rapid health changes triggered by social, economic, technological, and changes in lifestyles in poor and medium-income countries (LMICs), which have seen a significant rise in their numbers. (Galav et al., 2015). Primary and secondary planning and prevention of hypertension depend on estimating the frequency in urban and rural populations (Galav et al., 2015). Additionally, it is mentioned that hypertension is a silent killer because, in its early stages, very few symptoms are apparent until a major medical issue like a cardiac attack, stroke, or chronic kidney disease occurs. (Shikha et al., 2017). This is most likely the cause of the elevated death rate associated with this illness occurrence.

Numerous studies have been conducted on hypertension, either as a separate condition or in conjunction with other conditions. The majority of these studies focused on determining the disease's prevalence. A small number of people have attempted to model this disease based on its different risk factors. It is still important to note, though, that additional work needs to be done on modeling out the different categories so that the important non-hereditary risk factors or factors can be discovered in order to contribute to tangible decision-making in the effort to lower the imminent danger of this illness. A number of studies have been done on the global prevalence of high blood pressure, and sub-Saharan Africa has been found to have one of the highest prevalence rates, with an average of 3.2 million deceased annually. Nwoga (2023) assessed the risk factors for hypertension amongst staff of a tertiary institution in Nigeria, utilising some modifiable and non-modifiable risk factors study showed that both non-modifiable (family history of hypertension) and modifiable risk factors (intake of alcohol and exercise) were significant risk factors for hypertension. Nonetheless, there is still a dearth of clear and comprehensive evidence on the primary non-hereditary risk factors or factors for hypertension in Nigeria through in-depth modeling with some level of predictive model accuracy in Nigeria. Thus, by fitting a logistic regression model on the hypertension status of the cohorts under consideration given its risk factors and isolating identified risk factors for prediction and likelihood occurrence, this study aims to identify and model for prediction and control of hypertensive cases. The study will also use logistic regression analysis to determine the risk variables for hypertension.

MATERIALS AND METHODS

Data Source

One of the goals of this research is to assess some of the non-hereditary risk variables for hypertension among a number of variables, including height, index of body mass, diabetes, age, gender, status of employment and smoking, alcohol usage, length of treatment, and estimated hours of rest per day. The University of Ilorin Teaching Hospital provided secondary data. The data utilized in this study is relevant as the information is recent and may better reflect the current state of hypertension; the proximity of this teaching hospital is thought to be more suitable in connection with financial cost implications, time required, and, to some extent, the total objective of the study.

Methods

Due to the difficulty and type of data used, binary logistic regression was utilized in this investigation. Based on one or more independent variables, which may be continuous or categorical, a binomial logistic regression forecasts the likelihood that an observation will fall into one of two categories of a dichotomous dependent variable. Let Y stand for the binary answer variable, which can have two alternative outcomes: 0 (failure) and 1 (success).

The likelihoods P(Y = 1) = π and P(Y = 0) = (1-π) define the distribution of Y. The parameter π and index n define an n independent observations, the number of successes which follows a binomial distribution, n separate, identical trials with two likely results—"success" and "failure"—are the sources of the binomial distribution of categorical data. Specifically, the result of one trial has no bearing on the result of another.

These are frequently referred to as Bernoulli trials, where y is distributed with a binomial density function with parameter and index n, π with probability, assuming identical and n independent trials:

$$P(y) = \binom{n}{y}\pi^{y}(1-\pi)^{n-y} \quad for\ y = 0, 1, 2, . . ., n$$

$$= \frac{n!}{y!\,(n-y)!}\pi^{y}(1-\pi)^{n-y}\ for\ y = 0, 1, 2, . . ., n \tag{1}$$

The logit or log transforms—the natural logarithm of the likelihood that an event will happen —are the link functions in logistic regression. In other words:

$$\boldsymbol{\pi(x) = \frac{e^{\sum_{k=0}^{k} x_{ik}\beta_k}}{1 + e^{\sum_{k=0}^{k} x_{ik}\beta_k}}}$$

and

$$\log\left(\frac{\pi i}{1-\pi i}\right) = \sum_{k=0}^{k} x_{ik}\beta_k \quad \text{i} - 1, 2, . . ., \text{N} \tag{2}$$

Binary logistic regression with single explanatory variable

The probabilistic element of the (success, failure) outcomes in the logistic regression is modeled by a binomial distribution. The regression model is known as the logit model, where the logit function as the link function is represented as logit (π) and is the logarithm of the probability of success. The values of π fall between 0 and 1. The model is provided as follows:

$$\ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_o + \beta_1 X \tag{3}$$

From the above, we deduced that:

$$\left(\frac{\pi(x)}{1-\pi(x)}\right) = e^{\beta o + \beta 1\, X}$$

$$\pi(x) = e^{\beta o + \beta 1\, X} - \pi(x)\beta o + \beta 1\, X$$

$$\pi(x)\left(1 + e^{\beta o + \beta 1\, X}\right) = e^{\beta o + \beta 1\, X}$$

$$\pi(x) = \frac{e^{\beta o + \beta 1\, X}}{1 + e^{\beta o + \beta 1\, X}} \tag{4}$$

The log odds model called the *logit*has the linear relationship as follows:

$$\text{logit}\ \frac{\boldsymbol{\pi(x)}}{\boldsymbol{1-\pi(x)}} = \ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 X \tag{5}$$

By evaluating the effect of increasing X by one unit, the equation for logistic regression follows as:

$$\ln\left(\frac{\boldsymbol{\pi'(x)}}{\boldsymbol{1-\pi'(x)}}\right) = \beta_0 + \beta_1(X+1)$$

$$\ln\left(\frac{\boldsymbol{\pi'(x)}}{\boldsymbol{1-\pi'(x)}}\right) = \beta_0 + \beta_1 X + \beta_1) \tag{6}$$

By deducting (1) from (2), we have:

$$\ln\left(\frac{\boldsymbol{\pi'(x)}}{\boldsymbol{1-\pi'(x)}}\right) - \ln\left(\frac{\boldsymbol{\pi(x)}}{\boldsymbol{1-\pi(x)}}\right) = \beta_0 + \beta_1 X + \beta_1 - \beta_0 - \beta_1 X)$$

$$\beta_1 = ln\left(\frac{\frac{\boldsymbol{\pi'(x)}}{\boldsymbol{1-\pi'(x)}}}{\frac{\boldsymbol{\pi(x)}}{\boldsymbol{1-\pi(x)}}}\right)$$

$$\beta_1 = ln\left(\frac{odds'}{odds}\right) \tag{7}$$

This is to say that the log of the ratio of the odds at X+1 and X is β_1 and can also be stated as:

$$e^{\beta_1} = \left(\frac{odds'}{odds}\right)$$

The ratio of the odds of an increase of one unit in X to the initial probabilities at X is known as the regression coefficient or β1. The logit transform, or the log odds of the chance of success, is equivalent to the linear component in the logistic regression model:

$$log\left(\frac{\pi_i}{1-\pi_i}\right) = \sum_{k=0}^{k} x_{ik}\beta_k \quad i = 1, 2, 3, \quad , N \tag{8}$$

Parameter Estimation

Estimating K+1 unknown parameters, βs, in eq. 8 is the goal of the logistic regression. This is achieved by using the maximum likelihood estimation method, which entails identifying the set of criteria for which the observed data has the highest probability. The dependent variable's probability distribution is used to calculate the maximum likelihood equation.

Overall model estimations: The likelihood ratio test

A logistic model that outperforms the null, intercept-only model, which lacks a predictor variable, is expected to provide more appropriateness for the data. The likelihood ratio test is used for the overall significance of the independent variable coefficients of the model. The test is dependent on the statistic "G" if given the null hypothesis as:

$$H_0 : \beta_1 = \beta_2 = \cdots \beta_k = 0)$$

and the G statistic is obtained as follows:

$$G = \chi^2 = (-2\ ln\ likelihood\ of\ null\ model) - (-2\ ln\ likelihood\ of$$

$$G = -2ln\frac{likelihood\ of\ null\ model}{log\ likelihood\ of\ \ model\ with\ the\ variables} \tag{9}$$

Given the quantity of covariates in the logistic regression equation, the degrees of freedom, K, the distribution of "G" is Chi-squared, which indicates the amount to which each explanatory or independent variable influences the explanatory variable (Bewick et al., 2005). When the test for model significance is less than the standard α=0.05, then the p-value suggests that the alternative hypothesis is rejected for the entire model. This indicates that at least one of the explanatory variables is supported by evidence to play a role in outcome prediction.

Goodness of fit statistics: Hosmer – Lemeshow test

In subcategories of model populations, the test that serves to assess if the observed percentage of events matches the expected likelihood of occurrence is the Hosmer-Lemeshow test. (Hosmer et al., 2000). In order to evaluate the Hosmer-Lemeshow test, the projected likelihoods are divided into deciles (10 sets according to percentile ranks), and a Pearson Chi-square is then calculated by using a 2-by-10 table in comparing the observed and expected frequencies. The following is the test statistic:

$$H = \sum_{g=1}^{10}\frac{\left(O_g - E_g\right)^2}{E_g} \tag{10}$$

Where *Og* and *Eg* stand for the risk deciles group's observed and predicted events, respectively, the test statistic has eight degrees of freedom and is asymptotically distributed to a Chi-square statistic. Small values (with a large p-value near 1) indicate an outstanding fit to the data. This means that the model as a whole fits the data well. A poorly fitted data is indicated by large values (p-value < 0.05).

Wald statistic

The statistic employed to evaluate the logistic regression coefficients' significance is the Wald statistic (Bewick et al. 2005). When testing a single parameter, this test is typically conducted.

The Wald statistic is calculated as:

$$W_j = \frac{\hat{\beta}_j^2}{[SE(\hat{\beta}_j)]^2} \tag{11}$$

Where $\hat{\beta}_j$ represents theestimated coefficient and SE ($\hat{\beta}_j$) is its standard error. Assuming the null hypothesis $H_o : \beta_J = 0$ and the distribution of W_j follows a Chi-square.

Model Validation and Sensitivity Analysis

While sensitivity analysis evaluates how significantly the model's output differs with variations in input parameters, model validation examines how well a model fits the data and its prediction power.

RESULTS

Predictor variables including age, gender, working status (WS; sedentary or non-sedentary), smoking status (S), alcohol consumption status (CA), body mass index (BMI), duration of treatment (DT), diabetes status (DB) (yes or no), estimated hours of rest per day (HER), and height (H) comprise the data used for the study.

The model to fit for the data is of the following form:

$$\pi_{ij} = \frac{\exp(\beta_0+\beta_1 Age+\beta_2 G+\beta_3 WS+\beta_4 S+\beta_5 CA+\beta_6 DT+\beta_7 EHT+\beta_8 BMI+\beta_9 H+\beta_{10} DB)}{1+\exp(\beta_0+\beta_1 Age+\beta_2 G+\beta_3 WS+\beta_4 S+\beta_5 CA+\beta_6 DT+\beta_7 EHT+\beta_8 BMI+\beta_9 H+\beta_{10} DB)}$$

and the

$$\text{logit}(\pi_{ij}) = \log\left[\frac{\pi_{ij}}{1-\pi_{ij}}\right]$$

Table 1: Summary of Case processing

Unweighted Cases		N	Percent
Selected Cases	Included in Analysis	310	100
	Missing Cases	0	0
	Total	310	100
Unselected Cases		0	0
Total		310	100.0

Binary logistic regression model of case-controlled

The data set utilized in this composition has nine (10) predictor variables. This type of regression is to be selected based on the binary feature of the response variable once the underlying assumptions are satisfied.

Table 2: Table of the null regression (intercept only model) with intercept of 0.339, odds of 1.403, and a p-value of 0.003.

	B	S.E.	Wald	Df	Sig.	Exp (B)
Constant only	.339	.115	8.640	1	.003	1.403

Table 3: Variables used in the equation

							95% C.I.for EXP(B)	
	β	S.E.	Wald	Df	Sig.	Exp(β)	Lower	Upper
Age	.004	.015	.062	1	.804	1.004	.974	1.034
Gender	1.446	.343	17.801	1	.000	4.246	2.169	8.309
Working Status	-.595	.278	4.600	1	.032	.552	.320	.950
Smoking Status	-2.862	2.785	.000	1	.999	.000	.000	.
Alcohol Consumptn	.878	.605	2.110	1	.146	2.407	.736	7.876
Height	.139	1.791	.006	1	.938	1.149	.034	38.467
Body mass index	.958	.024	5.731	1	.017	2.606	1.011	3.111
Diabetes status	-.178	.269	.437	1	.508	.837	.494	1.419
Hour Of rest daily	.195	.311	.394	1	.530	1.215	.661	2.236
Duration of Trtmt	-.015	.034	.184	1	.668	.985	.921	1.054
Constant	1.539	2.785	.042	1	.039	4.662		

Table 4: The reduced model for hypertension with only significant response variables.

							95.0% C.I.for EXP(B)	
	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Gender	1.446	.343	17.801	1	.000	4.246	2.169	8.309
Working Status	-.595	.278	4.600	1	.032	.552	.320	.950
Body mass index	1.958	.024	5.731	1	.017	7.085	1.011	3.111
Constant	1.539	2.785	.042	1	.039	4.662		

The table above is an extract of the significant explanatory variables in the omnibus model. Gender, working status and body mass index are the significant predictors, based on the data used for the analysis.

The Logit (π_{12}) is:

$$\log\left[\frac{\pi_i}{1-\pi_i}\right] = 1.539 + 1.446 * \mathrm{G} - 0.595 * \mathrm{WS} + 0.058 * \mathrm{BMI}$$

Odds

The odds of developing hypertension is given as the ratio of π to $(1 - \pi)$, that is Odds (hypertension) is

$$\log\left[\frac{\pi_i}{1-\pi_i}\right] = \exp\left(1.539 + 1.446 * \mathrm{G} - 0.595 * \mathrm{WS} + 0.058 * \mathrm{BMI}\right)$$

The following is an extraction of the chances for each predictor variable from the aforementioned simplified model table:

Gender [Male (1) and Female (0)]. Odds of developing hypertension with respect to gender when other predictor variables are held constant is:

Exp (1.446) = 4.246 ≈ 4

Working Status [Sedentary (1) and non-sedentary (0)]. The odds of developing hypertension with respect to working status when other predictor variables are held constant is:

Exp (-0.595) = 0.552 ≈ 1

Body Mass Index: The probability of acquiring hypertension with regard to body mass index while other predictor variables are held constant is:
Exp (1.958) = 7.085 ≈ 7

Table 5: Model coefficients using Omnibus tests.

	Chi-square	Df	Sig.
Step	219.759	1	.000
Block	154.894	10	.000
Model	154.894	10	.000

Table 6: Model summary

Step	-2 Log likelihood	Cox &Snell R Square	Nagelkerke R Square
	64.865	.631	.812

Table 7: Full Model Assessment

Model	Model fitting criteria	Likelihood Ratio Test		
	-2Log Likelihood	Chi square	Df	Significance
Intercept only	219.759			
Final	154.894	64.865	10	0.000

Table 8: Classification of the Observed and Predicted hypertensive cases

	Predicted		
	HD		
Observed	YES	NO	Percentage Correct
YES	229	9	96.2
NO	46	26	36.1
Overall Percentage			82.3

Model Validation and Sentitivity

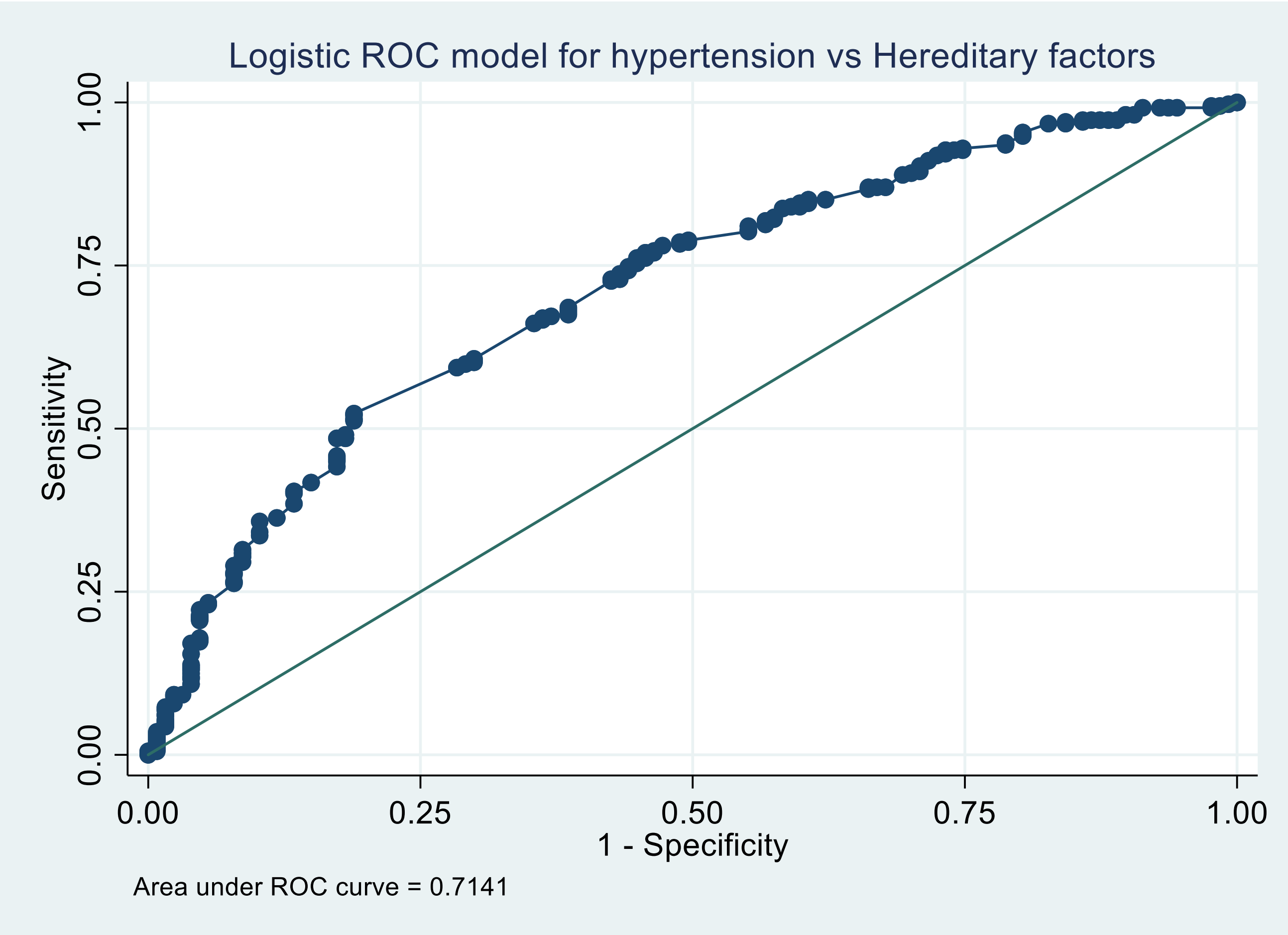

Figure 1: Logistic ROC model for hypertension vs hereditary Factors

Table 6 tells how much variability in hypertension (response variable) is explained by the model-this is an equivalent of R^2 in multiple linear regression.

Both model validation and sensitivity analysis are efforts to assess the appropriateness of a specific model specification and to identify the reliability of the inferences being made from it. The AUC = 0.7141 (71.4%) reflects the model's ability to distinguish between hypertension and the explanatory variables.

DISCUSSION

The null regression, or intercept-only model, is presented in Table 2 with a p-value of 0.003, an intercept of 0.339, and chances of 1.403. Given that α=0.05 exceeds the p-value (0.003), the null hypothesis is not accepted in favour of the alternative, and a conclusion is formed that the intercept-only model is significant. As indicated in the same table, the probability of getting hypertension naturally is 1.403. With all the predictor variables included, Table 3 displays the logistic regression of high blood pressure on its transmitting and risk agents. The simplified model for hypertension, which only includes significant predictor factors, is shown in Table 4. It demonstrates that gender (based on females), body mass index, and employment position (based on sedentary) are the main associated risks for hypertension.

The frequency and prevalence of hypertension rose with age from 40% (ages, 43–46 years) to 93% (ages, 91–94 years). For hypertensive people, the prevalence of uncontrolled hypertension was more in men (33%) than women (23%) at ages 43 to 46 years but became higher in women than men beginning at ages 61 to 64, with 56% of women and 40% men having unrestrained hypertension at ages 91 to 94 (Yeo et al., 2024). Our study in agreement showed that the odds of developing hypertension by gender and age, which can be seen from Table 4, indicates that four men are likely to be in danger of developing hypertension for every female (Beta = 1.45) and statistically significant (p-value < 0.005) age is also a predictive positive factor (β = 0.004).

Likewise, if every other predictor variable stays the same, the likelihood result for hypertension in relation to working status is 0.552. There is a 0.552 chance that someone who leads a sedentary lifestyle will acquire hypertension for every one person who does. This indicates that persons who lead active, non-sedentary lifestyles have a lower risk of having hypertension. Results from the sleep heart health study gave fascinating proofs that individuals sleeping between 8 and 9 hours per night have a similar risk of developing hypertension as those sleeping between 6 and 7 hours (19% increase in risk for both circumstances) (Grandner M et al., 2018). Our study corroborated with this as hours of sleep was found to be a protective risk factor for developing hypertension. Body Mass Index is found to be associated with blood pressure and vital capacity in medical students (Song et al., 2023), which aligned with our study that the odds of developing hypertension with respect to body mass index when other predictor variables are held constant is 7.085. In other words, those with a growing body mass index have a seven-fold increased risk of developing hypertension. Investigation of observational data discovered an inconsistent association between smoking and the risk of hypertension, where a lower risk was observed among current smokers when compared to non-smokers (Jareebi, 2024). In agreement, our study revealed that smoking status has a negative (OR = -2.862) and non-significant (p-value > 0.05) association with hypertension.

Tests of Model Coefficients using the omnibus test are shown in Table 5. It demonstrates that the null model was significant at every stage of the iteration used to fit the optimal model, that the model fits well once the explanatory variables have been added, and that at least one explanatory variable helped the model fit well. The degree of variability in the response variable, hypertension that the model can account for is shown in Table 6; this is the multiple linear regression equivalent of R2. Based on the model, we discovered that the explained variation in diabetes varies from 63.1% to 81.2%. According to some authors, it is preferable to state the Nagelkerke R2, which in this instance is 81.2%.

A summary of the likelihood ratio test may be seen in Table 7. The -2 Log Likelihood statistic for hypertension was lowered by 291.759 - 154.894 = 64.865 when the predictor variables were added. The lower the deviation value, the better the model fits. If the logistic model performs better than the null model, which lacks a predictor variable, it is regarded as providing a more accurate fit to the data..The effectiveness of the expected classification in comparison to the actual classification is evaluated in Table 8. About 82.3% were suitably classified of all the cases. The study also provides important information such as sensitivity, specificity and Receiver Operating Characteristic (ROC) curve for model validation. Sensitivity analysis and model validation are linked in that they are both attempts to assess the appropriateness of a particular model specification and appreciate the strength of the conclusions drawn from such a model (Salciccioli et al., 2016). Our study result showed that the AUC = 0.7141 (71.4%) reflects the model's ability to distinguish between hypertension and the explanatory variables.

As experimentally expected for any research study, the limitations of this study though not exhaustive, include that the sample size is small (n = 310) as the larger the sample size, the better the precision. Another limitation is that this study was constrained to a few individuals within a space. Better results should involve population study levels and bigger spaces or regions. Novel studies to investigate shared component models of hypertension and other non-communicable and hereditary diseases should be given attention in further research, as some concomittant health conditions may not be impossible.

CONCLUSION

This paper highlights the broad scope of issues related to hypertension. Everyone can contribute to lessening the impact on our planet. According to this study, one known risk factor for hypertension is an increasing or uncontrolled body mass indexAs a result, we advise people to pay attention to their body mass index, which always entails keeping an eye on and controlling their weight (body mass index (BMI) is computed by dividing weight in kilograms over height measured in meters squared).

This study also shows that four men are at risk of developing hypertension for every female who does. This means that the proportion of men to women is 1 to 4. Therefore, we advise men to take medical advice extremely seriously, stay away from all risk factors associated with prognosis as much as possible, and follow medical advice when and when it is offered. Smoking and excessive alcohol consumption should be discouraged to reduce the risk of hypertension. A sedentary lifestyle will always increase a person's risk of developing this disease, as one of the main risk variables identified in this study is working status (related to sedentary). For such individuals, physical activity is advised in lieu of sedentary behavior. According to the study, hypertension in the population under investigation was predicted by modifiable and non-modifiable risk variables. These two areas and many more should be the focus of risk factor screening and education initiatives.

ACKNOWLEDGMENTS

We appreciate the administrators, employees, and members of the statistics and research department, departments of endocrinology and cardiology of the University of Ilorin Teaching Hospital, (UITH), Kwara State for the data release.

DECLARATIONS

Availability of data and Materials

The cardiology and endocrinology departments of the University of Ilorin Teaching Hospital in Kwara State, Nigeria, provided the data for this study.

Conflicting interests

The authors declared no conflicts of interest.

Finances

This study was conducted without any funding.

Contributions of the Author

Davies was responsible for the study's conception, formulation, drafting, and data analysis. Wadai, Ode, and Omoha proofread the manuscript to ensure accuracy.

REFERENCES

Bewick, V., Cheek, L., & Ball, J. (2005). Statistics review 14: Logistic regression. *Critical Care, 9*(1). [Crossref]

Boutayeb, A., & Boutayeb, S. (2005). The burden of non-communicable diseases in developing countries. *International Journal for Equity in Health, 4*(1), 2. [Crossref]

Campbell, N. R., Lackland, D. T., & Niebylski, M. L. (2014). High blood pressure: Why prevention and control are urgent and important: A 2014 fact sheet from the World Hypertension League and the International Society of Hypertension. *Journal of Clinical Hypertension, 16*(8), 551–553. [Crossref]

Chobanian, A. V., Bakris, G. L., & Black, H. R. (2003). The seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure: The JNC 7 report. *JAMA, 289*(19), 2560–2572. [Crossref]

Contractor, A., Sarkar, B. K., Arora, M., et al. (2014). Addressing cardiovascular disease burden in low and middle income countries (LMICs). *Current Cardiovascular Risk Reports, 8*, Article 405. [Crossref]

Ezzati, M., Lopez, A. D., & Rodgers, A. (2002). Selected major risk factors and global and regional burden of disease. *The Lancet, 360*(9343), 1347–1360. [Crossref]

Forouzanfar, M. H., Liu, P., Roth, G. A., et al. (2017). Global burden of hypertension and systolic blood pressure of at least 110 to 115 mm Hg, 1990–2015. *JAMA, 317*(2), 165–182. [Crossref]

Galav, A., Bhatanajar, R., Megharal, S. C., & Jain, M. (2015). Prevalence of hypertension among rural and urban population in southern Rajasthan. *National Journal of Community Medicine, 6*(2), 174–178. https://njcmindia.com/index.php/file/article/view/1159

Grandner, M., Mullington, J. M., Hashmi, S. D., Redeker, N. S., Watson, N. F., & Morgenthaler, T. I. (2018). Sleep duration and hypertension: Analysis of >700,000 adults by age and sex. *Journal of Clinical Sleep Medicine, 14*(6). [Crossref]

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). John Wiley & Sons. [Crossref]

Jareebi, M. A. (2024). The association between smoking behavior and the risk of hypertension: Review of the observational and genetic evidence. *Journal of Multidisciplinary Healthcare, 17*, 3265–3281. [Crossref]

Manandhar, N., & Raman, T. P. (2016). Risk factors of hypertension: Logistic regression analysis. *SCIREA Journal of Health, 1*(1). http://www.scirea.org/journal/PMH

Katherinem, Mills, K. T., Stefanescu, A., & He, J. (2020). The global epidemiology of hypertension. *Nature Reviews Nephrology, 16*(4), 223–237. [Crossref]

Nwoga, H. O. (2023). Assessment of risk factors for hypertension amongst the staff of a tertiary institution in Nigeria. *European Journal of Medical and Health Sciences, 5*(5). [Crossref]

Omoleke, S. A. (2013). Chronic non-communicable disease as a new epidemic in Africa: Focus on The Gambia. *Pan African Medical Journal, 14*, 87. [Crossref]

Salciccioli, J. D., Crutain, Y., Komorowski, M., & Marshall, D. C. (2016). Sensitivity analysis and model validation. In *Secondary analysis of electronic health records* (pp. 289–306). Springer. [Crossref]

Shikha, S., Ravi, S., & Gyan, P. S. (2017). Prevalence and associated risk factors of hypertension: A cross-sectional study in urban Varanasi. *International Journal of Hypertension.* [Crossref]

Song, L., Li, J., Yu, S., Cai, Y., He, H., Lun, J., Zheng, L., & Ye, J. (2023). Body Mass Index is associated with blood pressure and vital capacity in medical students. *Lipids in Health and Disease.* [Crossref]

Syer, R., Teoh, X. Y., Aiman, W., Aiful, & Har, C. S. (2010). The prevalence of hypertension and its associated risk factors in two rural communities in Penang, Malaysia. *JSME Publication.*

World Health Organization. (2002). *Global report on diabetes and hypertension.*

Yeo, W.-J., Abraham, R., L., A., Surapaneni, S., Schlosser, P., Ballew, S. H., Ozkan, B., Flaherty, C. M., et al. (2024). Sex differences in hypertension and its management throughout life. *Hypertension, 81*(11). [Crossref]