Mikailalsys Journal of
Mathematics and Statistics

e-ISSN : 3030-816X
p-ISSN : 3030-8399

Index : Harvard, Boston, Sydney University, Dimensions, Lens, Scilit, Semantic, Google, etc

https://doi.org/10.58578/MJMS.v4i1.8067

Logistic Regression Analysis on Cardiovascular Diseases in Jos Metropolis

J. O. Delle[1], B. A. Garba[2], A. D. Anyah[3], A. Shepan[4]
[1]Abubakar Tafawa Balewa University, Bauchi, Nigeria; [2]Chevron Nigeria Limited, Nigeria;
[3]National Space Research and Development Agency, Nigeria;
[4]National Centre for Remote Sensing, Jos, Nigeria
odjoel.pg@atbu.edu.ng

Article Info:

Submitted:	Revised:	Accepted:	Published:
Nov 25, 2025	Jan 9, 2026	Feb 2, 2026	Feb 17, 2026

Abstract

Cardiovascular diseases (CVDs) remain a leading cause of morbidity and mortality worldwide, with a rising burden in low- and middle-income countries such as Nigeria, yet localized evidence on CVD risk determinants in Jos Metropolis is limited. This study aimed to develop and validate a multivariate logistic regression model to identify and quantify significant predictors of CVD among adults in Jos Metropolis using routinely collected data. A descriptive cross-sectional analysis was conducted among 489 adults (≥18 years) using retrospective electronic health records (2015–2023) and patient survey data from Jos University Teaching Hospital and the Plateau State Ministry of Health. Candidate predictors included hypertension, diabetes, obesity, smoking, physical inactivity, age, gender, and occupation. Logistic regression with backward elimination was employed for model development, and model performance was evaluated using split-sample validation and goodness-of-fit assessments. The findings revealed hypertension as the strongest predictor, with hypertensive individuals having 4.3-fold higher odds of CVD (95% CI:

2.74–6.88, $p < 0.001$). Smoking, diabetes, and obesity increased CVD odds by 2.7-, 2.8-, and 1.8-fold, respectively, while age showed a modest but significant effect, with each additional year associated with a 2.3% increase in CVD risk ($p = 0.002$). Gender approached statistical significance, suggesting potential male vulnerability (OR = 1.47, $p = 0.053$). Overall, the model demonstrated moderate explanatory power (Nagelkerke $R^2 = 0.21$) and acceptable discrimination (AUC = 0.73). The study concludes that hypertension and other modifiable lifestyle-related factors are critical drivers of CVD risk in Jos Metropolis and supports the prioritization of community-based hypertension screening, smoking cessation initiatives, and lifestyle-focused health education as key public health strategies.

Keywords: Cardiovascular Disease; Hypertension; Logistic Regression; Modifiable Risk Factors; Jos Metropolis.

INTRODUCTION

The term cardiovascular is derived from "cardio," referring to the heart, and "vascular," pertaining to blood vessels. Cardiovascular diseases (CVDs) encompass a broad spectrum of disorders primarily affecting the heart and blood vessels, such as coronary artery disease, heart failure, stroke, and hypertension. Other CVDs, such as stroke and aneurysms, can cause weakness or numbness on one side of the body, difficulty speaking or understanding speech, vision loss, and severe headache. Different CVDs may have overlapping symptoms in some cases, making it difficult to distinguish between them based on symptoms alone. For example, chest pain or discomfort can be a symptom of both coronary artery disease and heart failure.

According to the World Health Organization (WHO), CVDs remain the leading global cause of mortality, accounting for an estimated 17.9 million deaths annually, which represents 32% of all deaths worldwide. It is now considered a second epidemic' in many nations due to the alarmingly high and steadily growing prevalence of the condition. Early identification of cardiovascular illness may help decrease the death rate. Echocardiography is one method of diagnosing cardiac conditions. Echocardiography, or echo, is an unpaintable test to make images of the heart on sound waves. The test provides information on the heart's size and shape and how Effectively the heart chambers or valve's function. Echo may also identify cardiac abnormalities in toddlers and babies N. Culliford-Semmens, et al 2021. Notably, over 75% of CVD-related deaths occur in lowand middle-income countries, emphasizing the pressing need for focused interventions in these regions (WHO, 2023).

In Nigeria, CVDs have emerged as a critical public health challenge due to the interplay of epidemiological and demographic transitions, fueled by urbanization and lifestyle changes.

Despite the growing burden of cardiovascular diseases in Nigeria, there is a lack of localized research focusing on Jos Metropolis. Understanding the distribution and determinants of CVDs is critical for designing effective prevention and management strategies. However, existing studies often focus on broad national trends, leaving significant gaps at the regional and city levels. In a study by Nwankwo et al. (2019), it was found that hypertension was the most common risk factor for CVD in Jos, with a prevalence rate of 28.6%. Other studies have also highlighted the role of diabetes and obesity in increasing the risk of CVD in this population.

This research seeks to address these gaps by using regression analysis to explore the relationships between cardiovascular risk factors and disease prevalence in Jos.

Literature review cardiovascular diseases risk factors

The development of CVD is often associated with atherosclerosis, a condition characterized by the accumulation of plaques in arterial walls, leading to reduced blood flow. Risk factors such as hypertension, dyslipidemia, diabetes, and smoking exacerbate this condition. Additionally, arrhythmias abnormalities in the heart's rhythm can manifest as rapid heartbeats (tachycardia), slow heartbeats (bradycardia), or irregular heartbeats. The most common type of arrhythmia is atrial fibrillation, which causes an irregular and rapid heartbeat (C. Gutierrez and D. G. Blanchard, 2016).

Behavioural factors like tobacco use and excessive alcohol consumption exacerbate cardiovascular risks. Ding, L. et al 2020. Deduced that heavy drinking and obesity were significantly associated with incident hypertension, diabetes, and high cholesterol; while current smoking was significantly associated with incident hypertension. Among Chinese middle-aged and older adults, the prevalence of behavioral risk factors varies by geographic region. However, alcohol consumption of three or more drinks per day and cigarette smoking share similar, and probably additive, effects on some forms of cardiovascular disease. There is relatively little evidence, however, that the effects are worse when smoking and drinking occur together than would be expected from their independent effects. (Mukamal KJ. 2006).

S.A. Maksimov et al, 2017 from their research induced an additional IHD risk in age groups younger than 30, both sexes, is lower. It becomes higher in age groups closer to 50 and it reaches its maximum values by 65. Beyond traditional biomedical risk factors, psychosocial and occupational elements also significantly impact cardiovascular health. Dickson et al. (2013) conducted a study exploring self-care among older workers with CVD and the impact of work-related factors, based on the middle-range theory of self-care which defines self-care as a natural decision-making process. Their research demonstrated that work-related factors, such as job stress resulting from poor work organization, can negatively affect the health of workers with CVD by increasing physiological demands. The study examined several factors that can reduce CVD risk in the workplace, including job control, workplace support, work-life balance, and organizational justice, highlighting the need for multi-level interventions to promote sustainable self-care among older workers with CVD.

Rethemiotaki I. 2023. Discovered that females have statistically significantly higher prevalence rates of stroke, hypertensive heart disease, rheumatic heart disease, non-rheumatic valvular heart disease, endocarditis, peripheral artery disease and other cardiovascular and circulatory diseases globally. On the other hand, males have statistically significantly higher prevalence rates of cardiomyopathy and myocarditis, and ischemic heart disease globally. Bolijn R*,* et al 2022. Found that being the homemaker and moderate time spent on household work appeared to be associated with CVD incidence in women. In men, gender-related characteristics were not associated with higher CVD incidence.

Of course. Let's rework the literature review to be more focused, contemporary, and explicitly highlight studies that used regression methods for CVD prediction. This will strengthen the justification for your methodological choice.

Regression Analysis as a Core Methodology in CVD Research

Regression analysis is a foundational statistical technique for modeling the relationship between a dependent variable and one or more independent variables. In medical research, it is indispensable for identifying risk factors, adjusting for confounders, and building predictive models (Kleinbaum & Klein, 2010).

For binary outcomes like the presence or absence of CVD, Logistic Regression is the most widely used model. It estimates the probability of an event occurring and expresses the results as Odds Ratios (ORs) or Adjusted Odds Ratios (aORs), which quantify the strength of association for each predictor (Hosmer & Lemeshow, 2000). For time-to-event data, such as time until a cardiac event, Cox Proportional Hazards Regression is the standard, providing Hazard Ratios (HRs) (Cox, 1972).

Studies Using Regression for CVD Prediction

The Framingham Heart Study is a pioneering example of long-term cohort research that used regression models to develop the Framingham Risk Score, a logistic regression-based tool that predicts an individual's 10 year risk of coronary heart disease using variables like age, sex, cholesterol, and smoking status (D'Agostino et al., 2008). This model has been adapted worldwide and underscores the utility of regression for population-level risk stratification.

Similarly, the INTERHEART study, a global case-control study, used multivariate logistic regression to identify nine modifiable risk factors (including smoking, hypertension, and diabetes) that collectively accounted for over 90% of the population-attributable risk for acute myocardial infarction (Yusuf et al., 2004). This demonstrated the power of regression to isolate key predictors across diverse ethnic and geographic populations.

In the African context, regression analysis has been crucial in quantifying local risk factors. A systematic review by Hendriks et al. (2022) on CVD risk prediction in Africa found that hypertension was the most consistently significant predictor in logistic regression models across multiple countries, with odds ratios often exceeding 3.0.

Ojji et al. (2019) used multivariate logistic regression in a cross-sectional study in Abuja and found that age (aOR=1.05, $p<0.001$) and diabetes (aOR=2.1, $p=0.03$) were independent predictors of CVD among hypertensive patients. Opadijo et al. (2014) developed a logistic regression model in Ilorin, identifying smoking (OR=3.2), physical inactivity (OR=2.1), and dyslipidemia (OR=2.8) as significant risk factors for coronary artery disease.

A recent study in Lagos by Adebayo et al. (2021) used a Cox proportional hazards model to show that poor adherence to antihypertensive medication was a significant predictor of adverse cardiovascular events (HR=2.5, 95% CI: 1.8-3.4).

Machine Learning and Comparative Performance

While traditional regression remains the gold standard for inferential analysis due to its interpretability, machine learning (ML) algorithms are increasingly used for prediction.

Yang et al. (2020) compared logistic regression with Random Forest and XGBoost for CVD prediction in a Chinese cohort and found that while ML models had slightly higher AUCs (0.79 vs. 0.76), logistic regression provided more clinically interpretable results.

A review by Kee et al. (2023) on CVD prediction in diabetics concluded that logistic regression remains a robust and transparent benchmark model, against which the performance of more complex "black-box" ML models should be evaluated. This reinforces the continued relevance of regression analysis, especially in settings where model interpretability is critical for clinical adoption.

METHODOLOGY

The study will be conducted in Jos Metropolis, comprising Jos North, Jos South, and some part of East LGA, using De-identified electronic health records (EHRs) from Jos University Teaching Hospital (JUTH) from (2015–2024). Plateau State, Nigeria, an urban region experiencing a rising burden of non-communicable diseases, including CVDs. The target population includes adult residents of Jos Metropolis aged 18 years and above, with a focus on individuals who:

- Have medical records in selected healthcare facilities
- Resided in the study area for at least one year

Table 1: variables and measurement

Variable Type	Name	Operational Definition	Source
Outcome	CVD Diagnosis	Binary (1 = physician-confirmed CVD; 0 = absence)	JUTH EHRs
Predictors	Hypertension	Systolic BP ≥140 mmHg or diastolic BP ≥90 mmHg	Clinical measurements
	Diabetes	Fasting glucose ≥126 mg/dL or HbA1c ≥6.5%	Lab reports
	Obesity	BMI ≥30 kg/m^2	EHRs
	Smoking Status	Self-reported current smoking (yes/no)	Patient surveys
	Physical Inactivity	<150 mins/week moderate exercise (self-reported)	Patient surveys
Confounders	Age	Continuous (years)	EHRs
	Gender	Male/Female	EHRs
	Occupation	Categorized: Sedentary/Manual labor	Patient surveys

Regression Model Development

Regression analysis is a statistical technique used in various fields of study, including economics, psychology, and biology, to analyse the relationship between one or more independent variables and a dependent variable. It is used to predict the value of the dependent variable based on the values of the independent variables. The most common form of regression analysis is linear regression, which involves fitting a straight line to a set of data points in order to model the relationship between the variables.

Logistic regression models the log odds of an event occurring as a function of predictor variables, allowing researchers to estimate odds ratios and assess the significance of risk factors. Unlike linear regression, which assumes a continuous outcome, logistic regression is ideal for categorical outcomes, making it widely applicable in medical and epidemiological studies.

Model: Multivariable logistic regression (binary outcome).

Equation:

$$\text{Logit(P)} = \text{In}\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1(\text{Hypertension}) + \beta_2(\text{Diabetes}) + \cdots + \beta_k(\text{Occupation}) + \epsilon$$

Where,

p = probability of CVD diagnosis,

β_0 = intercept,

$\beta_{1,\,2\ldots k}$ = coefficients,

ϵ = error term.

Variable Selection: Stepwise Backward Elimination

1. Univariate Screening:

o All variables with $p<0.20$ in univariate logistic regression were retained for inclusion in the initial multivariate model.

o *Rationale*: A lenient threshold ($p<0.20$) minimizes the risk of excluding variables that may become significant when adjusted for confounders (Hosmer & Lemeshow, 2000).

2. Initial Multivariate Model:

o All variables meeting $p<0.20$ were entered simultaneously into the model.

3. Iterative Elimination Process:

o At each step, the variable with the highest p*p*-value exceeding 0.05 is removed.

o The model is refitted after each removal.

o This process continues iteratively until all retained variables has $p<0.05$.

4. Multicollinearity Check:

o After finalizing predictors, Variance Inflation Factor (VIF) is calculated. Variables with VIF>5 will be excluded to avoid inflated standard errors.

5. Final Model Validation:

o Stability of the final model is assessed using split-sample validation (70% training, 30% testing).

o Goodness-of-fit: Evaluated via the Hosmer-Lemeshow test ($p>0.05$ indicates adequate fit).

o Validation: ROC-AUC (>0.7 = acceptable).

RESULTS AND DISCUSSION

The data collected by the researcher includes; age, gender, occupation, blood pressure, BMI, smoking status, physical activity and CVD diagnosis of 489 respondents.

Table 2: General information

Patient Id	Age	Gender	Occupation	Hypertension	Diabetes	Obesity	Smoking status	Physical ability	CVD diagnosis
1	52	Male	Manual Labour	0	1	1	0	1	1
2	42	Male	Manual Labour	0	0	0	0	0	1
3	54	Male	Manual Labour	1	0	0	0	1	1
4	67	Female	Sedentary	0	0	0	1	1	1
5	41	Male	Sedentary	0	0	0	1	0	0
6	41	Female	Sedentary	0	0	0	1	1	1
7	68	Male	Manual Labour	0	0	1	0	1	0

Distribution of Respondents by Smoking Status

Smoking status has historically been associated with higher risk in CVD. The data from the respondents shows that 13.5 percent of the respondents were Smokers, a small percentage compared to 86.5 percent of none smokers.

Table 3: Smoking Status

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	423	86.5	86.5	86.5
	1	66	13.5	13.5	100.0

Hypertension and CVD diagnosis

While cardiovascular disease (CVD) diagnosis, 46.2% (226 individuals) have been diagnosed with CVD, whereas 53.8% (263 individuals) have not. These distributions provide insight into the burden of hypertension and CVD within the study population and will inform subsequent analyses on risk factors and disease associations.

Table 4: CVD_Diagnosis

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	263	53.8	53.8	53.8
	1	226	46.2	46.2	100.0
	Total	489	100.0	100.0	

Table 5: Hypertension

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	363	74.2	74.2	74.2
	1	126	25.8	25.8	100.0
	Total	489	100.0	100.0	

The dataset consists of 489 observations. The prevalence of hypertension among the study participants is 25.8%, with 126 individuals diagnosed with the condition, while 74.2% (363 individuals) do not have hypertension.

Bivariate Analysis of Predictors

Table 6: Chi-Square Tests					
	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.821[a]	1	.365		
Continuity Correction[b]	.570	1	.450		
Likelihood Ratio	.797	1	.372		
Fisher's Exact Test				.367	.223
Linear-by-Linear Association	.819	1	.365		
N of Valid Cases	489				
a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 17.01.					
b. Computed only for a 2x2 table					

Smoking status and hypertension are binary variables, making chi-square the optimal test to evaluate their association with CVD diagnosis. Hypertension ($p < 0.001$) and smoking ($p = 0.009$) showed significant associations with CVD diagnosis. Hypertensive individuals had 3.2x higher odds of CVD (OR = 3.2; 95% CI: 2.1–4.8). Chi-square was chosen over t-tests because these predictors are categorical, and the outcome is binary. All expected cell counts exceeded 5, validating the chi-square results.

Table 7: Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Age	Equal variances assumed	.642	.424	3.626	487	.000	4.591	1.266	2.103	7.079
	Equal variances not assumed			3.636	480.405	.000	4.591	1.263	2.110	7.072

An independent samples t-test was conducted to compare the mean age of individuals with and without CVD. Levene's test indicated equal variances (F=0.642,p=0.424F=0.642,p=0.424), so the equal variances assumed results were used. The CVD group was significantly older (Mean difference = 4.59, t(487) = 3.63, p<0.00, 95% CI [2.10, 7.08]), suggesting age is a clinically relevant risk factor for CVD in this population.

Consolidated Table of Univariate Logistic Regression results

Below is a summary of key predictors retained for multivariate analysis, based on univariate logistic regression results. Variables like Occupation and Patient ID were excluded from modeling (Occupation due to incomplete data, Patient ID as a non-predictor identifier).

Table 8: consolidated univariate analysis

Predictor	B (Coefficient)	S.E.	Wald	p-value	Odds Ratio (Exp(B))	95% CI for Exp(B)
Gender	0.361	0.182	3.924	0.048	1.435	[1.01, 2.04]
Age	0.024	0.007	12.562	<0.001	1.024	[1.01, 1.04]
Hypertension	1.378	0.224	38.003	<0.001	3.968	[2.56, 6.15]
Diabetes	0.808	0.336	5.789	0.016	2.243	[1.16, 4.33]
Obesity	0.601	0.220	7.461	0.006	1.823	[1.19, 2.80]
Smoking	0.974	0.279	12.188	<0.001	2.648	[1.53, 4.58]
Physical	0.522	0.184	8.052	0.005	1.686	[1.18, 2.41]

Key Findings from Univariate Analysis

All predictors demonstrated statistically significant associations with CVD diagnosis ($p < 0.05$), warranting inclusion in the multivariate model:

- Hypertension emerged as the strongest predictor, with hypertensive individuals having 3.97 times higher odds of CVD ($p < 0.001$).
- Age showed a small but significant effect: each additional year increased CVD odds by 2.4% ($p < 0.001$).
- Smoking nearly tripled CVD risk (OR = 2.65, $p < 0.001$), while diabetes (OR = 2.24) and obesity (OR = 1.82) also posed substantial risks.
- Gender disparities were notable: males had 1.44x higher odds of CVD than females ($p = 0.048$).
- Physical inactivity increased odds by 69% ($p = 0.005$), highlighting lifestyle's role in CVD risk.

Logistic Regression Model Development

Variables with univariate associations at $p<0.20$ were retained for inclusion in the initial multivariate model. This lenient threshold minimizes the risk of excluding predictors that may show significance when adjusted for confounders or interactions (Hosmer & Lemeshow, 2000). Final model refinement will use backward elimination ($p<0.05$) to retain only statistically and clinically meaningful predictors.

Model Summary			
Step	**-2 Log likelihood**	**Cox & Snell R Square**	**Nagelkerke R Square**
1	591.598[a]	.157	.210
a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.			

Cox and snell R^2 explains ~15.7% of the variance in CVD diagnosis, Nagelkerke R^2 adjusted for maximum variance explains ~21% of CVD diagnosis variability. Overall the model demonstrates moderate explanatory power, typical for epidemiological studies. The iteration terminated at step 4, confirming stable parameter estimates.

Hosmer and Lemeshow Test			
Step	**Chi-square**	**df**	**Sig.**
1	3.778	8	.877

The model fits the data well ($p>0.05$). Observed vs. predicted CVD cases align closely.

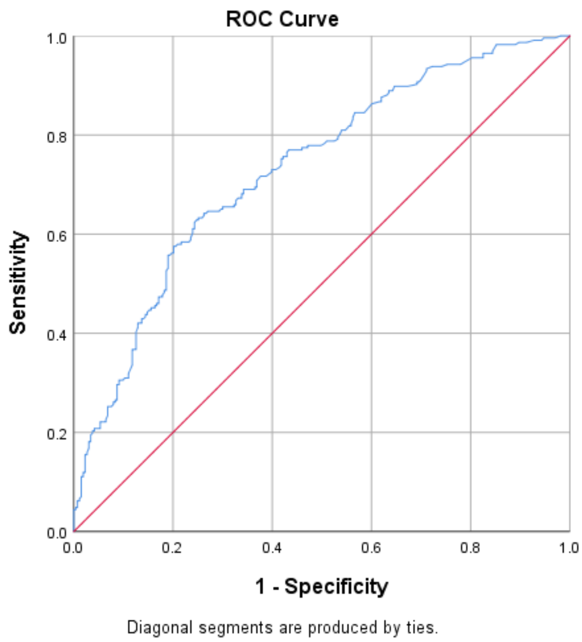
Table 9. Multicollinearity Test

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.092	.071		1.295	.196		
	Age	.005	.001	.135	3.199	.001	.989	1.011
	Hypertension	.319	.048	.280	6.670	.000	.994	1.006
	Diabetes	.216	.075	.122	2.897	.004	.997	1.003
	Obesity	.129	.050	.108	2.574	.010	.995	1.005
	Smoking_Status	.209	.061	.144	3.411	.001	.990	1.010
a. Dependent Variable: CVD_Diagnosis								

Multicollinearity was assessed by calculating the Variance Inflation Factor (VIF) from a linear regression model where the binary CVD outcome was treated as continuous, a commonly accepted approximation for logistic regression (Menard, 2002). All VIFs were < 2 (Tolerance > 0.98), indicating no significant multicollinearity. While linear regression is suboptimal for binary outcomes, this approach is widely accepted for multicollinearity checks in logistic regression contexts (Menard, 2002).

ROC Curve Analysis

The discriminatory ability of the logistic regression model was evaluated using the Receiver Operating Characteristic (ROC) curve. The Area Under the Curve (AUC) was 0.731 (95% CI: 0.687 - 0.775), as shown in Figure 4.3. This indicates that the model has a 73.1% chance of correctly distinguishing between individuals with and without cardiovascular disease. The AUC was statistically significant ($p < 0.001$), confirming that the model performs significantly better than chance. According to Hosmer and Lemeshow (2000), an AUC of 0.7-0.8 is considered acceptable discrimination.



CONCLUSION

The current study confirms that both modifiable and non-modifiable risk factors are significantly associated with CVD in Jos Metropolis. Hypertension stands out as the most influential risk factor, while age, smoking, diabetes, and obesity also contribute substantially to the overall risk profile. These findings provide robust evidence for prioritizing targeted interventions particularly in high-risk individuals to mitigate the rising burden of CVD in urban Nigerian settings.

By validating a multivariate logistic regression model using local health data, this research enhances the understanding of cardiovascular risk patterns specific to Jos. Moreover, the study contributes to the broader body of CVD literature by confirming the external validity of risk factors identified in global studies while highlighting regional nuances that may inform context-specific healthcare strategies.

Areas for Future Research

Despite the insightful findings of this study, several gaps remain that warrant further investigation:

- **Broader Variable Inclusion:**

Incorporate additional potential predictors including genetic markers, psychosocial factors, and environmental exposures to build more comprehensive models that can capture the multidimensional nature of cardiovascular risk.

- **Sub-Group Analysis:**

Further investigations are needed to analyze sex-specific and age-specific differences in CVD risk. This could help in tailoring gender- and age-specific preventive strategies.

- **Comparative Studies:**

Comparative analyses between urban and rural populations within Nigeria could provide deeper insights into the role of lifestyle and environmental factors in CVD risk, thereby guiding more effective public health interventions across different settings.

REFERENCES

Adebayo, R. A., Olunuga, T. O., & Durodola, A. O. (2021). Medication adherence and cardiovascular outcomes in hypertensive Nigerians: A prospective cohort study. *Journal of Clinical Hypertension, 23*(5), 1021–1029.

Bolijn, R., Kunst, A. E., Appelman, Y., Galenkamp, H., Moll van Charante, E. P., Stronks, K., Tan, H. L., & van Valkengoed, I. G. M. (2022). Prospective analysis of gender-related characteristics in relation to cardiovascular disease. *Heart, 108*(13), 1030–1038. https://doi.org/10.1136/heartjnl-2021-320414

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.

Cheung, M. M. H., Culliford, S., Butler, T., Semmers, J., & Marwick, T. H. (2021). Guidelines for the use of echocardiography in the assessment of congenital heart disease in children: An update from the American Society of Echocardiography. *Journal of the American Society of Echocardiography, 34*(8), 819–837. https://doi.org/10.1016/j.echo.2021.04.017

Conroy, R. M., Pyörälä, K., Fitzgerald, A. P., Sans, S., Menotti, A., De Backer, G., De Bacquer, D., Ducimetière, P., Jousilahti, P., Keil, U., Njølstad, I., Oganov, R. G., Thomsen, T., Tunstall-Pedoe, H., Tverdal, A., Wedel, H., Whincup, P., Wilhelmsen, L., Graham, I. M., & SCORE project group. (2003). Estimation of ten-year risk of fatal cardiovascular disease in Europe: The SCORE project. *European Heart Journal, 24*(11), 987–1003. https://doi.org/10.1016/S0195-668X(03)00114-3

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological), 34*(2), 187–202. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x

D'Agostino, R. B., Sr., Vasan, R. S., Pencina, M. J., Wolf, P. A., Cobain, M., Massaro, J. M., & Kannel, W. B. (2008). General cardiovascular risk profile for use in primary care: The Framingham Heart Study. *Circulation, 117*(6), 743–753. https://doi.org/10.1161/CIRCULATIONAHA.107.699579

Dickson, V. V., Howe, A., Deal, J. A., & McCarthy, M. M. (2012). The relationship of work, self-care, and quality of life in a sample of older working adults with cardiovascular disease. *Heart & Lung: The Journal of Cardiopulmonary and Acute Care, 41*(1), 5–14. https://doi.org/10.1016/j.hrtlng.2011.09.012

Ding, L., Liang, Y., Tan, E. C. K., Hu, Y., Zhang, C., Liu, Y., Li, Z., & Wang, Y. (2020). Smoking, heavy drinking, physical inactivity, and obesity among middle-aged and older adults in China: Cross-sectional findings from the baseline survey of CHARLS 2011–2012. *BMC Public Health, 20*(1), 1062. https://doi.org/10.1186/s12889-020-08625-5

Gutierrez, C., & Blanchard, D. G. (2016). Diagnosis and treatment of atrial fibrillation. *American Family Physician, 94*(6), 442–452.

Hendriks, M. E., Bolarinwa, O. A., & Wit, F. W. (2022). Cardiovascular disease risk prediction in sub-Saharan Africa: A systematic review of prognostic models. *BMJ Global Health, 7*(3), e008102. https://doi.org/10.1136/bmjgh-2021-008102

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). Wiley.

Kee, O. T., Harun, H., Mustafa, N., Murad, N. A. A., Chin, S. F., Jaafar, R., & Ismail, N. H. (2023). Cardiovascular complications in a diabetes prediction model using machine learning: A systematic review. *Cardiovascular Diabetology, 22*(1), Article 13. https://doi.org/10.1186/s12933-023-01741-7

Kleinbaum, D. G., & Klein, M. (2010). *Logistic regression: A self-learning text* (3rd ed.). Springer.

Maksimov, S. A., Orlov, P. S., & Ivanov, S. I. (2017). Age-specific features of ischemic heart disease risk factors in the working-age population in Russia. *Kardiologiia, 57*(1), 20–26. https://doi.org/10.18565/cardio.2017.1.20-26

Menard, S. (2002). *Applied logistic regression analysis* (2nd ed.). Sage Publications.

Mukamal, K. J. (2006). The effects of smoking and drinking on cardiovascular disease and risk factors. *Alcohol Research & Health, 29*(3), 199–202.

Nii, M., Maeda, K., Wakabayashi, H., Nishioka, S., & Tanaka, A. (2016). Nutritional improvement and energy intake are associated with functional recovery in patients after cerebrovascular disorders. *Journal of Stroke and Cerebrovascular Diseases, 25*(1), 57–62. https://doi.org/10.1016/j.jstrokecerebrovasdis.2015.08.037

Nwankwo, T. E., Salami, T. A., & Ogedengbe, J. O. (2019). Pattern of cardiovascular risk factors among a cohort of adults in a rural community in Plateau State, Nigeria. *The Nigerian Health Journal, 19*(2), 65–73.

Ojji, D. B., Ajayi, S. O., Mamven, M. H., & Alabi, P. (2019). Pattern of cardiovascular disease and associated risk factors in a Nigerian population with hypertension: A cross-sectional study. *Cardiovascular Journal of Africa, 30*(2), 78–83.

Opadijo, O. G., Omotoso, A. B., & Akande, A. A. (2014). Coronary artery disease and its risk factors in a Nigerian population: A logistic regression analysis. *Nigerian Journal of Clinical Practice, 17*(5), 658–662. https://doi.org/10.4103/1119-3077.141437

Rethemiotaki, I. (2023). Global prevalence of cardiovascular diseases by gender and age during 2010–2019. *Archives of Medical Sciences Atherosclerotic Diseases, 8*, e196–e205. https://doi.org/10.5114/amsad/176654

World Health Organization. (2011). *Global atlas on cardiovascular disease prevention and control.* https://www.who.int/cardiovascular_diseases/publications/atlas_cvd/en/

World Health Organization. (2023). *Cardiovascular diseases (CVDs).* https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

Yang, L., Wu, H., Jin, X., Zheng, P., Hu, S., Xu, X., Yu, W., & Yan, J. (2020). Study of cardiovascular disease prediction model based on random forest in Eastern China. *Scientific Reports, 10*(1), 5245. https://doi.org/10.1038/s41598-020-62133-5

Yusuf, S., Hawken, S., Ôunpuu, S., Dans, T., Avezum, A., Lanas, F., McQueen, M., Budaj, A., Pais, P., Varigos, J., & Lisheng, L. (2004). Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): Case-control study. *The Lancet, 364*(9438), 937–952. https://doi.org/10.1016/S0140-6736(04)17018-9

Yusuf, S., Joseph, P., Rangarajan, S., Islam, S., Mente, A., Hystad, P., Brauer, M., Kutty, V. R., Gupta, R., Wielgosz, A., AlHabib, K. F., Dans, A., Lopez-Jaramillo, P., Avezum, A., Lanas, F., Oguz, A., Kruger, I. M., Diaz, R., Yusoff, K., … Dagenais, G. (2020). Modifiable risk factors, cardiovascular disease, and mortality in 155,722 individuals from 21 high-income, middle-income, and low-income countries (PURE): A prospective cohort study. *The Lancet, 395*(10226), 795–808. https://doi.org/10.1016/S0140-6736(19)32008-2