

RESEARCH

Open Access



Machine learning approaches in identifying factors associated with hypertension and undiagnosed hypertension in adults in rural areas of Bangladesh

Farzana Akhter Bornee^{1†}, Mohammad Rocky Khan Chowdhury^{2,3†}, Md Zahidul Islam^{4,5}, Zarin Raihana⁶, Hasina Akhter Chowdhury³, Manzur Kader⁷ and Mamunur Rashid^{8*}

Abstract

Background Hypertension is a major cause of death and disability, and undiagnosed cases are particularly dangerous as they can cause severe damage without timely treatment. The aim of the study was to identify risk factors for hypertension and undiagnosed hypertension in rural areas of Bangladesh using advanced Machine Learning (ML) algorithms.

Methods This study involved 1,603 respondents, selected through a cross-sectional survey using a multistage cluster random sampling technique. Four ML algorithms, including Gradient Booster (GB), Logistic Regression (LR), Random Forest (RF) and Support Vector Machine (SVM), were used in this study. Risk factors for hypertension and undiagnosed hypertension were identified using the best-performing ML model, selected based on metrics such as accuracy, sensitivity, specificity, precision, F1 score, receiver operating characteristics-area under the curve (ROC-AUC), and calibration plot.

Results The prevalence of hypertension was 15.5%, slightly higher than the 15.4% for undiagnosed hypertension. In predicting the risk of both hypertension and undiagnosed hypertension, the LR model outperformed other ML models across most evaluation metrics. For hypertension, it achieved higher performance in terms of precision (0.580), F1 score (0.550), ROC-AUC (0.729; 95% CI: 0.677–0.779), and calibration. Similarly, for undiagnosed hypertension, the LR model showed better precision (0.580), ROC-AUC (0.596; 95% CI: 0.537–0.654), and calibration compared to other models. The risk factors for hypertension and undiagnosed hypertension differed notably. Key risk factors for undiagnosed hypertension included being overweight or obese, the absence of chronic diseases or cardiovascular disease (CVD), being male, non-use of tobacco, older age (above 50 years), being currently married, non-smoking status, having diabetes, and having no formal education.

[†]Farzana Akhter Bornee and Mohammad Rocky Khan Chowdhury contributed equally to this work.

*Correspondence:
Mamunur Rashid
mamunur.rashid@hig.se

Full list of author information is available at the end of the article



Conclusion The findings emphasize the urgent need for enhanced national and regional public health initiatives to improve the detection and awareness of hypertension in rural Bangladesh. Further research is important to validate the findings.

Keywords Hypertension, Unaware, Machine learning, Rural area, Bangladesh

Text box 1. Contributions to the literature

- Evidence on factors contributing to undiagnosed hypertension in rural Bangladeshi population remains limited, emphasizing the critical need for sustained and systematic monitoring of non-communicable disease trends in rural settings.
- Applications of machine learning (ML) offer strong potential to predict hypertension status and to identify the most influential factors associated with both hypertension and undiagnosed hypertension.
- Factors associated with hypertension and undiagnosed hypertension among rural Bangladeshi populations differ substantially, reflecting variations in health awareness, access to care, and underlying risk profiles.

Introduction

Hypertension has emerged as a critical global public health concern, affecting approximately one in four adults worldwide [1]. Hypertension and pre-hypertension cause over 8.5 million global deaths per year from stroke, heart disease, vascular, and renal conditions [2, 3]. The growing burden of hypertension is not affecting all regions equally. Low and middle-income countries, especially in South Asia, are seeing a significant rise in hypertension cases [4]. Disturbing trends, especially in countries like Bangladesh, underscore the urgent need for a comprehensive understanding of the risk factors associated with hypertension [5]. In Bangladesh as a developing nation, epidemiological and demographic shifts have contributed to a concerning rise in hypertension rates alongside other non-communicable diseases (NCDs) [6]. In 2018, the prevalence of hypertension among adults exceeds 28% in Bangladesh [7].

In Bangladesh, hypertension affects about 25% of the rural population, where nearly 67% of the country's people reside [8]. The rural population is facing an increasing incidence of chronic diseases, such as hypertension, which not only affects their health but also threatens their livelihoods [9]. In these communities, the lack of healthcare access, coupled with inadequate media coverage and information, significantly increases the risk of hypertension [10]. Furthermore, the financial strain of managing hypertension, particularly the high out-of-pocket (OOP) expenses, adds another layer of difficulty for rural populations [11]. In Bangladesh, healthcare is primarily financed through OOP payments, unlike developed countries, where government funding and private insurance cover most costs. This reliance on direct household spending places a heavy financial strain on rural families, often forcing them to delay treatment, borrow money, or

sell assets. The contrast underscores the lack of financial protection mechanisms in Bangladesh and highlights the vulnerability of rural households to economic hardship when faced with medical expenses [12, 13].

The existing evidence indicates that achieving awareness and control of blood pressure poses challenges in rural areas [14]. While many individuals with hypertension can manage their high blood pressure (BP) through lifestyle modifications, medications, or a combination of both, a significant number of adults are unaware of their condition. Often, symptoms do not manifest until damage has occurred [15]. That gap in hypertension status in rural areas is attributed to broader social and demographic determinants [5]. Research on the prevalence and factors influencing hypertension and undiagnosed hypertension in rural Bangladesh is limited. The national prevalence of hypertension and undiagnosed hypertension was approximately 12.2% and 15.8%, respectively [7]. In a rural district of Bangladesh, it was estimated that over 80% of hypertensive patients were previously unaware [16]. Furthermore, various factors, including social, demographic, and clinical aspects, were identified as associated with undiagnosed hypertension [17].

Most of the previous studies have relied on traditional regression models to identify factors associated with hypertension and undiagnosed hypertension in Bangladesh [7, 16, 18]. While the application of Machine Learning (ML) in predicting hypertension in Bangladesh has been evident [19–21], its use in determining factors associated with hypertension and undiagnosed hypertension, especially in rural communities, remains unexplored. Moreover, evidence suggests that ML methods outperform traditional regression models in prediction and factor selection [22]. Over time, machine learning has become increasingly popular in public health because of growing data availability, advances in computational power, and the need for predictive tools. It is now widely applied in areas such as disease surveillance, risk prediction, health behavior analysis, and resource allocation, reflecting a steady shift toward data-driven decision-making in the field [23, 24].

Given the diverse characteristics of rural communities across different regions within a country, it is essential to conduct more nuanced, community-based studies to capture local health dynamics effectively. Moreover, consistent monitoring of non-communicable disease patterns over time in rural settings is vital, particularly given their worsening trends. Such efforts are increasingly important

to inform targeted interventions, allocate resources efficiently, and address the rising burden of hypertension in these often-underserved populations. Therefore, this study aims to identify risk factors for hypertension and undiagnosed hypertension using ML algorithms in rural areas of Bangladesh, thereby enhancing our understanding of these conditions and enabling effective management to improve public health outcomes.

Materials and methods

Survey management

The recruitment process for field personnel and the training workshop (S1), the data collection instruments (S2), the pilot survey (S3), and the quality assurance (S4) were presented in the Additional file.

Sample size calculation/power analysis

The study collected blood pressure (BP) data from village residents and determined the sample size using the formula: $n = Z\alpha/2 \times p(1 - p) \div d^2$. Here, n represents the required sample size, p is the prevalence of hypertension in rural Bangladesh (approximately 27% based on a recent study), and d is the desired accuracy level (set at 3%) (8). With a standard normal deviate ($Z1 - \alpha/2$) of 1.96 for a 95% confidence level, the initial sample size was calculated to be 841 participants. Adjusting for a design effect of 1.75 to accommodate the multi-stage study design, the final sample size was determined to be 1472 [25, 26]. After adjusting the non-response rate, which was around 8.2%, the information was collected from 1603 participants in the actual study (adjusted = $1472 / (1 - 0.082) = 1603$).

Study design and sampling

This was a cross-sectional study conducted from August to November 2021. Geographically, Bangladesh is administratively divided into eight divisions: Barisal, Chittagong, Dhaka, Khulna, Mymensingh, Rajshahi, Rangpur, and Sylhet. The study followed a multi-stage random selection process, beginning with Khulna division. Within Khulna, Jhenaidah district, one of its ten districts, was randomly chosen. From there, Jhenaidah Sadar Upazila, one of the district's six Upazilas, was selected. The process continued with the selection of Naldanga Union, one of 14 Unions in Jhenaidah Sadar. Participants were drawn from 18 villages in Naldanga Union, and four additional villages were excluded once the target sample size was reached. At least 80 participants were interviewed from each village.

This study utilized the 'Kish Grid' method to select households and identify study participants [27]. The 'Kish Grid' is a systematic sampling technique used in survey research to ensure representative samples from geographically dispersed populations. It involves listing

eligible individuals at an address by age and selecting participants based on the address's serial number, ensuring equal selection probability within households. This method minimizes bias, ensures randomness, and provides comprehensive population coverage. Effective implementation requires consideration of factors like grid size, resolution, and potential clustering effects.

This method allowed interviewing only one member per selected household. Interviewers began with the household closest to the Union Parishad center and enrolled participants based on inclusion criteria: adults aged 18 and above residing in the household, consenting to participate [28]. Pregnant women, the mentally disabled, and individuals who had surgery within the last three months were excluded to minimize bias. If a household member declined or the household was inaccessible, it was marked as a refusal, and the next household was approached. Proportional representation was ensured by balancing respondents by sex (male and female) and age groups (young adults, adults, middle-aged, and older adults).

Outcome measures

The study's outcome variables were hypertension and undiagnosed hypertension. Hypertension was determined based on the BP classification criteria outlined in the Seventh Report of the Joint National Committee (JNC 7), where a systolic blood pressure (SBP) > 140 mmHg and/or diastolic blood pressure (DBP) > 90 mmHg indicated hypertension [29]. Participants were classified as hypertensive if they had a documented diagnosis from a registered medical professional or were taking prescribed antihypertensive medications at the time of data collection [30]. Undiagnosed hypertension referred to participants with an average SBP > 140 mmHg and/or DBP > 90 mmHg who were not undergoing antihypertensive treatment (determined by checking medical records, prescriptions or self-reported) at the time of data collection.

All anthropometric measurements and BP assessments were conducted by a registered medical nurse. Blood pressure was measured using a digital monitor (ELITE YE770A) equipped with a standard-sized cuff. The device was routinely calibrated by inspecting for any damage, connecting it to a certified reference manometer, and verifying accuracy at multiple pressure points (e.g., 50, 100, 150, and 200 mmHg). Readings were required to fall within an acceptable error margin of ± 3 mmHg compared to the reference standard.

Participants were asked to avoid drinking coffee or smoking for at least 30 min before the measurement. They were then instructed to sit in a relaxed position, leaning back with their legs uncrossed and feet flat on the floor, for five minutes. BP was measured on the arm free of excess clothing, positioned at heart level, and the

same arm was used for all measurements to ensure consistency. Each participant's BP was recorded three times, with a 15-minute interval between the first and second, and between the second and third measurements. The mean BP with 95% confidence interval (CI) for each measurement sample was calculated for comparison (Additional file: Supplementary Table 1) The final BP value was calculated as the average of the three readings.

Independent variables

The study included independent variables derived from findings in the existing published literature [7, 9, 31–33], such as demographic characteristics, which includes gender, age, marital status, educational status, employment status; lifestyle factors, such as smoking history, chewing tobacco; anthropometric characteristics such as, body mass index (BMI) and waist-hip measurement; clinical characteristics such as diabetes, cardiovascular disease (CVD), other chronic diseases, family history of hypertension, presence of anxiety symptoms, and presence of depression symptoms. The operational definitions and measurement scales of the independent variables were provided in Additional file: Supplementary Table 2.

Statistical analysis

Baseline characteristics were assessed through descriptive analysis. The relationship between the independent and dependent variables was examined using a Chi-square test, with significance set at $p < 0.05$. The first-degree interaction effect was examined among independent variables using logistic regression before final analysis.

Selection of the best ML algorithm

ML approaches are highly effective in handling complex, non-structured big data from patient databases. They excel at identifying intricate patterns within large datasets containing numerous variables, enabling the exploration of multiple interactions and nonlinear relationships with the outcome, relationships that traditional statistical methods may struggle to explain [34–36]. From a widespread literature search, some ML algorithms were found superior in developing models to predict hypertension, which include Gradient Booster (GB), Logistic Regression (LR), Random Forest (RF) and Support Vector Machine (SVM) [21]. These ML techniques can select variables after model fit [37].

The ML model was trained on 60% of the dataset (the training dataset), and the remaining 40% was used for validation (the test dataset). The class imbalance of outcome was addressed using the Adaptive Synthetic (ADASYN) resampling technique. Each trained model was optimised with hyperparameter tuning using a 5-fold cross-validation protocol during development

(Additional file: Supplementary Table 3). Subsequently, the trained model was validated in 40% validation dataset. The best ML algorithm was selected based on the comparison of performance metrics (accuracy, sensitivity/recall, specificity, precision, F1 score, receiver operating characteristic (ROC) area under the curve (ROC-AUC) curve with 95% CI, and calibration plot (Additional file: Supplementary Table 4) in the validation dataset [38, 39]. A schematic presentation of the best ML model selection is shown in Fig. 1.

Potential risk factor selection and interpretability in the model

The most significant risk factors for predicting hypertension and undiagnosed hypertension were identified using the SHapley Additive exPlanations (SHAP) technique, developed by Lundberg and Lee [40], which has proven effective in highlighting key risk factors. SHAP is based on game theory principles [41] and local explanation methods [42], and it estimates the contribution of each feature to the model's overall decision-making process.

Data analysis was carried out using statistical software packages, specifically Stata (version 17), R (latest version), and Python (latest version).

Results

Background profile

Approximately 26.7% of the respondents were aged 50 years and older. The gender distribution among respondents was nearly equal, with 51% being male. About 27.4% of respondents reported having no formal education. A smoking habit was present in 19.7% of respondents during the data collection period. Around 56.8% of respondents were classified as overweight or obese. Diabetes affected around 8% of respondents, while 5.4% had some form of CVD. A family history of hypertension was reported by 27.8% of respondents. Symptomatic anxiety was documented in approximately 16% of individuals, and around 18% experienced symptoms of depression. The detailed background profile was presented in Table 1.

Prevalence of hypertension and undiagnosed hypertension

The prevalence of hypertension stood at 15.5%, while undiagnosed hypertension was slightly lower at 15.4% in rural areas. Both hypertension and undiagnosed hypertension showed significantly higher rates among individuals aged 50 years and above, with percentages of 28.1% and 19.2%, respectively. Similarly, there was a significant increase in hypertension and undiagnosed hypertension among individuals classified as obese, with proportions of 26.3% and 21.6%, respectively (Table 2).

Moreover, hypertension was particularly prevalent among females (18.6%), individuals with no formal

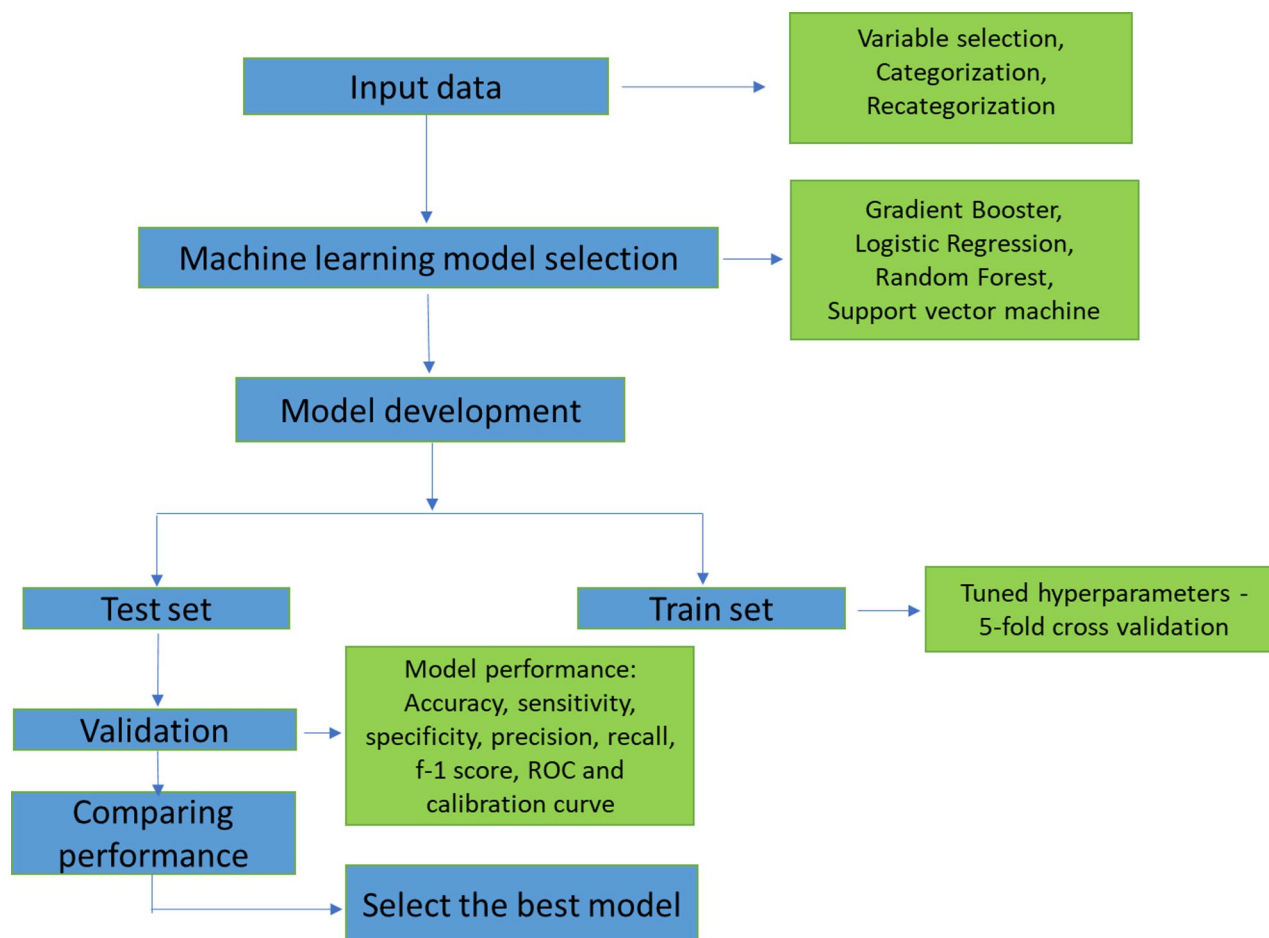


Fig. 1 Schematic presentation of selecting the best Machine Learning (ML) model

education (20.1%), those currently not employed (19.9%), non-smokers or former smokers (17.4%), those with a high waist-hip ratio (21.5%), individuals with diabetes (34.9%), CVD (40.7%), a family history of hypertension (26.0%), as well as those with symptomatic anxiety (24.2%) and depression (25.3%) (Table 2).

Best ML model selection

In predicting hypertension, the LR model outperformed other ML models across most metrics which included precision (0.580), F1 score (0.550) and ROC-AUC score (0.729, 95% CI: 0.677–0.779), as shown in Table 3 and Fig. 2. The calibration curve indicated improved calibration for LR model in predicting hypertension (Fig. 3). The accuracy (0.631), sensitivity (0.690) and specificity (0.620) of LR model were also high.

In predicting undiagnosed hypertension, the LR model outperformed the other ML models across most evaluation metrics, including a precision of 0.580 and a ROC-AUC score of 0.596 (95% CI: 0.537–0.654), as presented in Table 3 and Fig. 2. The calibration curve (Fig. 3.) also demonstrated that the LR model had better calibration

compared to the other ML models for this prediction. The LR model showed moderate accuracy (0.581), sensitivity (0.535), specificity (0.589) and F1 score (0.490) in predicting undiagnosed hypertension. Furthermore, the confusion matrix for all models was presented in the Additional file: Supplementary Fig. 1.

Influential variables and their interpretability in the RF model

In (Fig. 4), the SHAP plot provides an overview of the influence of variables, showing the direction of their impact through the distribution of red and blue dots. The vertical bar on the right represents the range of values for the independent variables, with blue indicating low values and red indicating high values. A positive SHAP value for a variable on the x-axis indicates a higher likelihood of the outcome, while a negative value suggests a lower likelihood, such as being normotensive. For instance, the variable “age” is coded as 0 for ≤ 30 years, 1 for 31–50 years and 2 for > 50 years (Additional file: Supplementary Table 1). A value of 0 corresponds to the blue colour, and 2 corresponds to red on the bar.

Table 1 Background characteristics of the study participants

Factors	Frequency	Percentages
<i>Demographic characteristics</i>		
Age (in years)		
30 and less	377	23.5
31–50	799	49.8
50 and above	427	26.7
Sex of the respondent		
Male	818	51.0
Female	785	49.0
Educational status		
No formal education	439	27.4
Primary	409	25.5
Secondary	572	35.7
Higher	183	11.4
Employment status		
Employed or self-employed	714	44.5
Housewife	713	44.5
Retired or student	176	11.0
Marital status		
Never married, separated, divorced or widowed	195	12.2
Currently married	1408	87.8
<i>Lifestyle factors</i>		
Smoking history		
Never and past smoker	1287	80.3
Current smoker	316	19.7
Chewing tobacco		
Never and past user	1218	76.0
Current user	385	24.0
<i>Anthropometric characteristics</i>		
Body mass index		
Normal and underweight	693	43.2
Overweight/pre-obesity	591	36.9
Obese	319	19.9
Waist–hip ratio		
Low	453	28.3
Moderate	313	19.5
High	837	52.2
<i>Clinical characteristics</i>		
Diabetes		
No	1474	92.0
Yes	129	8.0
Cardiovascular diseases		
No	1517	94.6
Yes	86	5.4
Other chronic disease		
No	1454	90.7
Yes	149	9.3
Family history of hypertension		
No	1157	72.2
Yes	446	27.8
Presence of anxiety symptoms		
No	1,347	84.0
Yes	256	16.0
Presence of depression symptoms		

Table 1 (continued)

Factors	Frequency	Percentages
No	1314	82.0
Yes	289	18.0

Age emerged as a significant risk factor for hypertension, as the red dots shifted towards positive SHAP values, indicating that individuals who belonged to the age group >50 years were more likely to develop hypertension. Similarly, factors such as having CVD, having a family history of hypertension, being overweight or obese, having higher waist-hip measurements, having diabetes, being a non-smoker, being retired or a student (currently not employed), having symptoms of depression, and having a chronic disease were associated with hypertension.

Similarly, in (Fig. 4.), influential factors associated with undiagnosed hypertension included being overweight or obese, absence of chronic diseases or CVD, being male, non-use of tobacco, older age (above 50 years), being currently married, non-smoking status, having diabetes, and having no formal education.

Discussion

There is limited documentation on the status of hypertension and undiagnosed hypertension in rural Bangladesh. Hypertension already diagnosed (15.5%) and undiagnosed (15.4%) hypertension were found to have similar prevalence rates and were notably higher in rural areas. Therefore, the present study explored key factors linked to both hypertension and undiagnosed hypertension among rural Bangladeshi adults, utilising cutting-edge ML algorithms previously recognised as the most effective in predicting hypertension across various populations [21]. Among ML models, the LR outperformed others in predicting hypertension and undiagnosed hypertension. Furthermore, hypertension and undiagnosed hypertension were linked to a range of demographic, behavioural, anthropometric, and clinical factors.

In identifying risk factors of hypertension and undiagnosed hypertension, LR model demonstrated remarkable performance across various metrics, including precision, F1 score, ROC-AUC, and calibration, outperforming other ML models. This aligns with prior research, which also highlighted the potential of the LR method for predicting hypertension, regardless of whether it is diagnosed or undiagnosed [43–45]. In this study, LR outperformed more complex models, such as RF and GB, despite having moderate discriminative ability. This is likely because the dataset was small, contained relatively weak predictors, and modest event rates, conditions under which LR, being a simpler and more controlled model, can perform more reliably. LR is often more robust when the signal-to-noise ratio is modest.

It can handle modest signal strength effectively under limited sample sizes. In contrast, RF and GB require larger datasets with stronger signals to capture complex interactions and are more prone to overfitting when the predictive signal is low. As a result, LR provided more stable and generalizable predictions, even though overall performance remained limited. Additionally, studies have identified other algorithms, such as GB, SVM, and RF, as highly effective methods for predicting hypertension [21, 46–49].

The findings suggest that ML models have the potential to provide highly accurate predictions of clinical outcomes and may outperform traditional statistical methods. Integrating ML-based prediction models into rural healthcare in Bangladesh can significantly enhance early detection and resource allocation if embedded into existing workflows rather than creating parallel systems. Community health workers could use mobile apps powered by ML to screen patients during household visits, enabling faster triage and referrals, while mobile health platforms could deliver personalized alerts and reminders to at-risk individuals. Telemedicine services could be strengthened by ML pre-screening that prioritises urgent cases and provides structured risk scores to doctors, improving efficiency in settings with limited physician availability [50]. However, challenges such as incomplete local datasets, infrastructure gaps, digital literacy, and privacy concerns must be addressed through pilot programs, hybrid decision-making that combines AI with human judgment, training for health workers, and government partnerships.

When deploying ML-based health prediction systems in low-resource rural settings like Bangladesh, ethical and privacy considerations become critical to ensure community trust and long-term sustainability. Sensitive health and lifestyle data must be collected, stored, and shared with strict safeguards to prevent misuse or breaches, especially given the limited digital infrastructure and weak regulatory frameworks in rural areas. Transparent communication about how data will be used, coupled with community engagement, can help build trust and reduce fears of surveillance or exploitation. Additionally, ensuring that ML tools are designed with equity in mind, avoiding biases that disadvantage marginalized groups, is essential. By prioritising secure data practices, culturally sensitive implementation, and clear accountability mechanisms, these systems can strengthen healthcare delivery while respecting the rights and dignity of rural populations [50].

Table 2 Prevalence of hypertension and undiagnosed hypertension

Factors	HTN		Undiagnosed HTN	
	N (%)	p-value	N (%)	p-value
Age (in years)				
30 and less	10 (2.7)	< 0.001	41 (10.9)	0.005
31–50	119 (14.9)		124 (15.5)	
50 and above	120 (28.1)		82 (19.2)	
Sex of the respondent				
Male	103 (12.6)	0.001	140 (17.1)	0.053
Female	146 (18.6)		107 (13.6)	
Educational status				
No formal education	89 (20.1)	0.009	76 (17.3)	0.616
Primary	63 (15.4)		62 (15.2)	
Secondary	73 (12.8)		83 (14.5)	
Higher	24 (13.1)		26 (14.2)	
Employment status				
Employed or self-employed	80 (11.2)	< 0.001	115 (16.1)	0.309
Housewife	134 (18.8)		100 (14.0)	
Retired or student	35 (19.9)		32 (18.2)	
Marital status				
Never married, separated, divorced or widowed	32 (16.4)	0.718	22 (11.3)	0.089
Currently married	217 (15.4)		225 (16.0)	
Smoking history				
Never and past smoker	224 (17.4)	< 0.001	207 (16.1)	0.131
Current smoker	25 (7.9)		40 (12.7)	
Chewing tobacco				
Never and past user	185 (15.2)	0.498	198 (16.3)	0.095
Current user	64 (16.6)		49 (12.7)	
Body mass index				
Normal and underweight	62 (9.0)	< 0.001	70 (10.1)	< 0.001
Overweight/pre-obesity	103 (17.4)		108 (18.3)	
Obese	84 (26.3)		69 (21.6)	
Waist–hip measurement				
Low	29 (6.4)	< 0.001	61 (13.5)	0.388
Moderate	40 (12.8)		52 (16.6)	
High	180 (21.5)		134 (16.0)	
Diabetes				
No	204 (13.8)	< 0.001	221 (15.0)	0.119
Yes	45 (34.9)		26 (20.2)	
Cardiovascular diseases				
No	214 (14.1)	< 0.001	235 (15.5)	0.701
Yes	35 (40.7)		12 (14.0)	
Other chronic disease				
No	228 (15.7)	0.611	232 (16.0)	0.058
Yes	21 (14.1)		15 (10.1)	
Family history of hypertension				
No	133 (11.5)	< 0.001	180 (15.6)	0.790
Yes	116 (26.0)		67 (15.0)	
Presence of anxiety symptoms				
No	187 (13.9)	< 0.001	205 (15.2)	0.630
Yes	62 (24.2)		42 (16.4)	
Presence of depression symptoms				
No	176 (13.4)	< 0.001	201 (15.3)	0.791
Yes	73 (25.3)		46 (15.9)	
Total	249 (15.5)		247 (15.4)	

HTN hypertension

Table 3 Performance metrics of Machine Learning algorithms in predicting hypertension and undiagnosed hypertension

Outcomes	Metrics of algorithms	Algorithms							
		GB		LR		RF		SVM	
		Test	Train	Test	Train	Test	Train	Test	Train
Hypertension	Accuracy	0.582	0.684	0.631	0.736	0.846	0.846	0.712	0.915
	Sensitivity	0.730	0.789	0.690	0.800	0.324	0.532	0.230	0.950
	Specificity	0.555	0.580	0.620	0.671	0.806	0.854	0.817	0.878
	Precision	0.580	0.690	0.580	0.740	0.550	0.650	0.520	0.920
	F-1 score	0.520	0.680	0.550	0.730	0.550	0.670	0.529	0.910
Undiagnosed hypertension	Accuracy	0.580	0.667	0.581	0.680	0.629	0.895	0.671	0.884
	Sensitivity	0.566	0.759	0.535	0.716	0.364	0.957	0.273	0.914
	Specificity	0.490	0.572	0.589	0.642	0.678	0.830	0.764	0.852
	Precision	0.520	0.670	0.530	0.680	0.510	0.900	0.510	0.890
	F-1 score	0.450	0.660	0.490	0.680	0.490	0.890	0.510	0.880

GB Gradient Booster, LR Logistic Regression, RF Random Forest, SVM Support Vector Machine

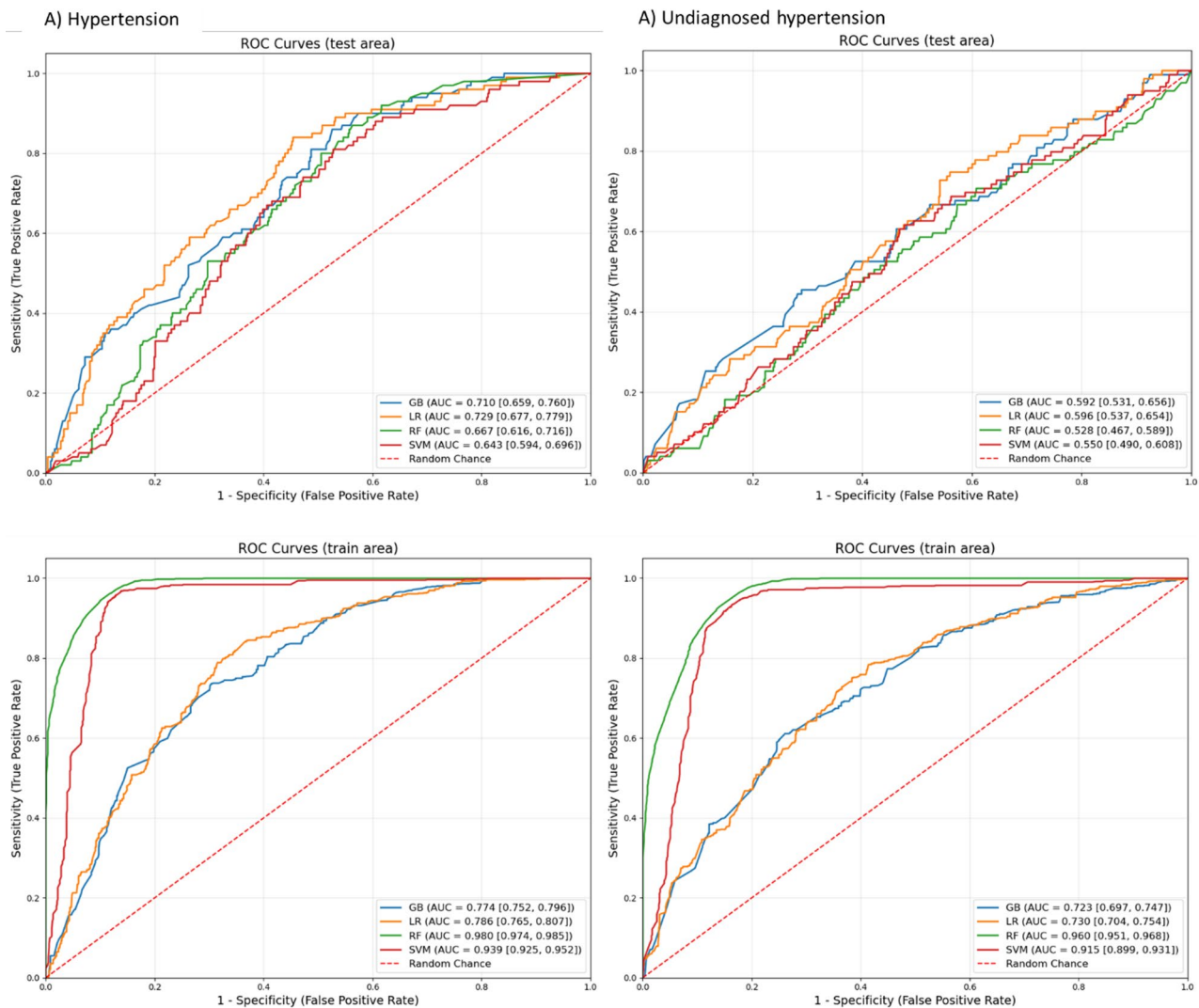


Fig. 2 Receiver operating characteristics (ROC) curve in predicting hypertension and undiagnosed hypertension. Note: GB, Gradient Booster; LR, Logistic Regression; RF, Random Forest; SVM, Support Vector Machine

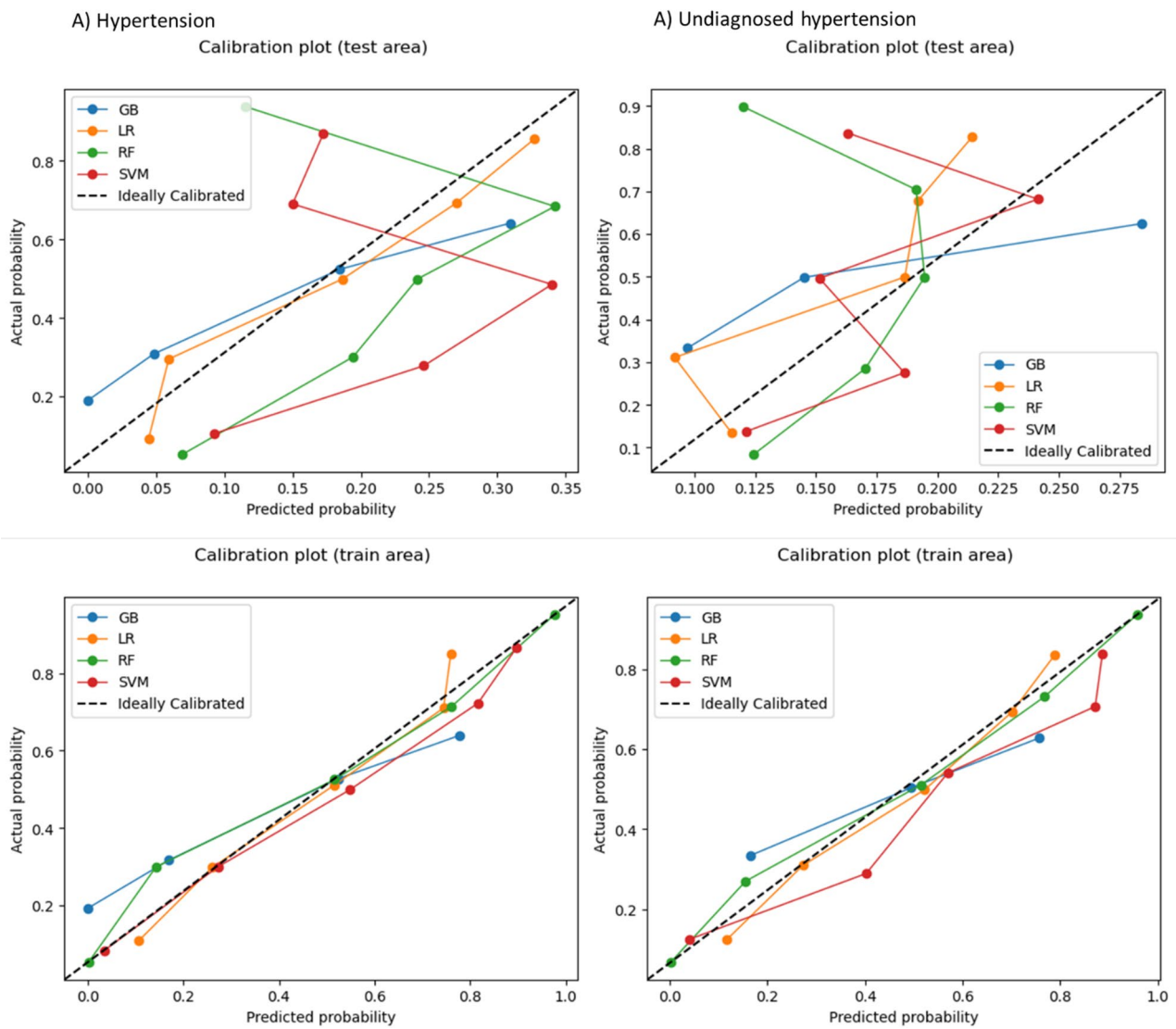


Fig. 3 Calibration curve in predicting hypertension and undiagnosed hypertension. Note: GB, Gradient Booster; LR, Logistic Regression; RF, Random Forest; SVM, Support Vector Machine

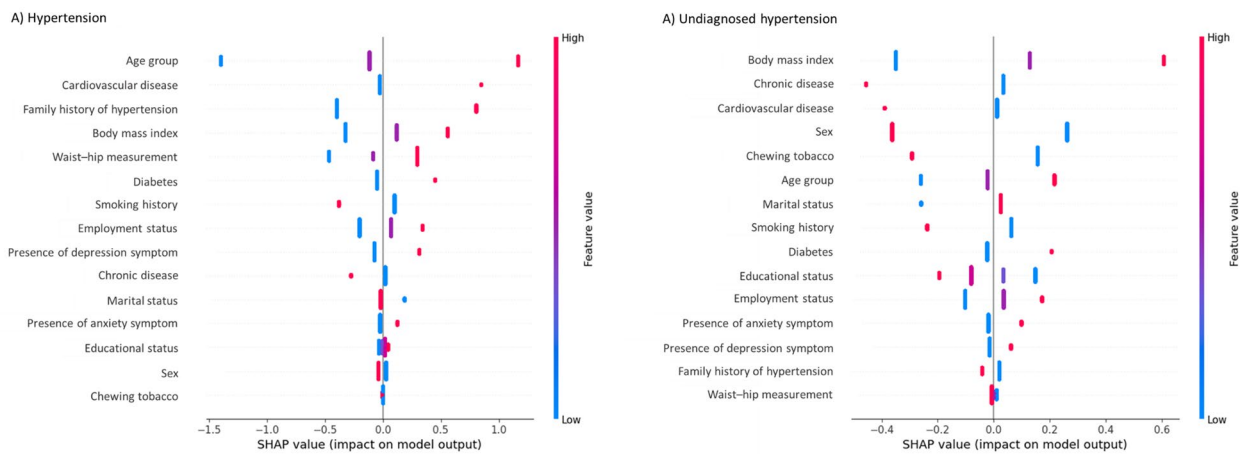


Fig. 4 Beeswarm plots for Logistic Regression model to explain the influential factors related to hypertension and undiagnosed hypertension

In this study, the LR algorithm identified several key factors associated with hypertension, including older age, having CVD, family history of hypertension, overweight or obesity, higher waist-hip measurement, having diabetes, non-smoking status, currently not employed, having symptoms of depression, and chronic disease. Regardless of the use of ML models, factors such as older age, currently not employed, being overweight, non-smoking status and having diabetes consistently emerged as significant contributors to hypertension in rural Bangladesh [5, 33, 51]. Similarly, the LR model selected being overweight or obese, absence of chronic diseases or CVD, being male, non-use of tobacco, older age (above 50 years), being currently married, non-smoking status, presence of diabetes, and having no formal education as influential factors related to undiagnosed hypertension. These findings are consistent with previous studies that primarily employed traditional regression analysis, identifying older age, being male, poor education, being currently married, overweight or obesity, having diabetes and comorbidities as significant risk factors for undiagnosed hypertension [7, 16, 46, 52, 53]. Overall, the risk factors for both hypertension and undiagnosed hypertension identified in this study were found to be important determinants of hypertension in several studies, regardless of their classification [7, 9, 31, 33]. These factors also demonstrated potential for predicting both hypertension and undiagnosed hypertension during ML model development.

The risk factors for hypertension and undiagnosed hypertension in rural Bangladeshi populations differed notably, which may be due to detection bias, health system interaction, health literacy, and personal health behaviour. There were only a few overlapping factors, such as, age over 50, being overweight or obese, having diabetes, and non-smoking status. Greater attention to these risk factors is crucial for better detection and management of both diagnosed and undiagnosed hypertension. Undiagnosed hypertension in rural Bangladeshi populations is influenced by a combination of individual, behavioural, and clinical factors [7, 16, 46, 52, 53]. Being overweight or obese and having diabetes increases the physiological risk of hypertension, yet individuals with these conditions may not seek or receive regular health check-ups [54]. In contrast, those without other illnesses or CVD are less likely to engage with healthcare services, reducing the chances of hypertension being detected [55]. Older adults are more prone to hypertension but may attribute symptoms to normal aging, leading to underdiagnosis [56]. Men, particularly those who are married, often engage less in preventive healthcare due to gender norms and family responsibilities, which lead to the perception that seeking care is unnecessary unless severe symptoms occur [57]. Interestingly, non-smokers

and non-tobacco users may be perceived as lower risk by both themselves and healthcare providers, potentially reducing the likelihood of being screened [58]. Furthermore, individuals with no formal education often have limited health literacy and awareness about hypertension and its potential symptoms and risks, which diminishes their likelihood of seeking preventive care [59]. These interconnected factors highlight systemic gaps in awareness, access, and targeted screening efforts, contributing to the persistence of undiagnosed hypertension in this population.

Consistent with the previous finding, approximately 50% of hypertensive patients in rural areas remain unaware of their hypertension status [16]. To address the high burden of undiagnosed hypertension in rural Bangladesh, targeted interventions are needed. These include implementing community-based screening programs, especially for high-risk groups like older adults, men, diabetics, and those with obesity or no formal education. Integrating blood pressure checks into existing chronic disease services and engaging male populations through workplace or market-based outreach can enhance early detection. Health education campaigns are essential for raising awareness among low-literacy populations, while equipping community health workers with training and equipment can improve outreach and referrals [60]. Additionally, mass screening during national health campaigns and strengthening primary healthcare services for routine hypertension monitoring are vital for sustainable control. Collectively, these strategies can bridge awareness, access, and detection gaps to reduce undiagnosed hypertension in rural settings.

This study has several strengths and limitations. Its main strength lies in the successful application of robust machine learning methods and a comprehensive evaluation of factors associated with hypertension and undiagnosed hypertension. However, in limitation, a slightly higher proportion of participants was aged ≥ 50 years than in the national population. Given that hypertension prevalence increases substantially with age, this skewed age structure may have led to an overestimation of the overall prevalence in this study. A locally popular, affordable BP device was used for measurements. To verify its reliability and accuracy, a subset of participants was re-evaluated using a clinically certified BP machine from a nearby clinic, and the readings were compared with those obtained from the survey device at the same time. BP was measured three times, and the average was used to determine hypertension, which could introduce inaccuracies because no variation between measurements was considered. Although taking the average of three readings is standard practice, the prevalence estimates may still be influenced by white-coat effects, as these were not accounted for due to difficulties obtaining clinical

records in community-based survey settings. The study was conducted in a single rural area of Bangladesh, which may limit its generalizability due to regional variations in healthcare access, diet, and socioeconomic factors. Caution should be exercised when applying these findings to other rural regions. Consequently, the findings should be interpreted with caution, as the external validity is limited.

Data were collected retrospectively using self-reports, which may have led to underreporting and recall bias. Due to the cross-sectional nature of data, causal relationships could not be assessed. In rural Bangladesh, collecting high-quality clinical and lifestyle data is challenging due to limited health infrastructure, inconsistent record-keeping, low digital literacy, and cultural barriers that affect accurate self-reporting. These gaps often lead to incomplete or unreliable datasets, which in turn weaken the predictive accuracy of ML models. Strengthening data collection through standardized digital tools for community health workers, culturally sensitive survey methods, and integration of mobile platforms for real-time data entry could greatly improve the richness and reliability of health information [61].

Although the LR model identified relevant factors, a ROC-AUC of less than 0.60 for predicting undiagnosed hypertension is generally regarded as non-predictive for screening or clinical decision-making, indicating the model performs only slightly better than chance. Although the LR model performed better than the other ML models, the overall discriminative performance across all models, particularly for undiagnosed hypertension, falls below thresholds typically considered acceptable for clinical prediction.

In real-world contexts, the low ROC-AUC for a model limits its direct applicability as a standalone tool for risk stratification or screening, as its low discrimination may lead to high false-positive and false-negative rates. However, this model can still provide value by highlighting the limitations of available predictors, identifying subgroups, such as older, obese males with low educational attainment, where predictive performance may be stronger, or serving as a baseline for future model development. This study did not compare the performance and findings of the ML-based LR model with those of the traditional LR model to avoid added complexity. Additionally, the ML models used cannot provide p-values or beta coefficients to evaluate associations between variables.

Conclusion

In Bangladesh, around half of the hypertensive patients were unaware of their status. The LR model demonstrated potential for improved predictions of both hypertension and undiagnosed hypertension

and identified influential factors. The risk factors for hypertension and undiagnosed hypertension differed notably, with only a few overlapping factors: age over 50 years, being overweight or obese, having diabetes, and non-smoking status. Other risk factors of undiagnosed hypertension included the absence of chronic diseases or CVD, being male, non-use of tobacco, being currently married, and having no formal education. These findings call for further evaluation of ML applications and risk factors. Moreover, the study highlights the urgent need for stronger national and regional public health initiatives to improve hypertension detection and awareness in rural Bangladesh.

Abbreviations

BMI	Body mass index
BP	Blood pressure
CVD	Cardiovascular disease
DBP	Diastolic blood pressure
GB	Gradient Booster
LR	Logistic Regression
ML	Machine Learning
NCDS	Non-communicable diseases
RF	Random Forest
ROC	Receiver Operating Characteristics
SBP	Systolic blood pressure
SHAP	SHapley Additive exPlanations
SVM	Support Vector Machine

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13690-026-01941-z>.

Additional file 1.

Acknowledgements

We would like to acknowledge the Monash University, Australia and Bangabandhu Sheikh Mujib Medical University, Bangladesh where the research was conducted.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Authors' contributions

FAB: Writing – Original Draft Preparation, Writing – Review & Editing. MRKC: Conceptualization, Data Curation, Formal Analysis, Methodology, Software, Supervision, Writing – Review & Editing. Writing – Review & Editing. MZL: Data Curation, Formal Analysis, Writing – Review & Editing. ZR: Data Curation, Writing – Review & Editing. HAC: Data Curation, Investigation, Project Administration, Writing – Review & Editing. MK: Writing – Review & Editing. MR: Supervision, Methodology, Writing – Review & Editing.

Funding

Open access funding provided by University of Gävle. Open access publication fees were covered by Dalarna University, Sweden, and received by the corresponding author, MK. The funders had no role in the study design, conduct, manuscript preparation, review, or approval process.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

This study was conducted as part of the “COVID-19 Vaccine Hesitancy and its Correlates in Rural Bangladesh” project. Ethical approval was granted by the Monash University Human Research Ethics Committee (project reference number: 29358). Consent was also obtained from local Bangladeshi authorities, including government officials, village leaders, and religious figures.

The research adhered to the principles outlined in the 2013 Declaration of Helsinki. Prior to conducting interviews, data collectors provided participants with detailed information about the study, including its purpose, data collection procedures, the voluntary nature of participation, the intended use of the data, and participants' privacy rights. Individuals who agreed to participate were given an explanatory statement, and any questions or concerns were addressed before participants signed the consent form and the interviews commenced.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Pediatrics, Bangabandhu Sheikh Mujib Medical University, Dhaka, Dhaka Division, Bangladesh

²Department of Population Science, Jatiya Kabi Kazi Nazrul Islam University, Mymensingh, Mymensingh Division, Bangladesh

³Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Clayton, Melbourne, VIC, Australia

⁴Department of Public Health, First Capital University of Bangladesh, Chuadanga, Khulna Division, Bangladesh

⁵Institute of Biological Sciences, University of Rajshahi, Rajshahi, Rajshahi Division, Bangladesh

⁶Department of Clinical Psychology, Faculty of Biological Sciences, University of Rajshahi, Rajshahi, Rajshahi Division, Bangladesh

⁷Department of Medical Science, School of Health and Welfare, Dalarna University, Falun, Sweden

⁸Faculty of Health and Working Life, Unit of Public Health Sciences, University of Gävle, Gävle, Sweden

Received: 10 October 2025 / Accepted: 2 May 2026

Published online: 18 May 2026

References

1. Hypertension. <https://www.who.int/news-room/fact-sheets/detail/hypertension>.
2. Olsen MH, Angell SY, Asma S, Boutouyrie P, Burger D, Chirinos JA, Damasceno A, Delles C, Gimenez-Roqueplo A-P, Hering D, et al. A call to action and a lifecourse strategy to address the global burden of raised blood pressure on current and future generations: the Lancet Commission on hypertension. *Lancet*. 2016;388(10060):2665–712.
3. Zhou B, Perel P, Mensah GA, Ezzati M. Global epidemiology, health burden and effective interventions for elevated blood pressure and hypertension. *Nat Reviews Cardiol*. 2021;18(11):785–802.
4. Kario K, Hoshida S, Mogi M. Hypertension treatment up-date on world hypertension day 2024: current status and future prospects in Asia. *Hypertens Res*. 2024;47(7):1763–5.
5. Parvin S, Akter S, Hossain MI, Ali MS, Soni MSM. Residential variations in hypertension prevalence and trends among adults in Bangladesh. *Res Health Serv Reg*. 2024;3(1):3.
6. Ahsan Karar Z, Alam N, Kim Streatfield P. Epidemiological transition in rural Bangladesh, 1986–2006. *Global health action*. 2009;2(1):1904.
7. Hossain A, Suhel SA, Chowdhury SR, Islam S, Akther N, Dhor NR, Hossain MZ, Hossain MA, Rahman SA. Hypertension and undiagnosed hypertension among Bangladeshi adults: Identifying prevalence and associated factors using a nationwide survey. *Front Public Health*. 2022;10:1066449.
8. Khan MN, Oldroyd JC, Chowdhury EK, Hossain MB, Rana J, Renzetti S, Islam RM. Prevalence, awareness, treatment, and control of hypertension in Bangladesh: Findings from National Demographic and Health Survey, 2017–2018. *J Clin Hypertens*. 2021;23(10):1830–42.
9. Khanam R, Ahmed S, Rahman S, Kibria GMA, Syed JRR, Khan AM, Moin SMI, Ram M, Gibson DG, Pariyo G, et al. Prevalence and factors associated with hypertension among adults in rural Sylhet district of Bangladesh: a cross-sectional study. *BMJ Open*. 2019;9(10):e026722.
10. Akter M. H <>Kabir 2023 Health Inequalities in Rural and Urban Bangladesh: The Implications of Digital Health. *Mayo Clin Proceedings: Digit Health* 1 2 201–2.
11. Rahman T, Gasbarro D, Alam K. Financial risk protection against noncommunicable diseases: trends and patterns in Bangladesh. *BMC Public Health*. 2022;22(1):1835.
12. Nazmul MH, Al-Amin AQ, Howlader SR, Kabir MA. Paying out of pocket for healthcare in Bangladesh—a burden on poor? *Iran J public health*. 2015;44(7):1024.
13. Rahman T, Gasbarro D, Alam K, Alam K. Rural–urban disparities in household catastrophic health expenditure in Bangladesh: a multivariate decomposition analysis. *Int J Equity Health*. 2024;23(1):43.
14. Lall D, Engel N, Devadasan N, Horstman K, Criel B. Challenges in primary care for diabetes and hypertension: an observational study of the Kolar district in rural India. *BMC Health Serv Res*. 2019;19(1):44.
15. Lackland DT, Weber MA. Global burden of cardiovascular disease and stroke: hypertension at the core. *Can J Cardiol*. 2015;31(5):569–71.
16. Islam FM, Bhuiyan A, Chakrabarti R, Rahman MA, Kanagasingam Y, Hiller JE. Undiagnosed hypertension in a rural district in Bangladesh: The Bangladesh Population-based Diabetes and Eye Study (BPDES). *J Hum Hypertens*. 2016;30(4):252–9.
17. Teshome DF, Balcha SA, Ayele TA, Atnafu A, Gelaye KA. Undiagnosed hypertension and its determinants among hypertensive patients in rural districts of northwest Ethiopia: a mediation analysis. *BMC Health Serv Res*. 2023;23(1):222.
18. Haider MR, Gupta RD. Inequalities in Undiagnosed Hypertension Among Adult Population in Bangladesh: Evidence from a Nationally Representative Survey. *High Blood Press Cardiovasc Prev*. 2022;29(1):57–64.
19. Islam MM, Rahman MJ, Chandra Roy D, Tawabunnahar M, Jahan R, Ahmed N, Maniruzzaman M. Machine learning algorithm for characterizing risks of hypertension, at an early stage in Bangladesh. *Diabetes Metab Syndr*. 2021;15(3):877–84.
20. Islam MM, Alam MJ, Maniruzzaman M, Ahmed NAMF, Ali MS, Rahman MJ, Roy DC. Predicting the risk of hypertension using machine learning algorithms: A cross sectional study in Ethiopia. *PLoS ONE*. 2023;18(8):e0289613.
21. Silva GFS, Fagundes TP, Teixeira BC, Chiavegatto Filho ADP. Machine Learning for Hypertension Prediction: a Systematic Review. *Curr Hypertens Rep*. 2022;24(11):523–33.
22. Tiruneh SA, Vu TTT, Rolnik DL, Teede HJ, Enticott J. Machine Learning Algorithms Versus Classical Regression Models in Pre-Eclampsia Prediction: A Systematic Review. *Curr Hypertens Rep*. 2024;26(7):309–23.
23. Hang C-N, Tsai Y-Z, Yu P-D, Chen J, Tan C-W. Privacy-enhancing digital contact tracing with machine learning for pandemic response: a comprehensive review. *Big Data Cogn Comput*. 2023;7(2):108.
24. Aklah Z, Al-Safi A, Abdali MH, Al-jabery K. A machine learning model for automated contact tracing during disease outbreaks. *Healthc Analytics*. 2025;7:100389.
25. Henry GT. *Practical sampling*. Sage; 1990;21.
26. Islam RM, Bell RJ, Billah B, Hossain MB, Davis SR. The prevalence of symptomatic pelvic floor disorders in women in Bangladesh. *Climacteric*. 2016;19(6):558–64.
27. Kish L. A Procedure for Objective Respondent Selection within the Household. *J Am Stat Assoc*. 1949;44(247):380–7.
28. Khalequzzaman M, Chiang C, Chowdhury SR, Yatsuya H, Al-Mamun MA, Al-Shoabi AAA, Hirakawa Y, Hoque BA, Islam SS, Matsuyama A, et al. Prevalence of non-communicable disease risk factors among poor shantytown residents in Dhaka, Bangladesh: a community-based cross-sectional survey. *BMJ Open*. 2017;7(11):e014710.
29. The Seventh Report of the Joint National Committee on Prevention. Detection, Evaluation, and Treatment of High Blood Pressure [<https://www.ncbi.nlm.nih.gov/books/NBK9630/>].
30. National Guidelines for Management of Hypertension in Bangladesh. [<https://iris.who.int/bitstream/handle/10665/279486/9789843368553-eng.pdf>].

31. Jahan I, Siddique SM, Anik AI, Salwa M, Mousum S, Towhid MI, Islam MT, Mroy WW, Khan MMH, Haque MA. Prevalence, risk factors, awareness, and control of hypertension: A cross-sectional study in an urban slum area of Bangladesh. *J Health Med Sci*. 2022;5(2):26–34.
32. Rahman MA, Parvez M, Halder HR, Yadav UN, Mistry SK. Prevalence of and factors associated with prehypertension and hypertension among Bangladeshi young adults: An analysis of the Bangladesh Demographic and Health Survey 2017–18. *Clin Epidemiol Global Health*. 2021;12:100912.
33. Khanam MA, Lindeboom W, Razzaque A, Niessen L, Milton AH. Prevalence and determinants of pre-hypertension and hypertension among the adults in rural Bangladesh: findings from a community-based study. *BMC Public Health*. 2015;15:203.
34. Goldstein BA, Navar AM, Pencina MJ, Ioannidis JP. Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. *J Am Med Inf Assoc*. 2017;24(1):198–208.
35. Shickel B, Tighe PJ, Bihorac A, Rashidi P. Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis. *IEEE J Biomedical Health Inf*. 2018;22(5):1589–604.
36. Cheng Y, Wang F, Zhang P, Hu J. Risk prediction with electronic health records: A deep learning approach. In: *Proceedings of the 2016 SIAM international conference on data mining*; 2016: SIAM; 2016: 432–440.
37. Mortazavi BJ, Bucholz EM, Desai NR, Huang C, Curtis JP, Masoudi FA, Shaw RE, Negahban SN, Krumholz HM. Comparison of Machine Learning Methods With National Cardiovascular Data Registry Models for Prediction of Risk of Bleeding After Percutaneous Coronary Intervention. *JAMA Netw Open*. 2019;2(7):e196835.
38. Hossin M, Sulaiman MN. A review on evaluation metrics for data classification evaluations. *Int J data Min Knowl Manage process*. 2015;5(2):1–11.
39. Nahm FS. Receiver operating characteristic curve: overview and practical use for clinicians. *Korean J Anesthesiol*. 2022;75(1):25–36.
40. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach, California, USA: Curran Associates Inc.; 2017: 4768–4777.
41. Štrumbelj E, Kononenko I. Explaining prediction models and individual predictions with feature contributions. *Knowl Inf Syst*. 2014;41(3):647–65.
42. Ribeiro MT, Singh S, Guestrin C. Why should i trust you? Explaining the predictions of any classifier. In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*; 2016; 2016: 1135–1144.
43. Islam SMS, Talukder A, Awal MA, Siddiqui MMU, Ahamad MM, Ahammed B, Rawal LB, Alizadehsani R, Abawajy J, Laranjo L, et al. Machine Learning Approaches for Predicting Hypertension and Its Associated Factors Using Population-Level Data From Three South Asian Countries. *Front Cardiovasc Med*. 2022;9:839379.
44. Asadullah M, Hossain MM, Rahaman S, Amin MS, Sumy MSA, Parh MYA, Hossain MA. Evaluation of machine learning techniques for hypertension risk prediction based on medical data in Bangladesh. *Indonesian J Electr Eng Comput Sci*. 2023;31(3):1794–802.
45. Kurniawan R, Utomo B, Siregar KN, Ramli K, Besral B, Suhatriil RJ, Pratiwi OA. Hypertension prediction using machine learning algorithm among Indonesian adults. *IAES Int J Artif Intell*. 2023;12(2):776–84.
46. Moreira MWL, Rodrigues JJPC, Oliveira AMB, Saleem K, Neto AJV. Predicting hypertensive disorders in high-risk pregnancy using the random forest approach. In: *2017 IEEE International Conference on Communications (ICC)*: 21–25 May 2017 2017; 2017: 1–5.
47. Elshawi R, Al-Mallah MH, Sakr S. On the interpretability of machine learning-based model for predicting hypertension. *BMC Med Inf Decis Mak*. 2019;19(1):146.
48. Ren Z, Rao B, Xie S, Li A, Wang L, Cui G, Li T, Yan H, Yu Z, Ding S. A novel predicted model for hypertension based on a large cross-sectional study. *Sci Rep*. 2020;10(1):10615.
49. Machine Learning Approaches for Estimating Prevalence of Undiagnosed Hypertension among Bangladeshi Adults. Evidence from a Nationwide Survey [<http://hdl.handle.net/10464/17629>].
50. Kenigson J. Statistical Remarks on Rural Healthcare: A Detailed Study of AI Medical Systems in Bangladesh. *J AI-Powered Med Innovations (International online ISSN 3078 – 1930)*. 2024;3(1):153–67.
51. Hasan M, Khan MSA, Sutradhar I, Hossain MM, Hossain M, Yoshimura Y, Choudhury SR, Sarker M, Mridha MK. Prevalence and associated factors of hypertension in selected urban and rural areas of Dhaka, Bangladesh: findings from SHASTO baseline survey. *BMJ Open*. 2021;11(1):e038975.
52. Prevalence of undiagnosed hypertension, and prehypertension among bankers of Bangladesh: a cross-sectional study [<https://doi.org/10.1101/2023.01.09.23284329>].
53. Ahmed S, Tariqujjaman M, Rahman MA, Hasan MZ, Hasan MM. Inequalities in the prevalence of undiagnosed hypertension among Bangladeshi adults: evidence from a nationwide survey. *Int J Equity Health*. 2019;18(1):33.
54. Welsh A, Hammad M, Piña IL, Kulinski J. Obesity and cardiovascular health. *Eur J Prev Cardiol*. 2024;31(8):1026–35.
55. Zhou B, Carrillo-Larco RM, Danaei G, Riley LM, Paciorek CJ, Stevens GA, Gregg EW, Bennett JE, Solomon B, Singleton RK, et al. Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. *Lancet*. 2021;398(10304):957–80.
56. Skinner TR, Scott IA, Martin JH. Diagnostic errors in older patients: a systematic review of incidence and potential causes in seven prevalent diseases. *Int J Gen Med*. 2016;9:137–46.
57. Mursa R, Patterson C, Halcomb E. Men's help-seeking and engagement with general practice: An integrative review. *J Adv Nurs*. 2022;78(7):1938–53.
58. Pankova A, Kralikova E, Fraser K, Lajka J, Svacina S, Matoulek M. No difference in hypertension prevalence in smokers, former smokers and non-smokers after adjusting for body mass index and age: a cross-sectional study from the Czech Republic, 2010. *Tob Induc Dis*. 2015;13(1):24.
59. Du S, Zhou Y, Fu C, Wang Y, Du X, Xie R. Health literacy and health outcomes in hypertension: An integrative review. *Int J Nurs Sci*. 2018;5(3):301–9.
60. Yusuf SS, Acharya K, Ahmed R, Ahmed A. Understanding general health service readiness and its correlates in the health facilities of Bangladesh: evidence from the Bangladesh Health Facility Survey 2017. *J Public Health*. 2022;30(12):2835–46.
61. Gupta S, Mamodiya U. Machine Learning Algorithms for Early Detection of Epidemics in Rural Settings. *Cognitive Security and Predictive Health Monitoring for Rural Development*. edn.: IGI Global Scientific Publishing; 2026. pp. 25–40.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.