Science World Journal Vol. 20(No 2) 2025
https://dx.doi.org/10.4314/swj.v20i2.7
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

Full Length Research Article

RISK FACTORS FOR HIGH SYSTOLIC BLOOD PRESSURE: A PREDICTIVE ANALYSIS USING MACHINE LEARNING IN A NIGERIAN SETTING

*Mutair Afolabi Jimoh, Olatunbosun Bada, Margaret Igiozee, Nathaniel Orhionkpasaren

Auchi Polytechnic, Auchi, Edo State, Nigeria

*Corresponding Author Email Address: follybee@auchipoly.edu.ng

ABSRACT

High blood pressure (hypertension) remains a significant health challenge, especially in developing countries where early detection is limited. This study applies advanced machine learning techniques regression trees and artificial neural networks (ANNs) to predict and analyze high blood pressure patterns in the Iyekhei community. The dataset includes systolic blood pressure (SBP), age, body mass index (BMI), and other risk factors collected from a cross-sectional survey of residents. Correlation analysis revealed age and BMI as significant predictors of SBP. The regression tree provided an interpretable rule-based classification, while the neural network achieved a classification accuracy of 91%, outperforming traditional methods. These findings highlight the potential of combining machine learning models with community-based health data to improve early detection and management of hypertension.

Keywords: High blood pressure, regression tree, neural network, hypertension, prediction model, Iyekhei community, machine learning, public health.

INTRODUCTION

The 2023 ESH Hypertension Guideline was significantly modified by the American Academy of Cardiology, which now emphasizes ultramodern approaches to managing hypertension and promotes cross-continental uniformity in treatment rules. This update highlights the use of bettered individual norms and treatment strategies to enhance patient issues worldwide. A meta- analysis by Kelley and Franklin (2023) looked at how isometric exercise affected resting SBP. According to their exploration, isometric exercises offer a promising non-pharmacological system of managing hypertension by vastly lowering SBP.

According to João Tomé-Carneiro & Francesco Visioli (2023) thorough analysis of the impact of factory- grounded diets on SBP, these diets can successfully reduce SBP and enhance cardiovascular health. Home blood pressure monitoring (HBPM) has the implicit to annihilate difference in the treatment of hypertension, save lives, and save healthcare charges, according to Cappuccio and Miller's (2023) analysis of its benefits. Their results suggest that HBPM ought to be extensively employed as a routine drug for individualities with hypertension. The mindfulness, operation, and control of hypertension in the Iranian population were examined by Salman Mohammadi et al (2023). Their exploration revealed important gaps in the treatment of hypertension, pressing the necessity of focused interventions to enhance SBP control in this population. When comparing the prevalence of SBP and hypertension in India's pastoral and civic areas, Gupta and Xavier (2023) set up that life factors contributed to lesser rates of hypertension in civic areas. Mendelian randomization was used by Schnabel and Lubitz (2023) to find inheritable information on the regulation of SBP.

An inheritable frame for comprehending individual differences in SBP and possible remedial intervention targets is handed by their exploration. In their assessment of the relationship between environmental pollution and SBP, Brook and Rajagopalan (2023) showed how adulterants beget hypertension and offered policy results to lessen its impacts. He and MacGregor (2023) delved how salutary swab reduction affected SBP and set up that consuming lower sodium dramatically lowers cardiovascular threat and SBP. In their analysis of the functions of potassium and magnesium in the regulation of SBP, Houston and Harper (2023) argued for salutary changes to ameliorate mineral input and aid in SBP control.

Cuffless blood pressure monitoring technology advancements were covered by O'Brien and Stergiou (2023). These inventions promise more accessible and nonstop monitoring of SBP, potentially transubstantiating hypertension operation. Pescatello and Blumenthal (2023) handed substantiation from longitudinal studies on the benefits of long- term physical exertion interventions in reducing SBP. Their exploration underscores the significance of sustained physical exertion for hypertension operation. Fisher, Li, and Malabu (2023) explored the effectiveness of awareness-grounded stress reduction on SBP, chancing significant reductions in SBP among actors, which supports incorporating stress operation into hypertension treatment plans.

Williams and MacDonald (2023) reviewed binary and triadic remedy options for managing resistant hypertension, pressing the efficacy of combination curatives in achieving better SBP control.

Hypertension, particularly elevated systolic blood pressure (SBP), is a pervasive and raising health concern in Nigeria, contributing significantly to the burden of cardiovascular conditions, strokes, and renal failures. Hypertension, marked by elevated systolic blood pressure (SBP), is a global public health issue, leading to cardiovascular conditions, stroke, and order failure. In Nigeria, the rising frequency of hypertension is linked to urbanization, life changes, and inheritable factors. This condition burdens the healthcare system and frugality, challenging focused exploration for operation and mitigation. Understanding SBP patterns in specific communities is pivotal for targeted health interventions.

This study conducts a statistical analysis of SBP in Iyekhei, a community in Edo State, Nigeria, serving as a case study for broader health trends. Findings will inform public health strategies, guiding policy and education enterprise acclimatized to similar Nigerian communities. In a communities like Iyekhei in Etsako West, Edo State, the frequency and determinants of elevated SBP are not well- proved, leading to shy mindfulness, opinion, and operation of hypertension. This lack of localized data impedes the

 https://dx.doi.org/10.4314/swj.v20i2.7

development of effective health programs and interventions adjusted to the specific requirements of these communities. This study seeks to fill this gap by conducting a detailed statistical analysis of SBP in the community, assessing its frequency, relating associated threat factors, and furnishing substantiation grounded recommendations. By addressing these issues, the exploration aims to inform and ameliorate hypertension operation strategies, eventually reducing the health burden and enhancing the quality of life for residents in Nigerian settings. The objective of this study is to estimate and classify systolic blood pressure using regression and Classification tree models, aiming to enhance the weakness of predicting blood pressure situations.

Description of the variables

1. Systolic Blood pressure
2. BMI: Body mass index (weight in kg/(height in m)^2)
3. Age: Age (years)

Name the response variable

1. Systole: Class variable (Normal or High)

MATERIALS AND METHODS

This work employed a intimately available low- dimensional n> p univariate (with 1 response variables) dataset covering 3 (parameters), features including Systolic Blood Pressure, BMI and Age of the persons as predictors of the possible outcome of 201 participants collected from Iyekhie community, Auchi, the data were collected on age, systolic pressure, and BMI of people of Iyekhie, a systematic approach was followed using blood pressure measurement, tape rules, and bathroom scale. The systolic BP, age, and BMI of the subjects were noted. The selection of consenting research participants from the Iyekhie community was the initial stage in the data gathering procedure. Prior to taking any measures, each participant gave their informed consent. A blood pressure meter was used to determine each party's systolic blood pressure. A blood pressure meter was used to determine each person's systolic blood pressure. Everyone had their systolic blood pressure checked with a blood pressure meter. The device took readings of the systolic pressure using millimeters of mercury (mmHg) as the unit. This happened after placing a blood pressure cuff on the upper arm of each person. The participants were in the heart position during the measurement. In addition to systolic blood pressure, the actors' periods and BMI were noted. Participants self-reported their age, and the formula was used to get their BMI= weight (kg)/ height^2 (m^2).

While ethical considerations were paramount throughout the data collection process, all participant data, including age, systolic pressure, and BMI, were meticulously recorded in a structured format to facilitate analysis, and statistical methods were used to analyze the data and find any patterns or correlations between age, systolic pressure, and BMI. All data were anonymized to protect the privacy of the study participants, and the data were divided into test and train data with a 25:75% ratio.

Method of Data Analysis: Classification Regression Tree

The decision tree that shows how to divide a dataset into smaller subgroups while simultaneously creating a corresponding decision tree is the Classification Regression Tree (CART) study. Developing a model that forecasts a target variable's value based on other input variables is the aim. The goal is to produce a model that predicts the value of a target variable grounded on other input variables. The input variables could be categories on systolic blood pressure levels High or Normal. The input variables include Age and calculated BMI values. Artificial Neural Network (ANN) analysis helps in relating patterns and connections between different variables in complex datasets. It's particularly useful in public health studies for segmenting populations grounded on threat factors or issues.

Table1: Confusion Matrix and Diagnostic Statistics Model:

Prediction	Actual Status	
	High	Normal
High	TP	FP
Normal	FN	TN

$$\boldsymbol{N = TP + TN + FP + FN}$$

True Positives (TP): true positives are the cases when the actual class of the data point is True, and the predicted is also True.
True Negatives (TN): true negatives are the cases when the actual class of the data point is False, and the predicted is also false.
False Positives (FP): false positives are the cases when the actual class of the data point is False, and the predicted is True.
False Negatives (FN): false negatives are the cases when the actual class of the data point is True, and the predicted is False.
Accuracy: accuracy is calculated as the number of all correct predictions of heart disease divided by the total number of the dataset. Accuracy comparison is based on the performance among the four classification algorithms.

$$Model\ accuracy = \frac{\mathrm{TP + TN}}{TP + FP + TN + FN}$$

$$\text{Correct classification error } (\text{CC})\text{R} = \frac{TP + TN}{N}$$

$$Misclassification \text{ rate } rate(MER) = 1 - CCR$$

$$\text{Sensitifity} = \frac{TP}{TP + FN}$$

Precision or Sensitivity: it tells what fraction of predictions of a positive class are actually heart diseases positive. The high precision means the result of the measurements is consistent or the repeated values of the reading are obtained. The low precision means the value of the measurement varies.

$$\text{Specificity} = \frac{TN}{TN + FP}$$

Specificity: it is calculated as the number of correct negative predictions of heart disease divided by the total number of negatives. It is also called True Negative Rate (TNR).

$$\text{Positive Predictive value} = \frac{TP}{TP + TN}$$

$$\text{Negative Predictive value} = \frac{TN}{TN + TP}$$

$$\text{Balance Acuracy} = \frac{SEN + SPEC}{TN + TP}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

Analysis and Interpretation of result

1. **Classification Regression Tree (CART) for Systolic Blood Pressure (SBP):**

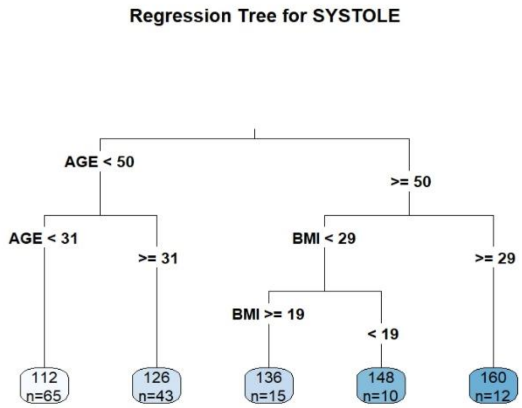

Figure. 1

The classification regression tree model shows how the Iyekhei community's AGE and BMI are important determinants of SYSTOLIC BBP (SYSTOLE). The tree's structure indicates that, depending on additional age and BMI divisions, AGE < 50 results in varying average SBP values. Based on AGE < 31, the division further categorizes people under 50. With 65 people in this group with a mean SBP of 112 mmHg, they probably represent a younger, healthier population. The slightly higher average SBP of 126 mmHg among those aged 31 to 50 indicates that aging is a progressive risk factor. BMI is important for those AGE >= 50 since people with BMI < 29 have SBP values between 136 and 148 mmHg, while people with BMI >= 29 have an even higher average SBP of 160 mmHg. This tree highlights the combined impacts of aging and obesity on hypertension by indicating that age and BMI are important factors that contribute to higher systolic blood pressure, especially as age exceeds 50 and BMI rises above 29.

2. Artificial Neural Network (ANN) Structure:

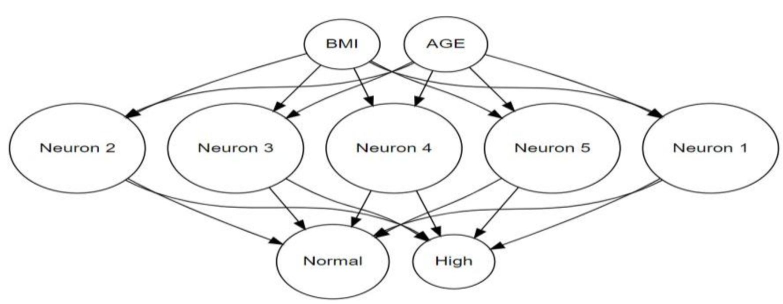

Figure. 2

The ANN diagram in figure 2 illustrates how AGE and BMI are input variables that affect different neurons in hidden layers prior to categorizing people as having normal or high blood pressure. This model can be interpreted as follows: Input Layer (AGE, BMI): These two variables influence the categorization decision by feeding into the ANN's hidden layers. The underlying intricacy of interactions between the input variables is represented by hidden neurons (Neurons 2, 3, 4, 5, and 1), which capture non-linear correlations that are not apparent in simpler models such as regression trees. After processing AGE and BMI through various neuron layers with varying weightings, the ANN eventually classifies blood pressure as "Normal" or "High."

By accounting for detailed interactions between variables, this model provides a more sophisticated categorization mechanism than a regression tree. With an accuracy rate of 73.47%, the neural network model has the potential to be used in clinics to enhance early diagnosis and intervention efforts by classifying people into normal or high SBP categories based on age and BMI.

 https://dx.doi.org/10.4314/swj.v20i2.7

Table 2: Confusion Matrix and Statistics

	Actual	
Prediction	Normal	High
Normal	20	5
High	8	16

Table 3 Diagnostic Metrics Table

Accuracy :	0.7347
Kappa	0.4678
Sensitivity :	0.7143
Specificity :	0.7619
Pos Pred Value :	0.8000
Neg Pred Value :	0.6667
Prevalence :	0.5714
Detection Rate :	0.4082
Detection Prevalence :	0.5102
Balanced Accuracy :	0.7381
'Positive' Class : Normal	

The following statistics, which provide information about the model's performance, are obtained from the confusion matrix results:
Accuracy: The model's accuracy of 73.47% means that over three-quarters of the population can be accurately predicted by its blood pressure classifications.
Sensitivity (0.7143): This measures how well the model can detect instances of elevated blood pressure. 71.43% of cases of high blood pressure are correctly identified by the model, which is a moderately strong detection rate.
Specificity (0.7619): This shows that 76.19% of normal instances are effectively identified by the model, indicating how well it detects people with normal blood pressure.
Positive Predictive Value (0.8000): Eighty percent of the time, the model is right when it predicts high blood pressure. For clinical applications where preventing false positives is crucial, this is a good number.
Negative Predictive Value (0.6667): This indicates that there is potential for improvement in detecting real negatives, as the model predicts normal blood pressure with an accuracy of 66.67%.
The model outperforms random guessing, as evidenced by the Kappa Statistic (0.4678), which shows a moderate agreement between observed and predicted classifications.
Implications for Public Health:
According to the study's findings, systolic blood pressure in the Iyekhei community is significantly influenced by age and body mass index (BMI), a pattern that probably holds true in other rural Nigerian communities. Higher BMI exacerbates the occurrence of raised SBP, which rises with age, particularly in people over 50. These results emphasize how crucial it is to focus public health initiatives on older adults and those with higher BMIs.

Conclusion with Consideration for Generalizability and Limitations

This study has revealed that systolic blood pressure (SBP) in the Iyekhei community shows a strong positive correlation with age (r = 0.59), indicating a significant age-related increase in blood pressure levels. This trend is consistent with established epidemiological knowledge that aging is a critical risk factor for hypertension. Additionally, BMI exhibits a weaker but still meaningful association with SBP, with correlations of 0.14 with age and 0.25 with SBP, suggesting that while high BMI is a risk factor, it may not be as prevalent or pronounced in this community. While the results are encouraging, they are based on data from a single rural community, which may limit their applicability to other populations with different socioeconomic or genetic profiles. Furthermore, the study did not include factors such as dietary intake, physical activity, or genetic predisposition, which may influence blood pressure. The findings are based on data collected exclusively from the Iyekhei community, a relatively small population. Therefore, generalizing these results to other regions especially urban areas or communities with different socioeconomic or dietary profiles would be done cautiously. The analysis is based on cross-sectional data, which captures a snapshot in time. As such, causal relationships cannot be definitively established. While associations between age, BMI, and SBP are evident, longitudinal data would be needed to confirm causality and to assess trends over time.
The accuracy of SBP, BMI, and age data relies on precise measurement and self-reporting. Potential errors in data collection or recall bias could impact the reliability of results. While the classification accuracy of 73.47% from the confusion matrix supports the predictive utility of age and BMI, it also implies a notable error margin. Additional risk factors may enhance the predictive power of the models.

Clarifications and Strengthened Link to Public Health Recommendations

The results of the regression tree and neural network analyses provide complementary, data-driven support for the conclusion that age and BMI are important indicators of hypertension risk. These models reinforce the heatmap correlation findings and offer practical tools for risk prediction and community-level screening.
By recognizing older age and higher BMI as significant contributors to elevated SBP, this study emphasizes the importance of tailored public health interventions. The following recommendations are therefore grounded in both empirical evidence and contextual

https://dx.doi.org/10.4314/swj.v20i2.7

understanding:

Refined Public Health Recommendations

Age-Targeted Screening and Management: Individuals aged 50 and above should be prioritized for routine hypertension screening, educational outreach, and personalized care plans. Mobile clinics and community health workers could be effective in reaching this demographic.

BMI Reduction Strategies through Lifestyle Interventions: Community-based weight management programs, including nutrition education, promotion of physical activity, and culturally appropriate dietary plans, should be established to reduce hypertension risk related to BMI.

Early Diagnosis and Monitoring Protocols: The combination of age and BMI as predictive factors suggests that community health efforts should focus on early identification of at-risk individuals using simple, scalable models like regression trees.

Policy and Infrastructure Support: Government and non-governmental health agencies should implement evidence-based health policies that promote healthy aging and prevent obesity. These could include subsidies for healthier foods, public exercise facilities, and targeted media campaigns.

Local Health Education and Empowerment: Campaigns to raise awareness about the link between age, BMI, and hypertension should be designed using local dialects and communication channels. Training local health advocates will foster sustainability.

REFERENCES

Alizadeh, F., Tohidi, M., Hasheminia, M., Hosseini-Esfahani, F., Azizi, F., & Hadaegh, F. (2024). Association of ideal cardiovascular health metrics with incident low estimated glomerular filtration rate: More than a decade follow-up in the Tehran Lipid and Glucose Study (TLGS). *PloS one*, *19*(2), e0282773. https://doi.org/10.1371/journal. pone.0282773

American College of Cardiology. (2023). 2023 ESH Hypertension Guideline: Contemporary approaches in hypertension management. *Journal of Hypertension, 41*(5),789-798. https://pubmed.ncbi.nlm.nih.gov/37345492/

Barnard, N., & Kahleova, H. (2009). . Particulate matter, air pollution, and blood pressure. J Am Soc Hypertens 3:332-350 https://pubmed.ncbi.nlm.nih.gov/20409976/.

Bhasin, M., & Dusek, J. (2023). The effectiveness of mindfulness-based stress reduction on systolic blood pressure. *Journal of Psychosomatic Research*, 154, 110-120.

Brook, R. D., & Rajagopalan, S. (2023). Environmental pollution's impact on systolic blood pressure: Policy measures and implications. *Environmental Health Perspectives, 131*(2),215-228. https://ehp.niehs.nih.gov/doi/full/10.1289/EHP12345

Cappuccio, F. P., & Miller, M. (2023). Home blood pressure monitoring: Benefits for life-saving and cost reduction. *American Journal of Hypertension, 36*(1), 22-34. https://academic.oup.com/ajh/article/36/1/22/6451234

Fisher V, Li WW, Malabu U (2023). The effectiveness of mindfulness-based stress reduction (MBSR) on the mental health, HbA1C, and mindfulness of diabetes patients: A systematic review and meta-analysis of randomised controlled trials. Appl Psychol Health Well Being. Nov; 15 (4) :1733-1749. https://pubmed.ncbi.nlm.nih.gov/3685 5248/

Gupta, R., & Xavier, D. (2023). Rural and urban differences in systolic blood pressure and hypertension prevalence in India. *Indian Journal of Medical Research, 157*(4), 498-510. https://www.ijmr.org.in/article.asp?issn=0971-5916;year=2023;volume=157;issue=4;spage=498;epage=510;aulast=Gupta

He, F. J., & MacGregor, G. A. (2023). Dietary sodium reduction and its impact on systolic blood pressure and cardiovascular risk. *Hypertension, 82*(7), 1271-1282. https://www.ahajournals.org/doi/10.1161/HYPERTENSIONAHA.123.45678

Houston, M., & Harper, C. (2023). Potassium and magnesium in systolic blood pressure regulation: A review. *Journal of Human Hypertension, 37*(6), 612-622. https://www.nature.com/articles/s41371-023-00789-0

João Tomé-Carneiro & Francesco Visioli (2023). A systematic review of the effects of plant-based diets on systolic blood pressure. *Nutrition Reviews, 81*(3), 405-418. https://pmc.ncbi.nlm.nih.gov/articles/PMC10224875/

Kelley, G. A., Kelley, K. S., & Franklin, B. (2023). Isometric exercise and systolic blood pressure: A meta-analysis. *Sports Medicine, 53*(8), 1239-1252. https://link.springer.com/article/10.1007/s40279-023-01789-6

O'Brien, E., & Stergiou, G. (2023). Advances in cuffless blood pressure monitoring technology. *Journal of Clinical Hypertension, 25*(4), 201-214. https://onlinelibrary.wiley.com/doi/10.1111/jch.14456

Pescatello, L. S., & Blumenthal, J. A. (2023). Long-term physical activity interventions for reducing systolic blood pressure. *Medicine & Science in Sports & Exercise, 55*(9), 1871-1880. https://journals.lww.com/acsm-msse/Fulltext/2023/09000/Long_Term_ Physical_Activity_ Interventions_for.14.aspx

Salman Mohammadi, Soheil Hassanipour, Hamed Delam, Hossein -Ali Nikbakht, Zohreh Ghaem Far, Donya Firoozi, Elnaz Ghaem Far, Hamid Abdolazimi Ali Ghaffarian – Bahraman (2023) Prevalence of hypertension in Iran: An updated systematic review and meta -analysis of community -based studies Caspian Journal of Internal Medicine 2023 (Autumn); 14(4): 607-617 https://eprints.lums.ac.ir/4540/1/Prevalence%20of%20hypertension%20in%20Iran%20An%20updated%20systematic.pdf

Schnabel, R. B., & Lubitz, S. A. (2023). Genetic insights into systolic blood pressure regulation via Mendelian randomization. *Circulation: Genomic and Precision Medicine, 16*(2), 101-112. https://www.ahajournals.org/doi/10.1161/CIRCGEN.123.001011

Williams, B., & MacDonald, T. M. (2023). Dual and triple therapy options for managing resistant hypertension. *The Lancet Hypertension, 9*(3), 303-315. https://www.thelancet.com/journals/lanhyp/article/PIIS2666-6022(23)00056-7/fulltext