

## Article

# Integrating Health Data-Driven Machine Learning Algorithms to Evaluate Risk Factors of Early Stage Hypertension at Different Levels of HDL and LDL Cholesterol

Pen-Chih Liao <sup>1,†</sup>, Ming-Shu Chen <sup>2,†</sup> , Mao-Jhen Jhou <sup>3</sup>, Tsan-Chi Chen <sup>4</sup> , Chih-Te Yang <sup>5</sup>   
and Chi-Jie Lu <sup>3,6,7,\*</sup> 

<sup>1</sup> Division of Cardiology, Cardiovascular Center, Far Eastern Memorial Hospital, New Taipei City 220, Taiwan

<sup>2</sup> Department of Healthcare Administration, College of Healthcare and Management, Asia Eastern University of Science and Technology, New Taipei City 220, Taiwan

<sup>3</sup> Graduate Institute of Business Administration, Fu Jen Catholic University, New Taipei City 242, Taiwan

<sup>4</sup> Department of Medical Research, Far Eastern Memorial Hospital, New Taipei City 220, Taiwan

<sup>5</sup> Department of Business Administration, Tamkang University, New Taipei City 251, Taiwan

<sup>6</sup> Artificial Intelligence Development Center, Fu Jen Catholic University, New Taipei City 242, Taiwan

<sup>7</sup> Department of Information Management, Fu Jen Catholic University, New Taipei City 242, Taiwan

\* Correspondence: 059099@mail.fju.edu.tw; Tel.: +886-2-2905-2973

† Pen-Chih Liao and Ming-Shu Chen are co-first authors and contributed equally to this work.



**Citation:** Liao, P.-C.; Chen, M.-S.; Jhou, M.-J.; Chen, T.-C.; Yang, C.-T.; Lu, C.-J. Integrating Health Data-Driven Machine Learning Algorithms to Evaluate Risk Factors of Early Stage Hypertension at Different Levels of HDL and LDL Cholesterol. *Diagnostics* **2022**, *12*, 1965. <https://doi.org/10.3390/diagnostics12081965>

Academic Editors: Shang-Ming Zhou, Haider Raza and Honghan Wu

Received: 14 July 2022

Accepted: 11 August 2022

Published: 14 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Purpose: Cardiovascular disease (CVD) is a major worldwide health burden. As the risk factors of CVD, hypertension, and hyperlipidemia are most mentioned. Early stage hypertension in the population with dyslipidemia is an important public health hazard. This study was the application of data-driven machine learning (ML), demonstrating complex relationships between risk factors and outcomes and promising predictive performance with vast amounts of medical data, aimed to investigate the association between dyslipidemia and the incidence of early stage hypertension in a large cohort with normal blood pressure at baseline. Methods: This study analyzed annual health screening data for 71,108 people from 2005 to 2017, including data for 27 risk-related indicators, sourced from the MJ Group, a major health screening center in Taiwan. We used five machine learning (ML) methods—stochastic gradient boosting (SGB), multivariate adaptive regression splines (MARS), least absolute shrinkage and selection operator regression (Lasso), ridge regression (Ridge), and gradient boosting with categorical features support (CatBoost)—to develop a multi-stage ML algorithm-based prediction scheme and then evaluate important risk factors at the early stage of hypertension, especially for groups with high-density lipoprotein cholesterol (HDL-C) and low-density lipoprotein cholesterol (LDL-C) levels within or out of the reference range. Results: Age, body mass index, waist circumference, waist-to-hip ratio, fasting plasma glucose, and C-reactive protein (CRP) were associated with hypertension. The hemoglobin level was also a positive contributor to blood pressure elevation and it appeared among the top three important risk factors in all LDL-C/HDL-C groups; therefore, these variables may be important in affecting blood pressure in the early stage of hypertension. A residual contribution to blood pressure elevation was found in groups with increased LDL-C. This suggests that LDL-C levels are associated with CPR levels, and that the LDL-C level may be an important factor for predicting the development of hypertension. Conclusion: The five prediction models provided similar classifications of risk factors. The results of this study show that an increase in LDL-C is more important than the start of a drop in HDL-C in health screening of sub-healthy adults. The findings of this study should be of value to health awareness raising about hypertension and further discussion and follow-up research.

**Keywords:** health data-driven; high-density lipoprotein cholesterol (HDL-C); low-density lipoprotein cholesterol (LDL-C); hypertension; machine learning

## 1. Introduction

Cardiovascular disease (CVD) is a major worldwide health burden today. Several large cohort studies, including the Framingham Heart Study, demonstrate that hypertension and dyslipidemia (high LDL cholesterol (LDL-C) and low HDL cholesterol (HDL-C)) are important risk factors of future CVD [1]. A systemic review and meta-analysis suggest that the lowering of blood pressure to the normotensive range should be considered for the prevention of CVD [2]. Hypertension and dyslipidemia have an additive effect on the incidence of coronary heart disease in subjects with both conditions compared to those with only one or the other [3]. In modern management of coronary heart disease and cerebrovascular disease, the lowering of both blood pressure and LDL-C is important [4–6].

In recent years, the definition of normotensive has become more stringent (less than 130/85 mmHg) [7], because raised blood pressure is the leading cause of death globally [8]. Several studies have found an association between CVD and early stage hypertension [9–11]. In addition, there is a positive association between abnormal serum cholesterol levels and hypertension [12,13]. It is noteworthy that early prediction of hypertension is an important issue for individuals with dyslipidemia. However, the relationship between early stage hypertension (or prehypertension) and dyslipidemia remains unclear [14]. The medication of blood statin effect pressure-lowering has been confirmed by a meta-analysis published paper [15], which showed a small reduction ( $-2.62$  mmHg) in systolic blood pressure (95% CI:  $-3.41$  to  $-1.84$ ;  $p < 0.001$ ). In addition, Borghi et al. [16] also found that better control of LDL-C is associated with lower antihypertensive treatment in a large cohort study. These studies suggested that LDL-C is associated with hypertension and early stage hypertension. The association between HDL-C and hypertension is poorly understood. The Framingham Heart Study considered HDL-C to have a cardio-protective effect. The relationship between HDL-C and blood pressure is less clear. A positive linear relationship has been reported [17], but some reports have shown a slightly U-shaped relationship, or an inverted J-shaped relationship [18,19].

As the risk factors of CVD, both hypertension and dyslipidemia are mostly mentioned. Blood pressure is an important leading indicator of health hazards, especially for CVD or other related chronic diseases. Predicting the presence of early stage hypertension could be provided the possibility to prevent future CVD or chronic diseases. As mentioned above, the result of most studies demonstrated the correlation between BP and either LDL-C or HDL-C. It means that there is a potential relationship between BP and dyslipidemia. Dyslipidemia and high LDL-C and/or low HDL-C are associated with atherosclerosis and could lead to a change in blood pressure. This study aimed to investigate the association between dyslipidemia and the incidence of early stage hypertension in a large cohort with normal blood pressure at baseline.

The application of data-driven machine learning (ML) algorithms to the analysis of healthcare data and/or medical records is not uncommon, and there is even an increasing trend in publications introducing artificial intelligence technology [20–23]. The advantages of ML algorithms include the effective investigation of complex relationships between risk factors and outcomes, and promising predictive performance with vast amounts of medical data [22–26]. Our study used five ML techniques—stochastic gradient boosting (SGB), multivariate adaptive regression splines (MARS), least absolute shrinkage and selection operator logistic regression (Lasso), ridge logistic regression (Ridge), and gradient boosting with categorical features support (CatBoost)—to develop a multi-stage ML algorithm-based prediction scheme.

SGB is a model creates multiple additive regression trees with the decision tree algorithm by combining bagging and boosting techniques [27]. MARS is a nonlinear spline regression and a non-parametric form of the regression analysis algorithm [28]. Lasso and Ridge are both improved conventional logistic regression models using shrinkage regularization techniques [29]. CatBoost is an algorithm of integrating gradient boosting and multiple categorical variables based on gradient boosting decision tree framework. [30]. These five ML methods have been widely used in various healthcare and/or medical

informatics applications [31–40] as they could generate more effective predictive models than classical logistic regression model. They have also successfully been applied the field of predicting hypertension [41–44]. For example, Chang et al. [41] constructed a multiple predictive model for hypertension and hyperlipidemia using MARS. Lee et al. [38] used CatBoost method to predict intracranial hypertension and arterial blood pressure in patients with acute phase traumatic brain injury. Ang et al. [42] applied Lasso method to predict non-contact hypertension by the facial characteristics data of subjects. Shan et al. [43] utilized the ridge method to evaluate intracranial hypertension in traumatic brain injury patient. Chai et al. [44] used SGB and CatBoost methods to develop adolescent hypertension prediction model based on anthropometric measurements data. This study aimed to investigate the association between dyslipidemia and the incidence of early stage hypertension in a large cohort with normal blood pressure at baseline. The proposed scheme was used for each of four subgroups grouped by HDL-C and LDL-C criteria to predict early stage hypertension, evaluate relatively important risk factors, and then integrate feature selection results.

## 2. Materials and Methods

### 2.1. Data

The subjects of this study were data tracked continuously for a long time in Taiwan. It belongs to the annual health examination data of sub-health groups. The data is of excellent quality and dozens of international journal papers have been published successively. In this study, health screening was applied to the data of sub-healthy adults. The research results are more helpful to provide government health units with policy directions for preventive population health and health promotion. The data sets used were sourced from the MJ Group (Taipei, Taiwan)—a major health screening center in Taiwan—for the years 2005 to 2017. Many studies from Taiwan published in international journals have used the MJ Health Checkup-Based Population Database (MJPD) and are collated in <http://www.mjhrf.org/main/page/resource/en/#resource07> (accessed on 18 April 2022). These include studies on metabolic syndrome [45–47] and chronic kidney disease [24,48]. The MJPD database includes data collected from four MJ clinics that carry out periodic health examinations of the center’s approximately 71,000 members. The database can be accessed by academic researchers on request. All the data sets used in this study were authorized by and received from the MJ Health Research Foundation (Approval No.: MJHRF-2016005A). The data application procedures are described at <http://www.mjhrf.org/main/page/release1/en/#release01> (accessed on 18 April 2022). In the case of ethical issues regarding the use of data in the database, the protocol of this study was evaluated and deemed acceptable by the Research Ethics Review Committee of Far Eastern Memorial Hospital (FEMH-IRB-107127-E, Protocol Version1, 15 February 2022) and the MJ Health Research Foundation, and approved by ClinicalTrials.gov (ID: NCT05225454). The study was conducted according to the guidelines of the Declaration of Helsinki and fulfilled the Institutional Review Board ethics requirements by anonymizing all data before analysis.

Figure 1 shows the subject identification process of this study. The data consisted of the health examination indices and questionnaire records of the 71,108 members in the MJPD database from 2005 to 2017. Table 1 shows the 27 health examination indices and questionnaire variables developed in this study. Because every member may have multiple records, only the latest records were analyzed for subjects who had undergone multiple health examinations. In all, 40,853 subjects were removed because they had missing data for certain variables. After data processing, 30,255 eligible subjects remained. Table 2 shows the demographics and statistical analysis of subjects’ characteristics.

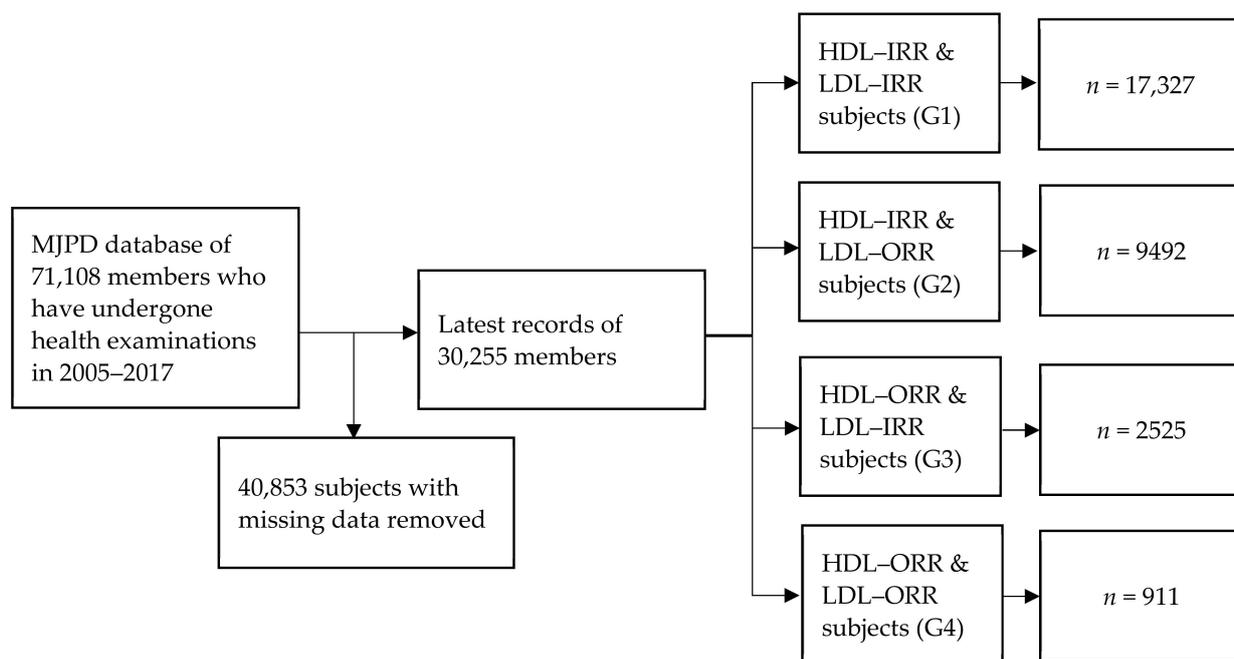


Figure 1. The enrollment flowchart for subject identification.

Table 1. Characteristics or laboratory indices of participants for predicting early stage hypertension.

Abbreviation	Variable (Unit)	Description/Reference Range (RR)
SEX	Gender (sex)	(1) Male; (2) Female
Age	Age (y/o)	Number; Years old (y/o)
MS	Marital status	(1) Single; (2) Married, remarried, cohabiting; (3) Divorced; (4) Widowed
EL	Education level	(1) No formal education; (2) Elementary school; (3) Secondary school; (4) High school; (5) College; (6) University; (7) Graduate school
FI	Family income (NTD)	(1) Unwaged; (2) ≤200,000; (3) 200,001–400,000; (4) 400,001–800,000; (5) 800,001–1,200,000; (6) 1,200,001–1,600,000; (7) 1,600,001–2,000,000; (8) >2,000,000
BMI	Body mass index (kg/m <sup>2</sup> )	Number; Body weight/Body height <sup>2</sup>
BF	Body fat (%)	Number; Data collection from ©OMRON: HBF-702t
WC	Waist circumference (cm)	Number; WC measured with a tape measure by SOP.
HC	Hip circumference (cm)	Number; HC measured with a tape measure by SOP.
WHR	Waist-to-hip ratio (%)	Number; Waist circumference/Hip Circumference
Hb	Hemoglobin (g/dl)	Number; RR: Male: 13.5 < Hb < 17.5; Female: 12.0 < Hb < 16.0
FPG	Fasting plasma glucose (mg/dL)	Number; RR: 70 < FPG < 100
TG	Triglycerides (mg/dL)	Number; RR: TG ≤ 150
T-Cho	Total cholesterol (mg/dL)	Number; RR: 130 < T-Cho < 200
FT4	Free thyroxine 4 (ng/dL)	Number; RR: 0.70 < FT4 < 1.48
TSH	Thyroid-stimulating hormone (μIU/mL)	Number; RR: 0.47 < TSH < 5.00
CRP	C-reactive protein (mg/dL)	Number, RR: CRP < 0.5
UP	Urine protein	Qualitative test; (1) none (2) trace (+/-) (3) + (4) ++ (5) +++ (6) ++++
CS	Current smoker	(1) Never; (2) Passive smoking; (3) Quit; (4) Occasional; (5) Addicted
AD	Alcohol drinker	(1) Never; (2) Quit; (3) 1–2 times a week; (4) 3–4 times a week; (5) 5–6 times a week; (6) Addicted

**Table 1.** Cont.

Abbreviation	Variable (Unit)	Description/Reference Range (RR)
CBN	Chews betel nut ( <i>Areca catechu</i> )	(1) Never; (2) Quit; (3) 1–3 times a week; (4) 4–5 times a week; (5) Addicted
MB	Mealtime behavior	(0) Irregular; (1) Regular
ET	Excise time (hours)	Time spent exercising in the past two weeks. (1) <0.5; (2) 0.5–1; (3) 1–2; (4) >2
ST	Sleep time (hours)	Average sleeping time at night. (1) <4; (2) 4–6; (3) 6–7; (4) 7–8; (5) 8–9; (6) >9
HDL-C	High-density lipoprotein cholesterol (mg/dL)	Number; RR: Male: HDL-C > 40; Female: HDL-C > 50. IRR–HDL and/or ORR–HDL: the different RR values for males and females were considered.
LDL-C	Low-density lipoprotein cholesterol (mg/dL)	Number, RR: LDL-C < 130
HTN	Hypertension in early stage # SBP: Systolic blood pressure (mmHg) DBP: Diastolic blood pressure (mmHg)	(0) Normal subjects: SBP < 120 and DBP < 80 (1) HTN subjects: SBP ≥ 120 and DBP ≥ 80

Note: The laboratory data in the subject databases were obtained using the same biochemical examination apparatus (an automatic biochemical analyzer was provided by Hitachi Medical Device Co., Ltd., ©Hitachi-7600, Tokyo, Japan). HDL: IRR and/or ORR, meaning data were within the reference range (IRR) and/or data were out of the reference range (ORR). #: Hypertension in the early stage, the criterion refers to the guidelines of the American Heart Association (AHA).

**Table 2.** The demographics statistical analysis of subjects’ characteristics.

Ordinal Variable (Unit)		N (%)	Ordinal Variable (Unit)		N (%)
Gender	Male	15,628 (51.65%)	Chews betel nut ( <i>Areca catechu</i> )	Never	28,784 (95.14%)
	Female	14,627 (48.35%)		Quit	1053 (3.48%)
Marital status	Single	4906 (16.22%)	Chews betel nut ( <i>Areca catechu</i> )	1–3 times a week	264 (0.87%)
	Married, remarried, cohabiting	22,948 (75.85%)		4–5 times a week	50 (0.17%)
	Divorced	1144 (3.78%)		Addicted	104 (0.34%)
	Widowed	1257 (4.15%)	Mealtime behavior	Irregular	8384 (27.71%)
Education level	No formal education	438 (1.45%)	Mealtime behavior	Regular	21,871 (72.29%)
	Elementary school	1958 (6.47%)		Excise time (hours)	<0.5
	Secondary school	1251 (4.13%)	0.5–1		13,513 (44.66%)
	High school	5655 (18.69%)	1–2		6409 (21.18%)
	College	6394 (21.13%)	>2		1972 (6.52%)
	University	9362 (30.94%)	Sleep time (hours)	<4	471 (1.56%)
Graduate school	5197 (17.18%)	4–6		7375 (24.38%)	
Family income (NTD)	Unwaged	1787 (5.91%)		Sleep time (hours)	6–7
	<200,000	2878 (9.51%)	7–8		6499 (21.48%)
	200,001–400,000	NA	8–9		NA
	400,001–800,000	6950 (22.97%)	>9		NA
	800,001–1,200,000	8256 (27.29%)	<b>Interval Variable (Unit)</b>		<b>Mean ± SD</b>
	1,200,001–1,600,000	4008 (13.25%)	Age (y/o)	47.25 ± 12.41	
1,600,001–2,000,000	2601 (8.60%)	Body mass index (kg/m <sup>2</sup> )	23.66 ± 3.59		
>2,000,000	3775 (12.48%)	Body fat (%)	26.76 ± 6.86		
Urine protein	none	29,364 (97.06%)	Waist circumference (cm)	78.84 ± 10.17	
	trace (+/–)	521 (1.72%)	Hip circumference (cm)	95.37 ± 6.31	
	+	254 (0.84%)	Waist-to-hip ratio (%)	0.83 ± 0.08	
	++	87 (0.29%)	Hemoglobin (g/dL)	14.14 ± 1.51	
	+++	29 (0.10%)	Fasting plasma glucose (mg/dL)	103.2 ± 19.35	
	++++	NA	Triglycerides (mg/dL)	115.79 ± 89.01	
Current smoker	Never	22,339 (73.84%)	Total cholesterol (mg/dL)	196.99 ± 34.40	
	Passive smoking	1066 (3.52%)	Free thyroxine 4 (ng/dL)	1.08 ± 0.15	
	Quit	2450 (8.10%)	Thyroid-stimulating hormone (μIU/mL)	1.73 ± 1.77	
	Occasional	1062 (3.51%)	C-reactive protein (mg/dL)	0.21 ± 0.39	
	Addicted	3338 (11.03%)			

**Table 2.** *Cont.*

Ordinal Variable (Unit)		N (%)	Ordinal Variable (Unit)		N (%)
Alcohol drinker	Never	24,832 (82.08%)	<b>Control Variable (Unit)</b>		<b>Mean ± SD</b>
	Quit	650 (2.15%)	High-density lipoprotein cholesterol (mg/dL)		59.01 ± 14.92
	1–2 times a week	3225 (10.66%)	Low-density lipoprotein cholesterol (mg/dL)		118.77 ± 32.2
	3–4 times a week	1045 (3.45%)	<b>Dependent Variable (Unit)</b>		<b>N (%)</b>
	5–6 times a week	NA	Hypertension in early stage (HTN)	SBP < 120 and DBP < 80	23,180 (76.62%)
	Addicted	503 (1.66%)		SBP ≥ 120 and DBP ≥ 80	7075 (23.38%)

Using the HDL-C and LDL-C reference ranges of the American Heart Association as baseline values, the data were then categorized into four subgroups based on whether the HDL-C and LDL-C values were within the reference range (IRR) or out of the reference range (ORR). An irregular HDL value was identified based on the reference range for the individual’s gender. The ORR values of HDL-C for men and women are more than 40 and 50 mg/dL, respectively. The ORR value of LDL is below 130 mg/dL. The four subgroups analyzed were as follows: 17327 HDL–IRR and LDL–IRR subjects (G1); 9492 HDL–IRR and LDL–ORR subjects (G2); 2525 HDL–ORR and LDL–IRR subjects (G3); and 911 HDL–ORR and LDL–ORR subjects (G4).

Past studies have not investigated whether the HDL/LDL indicators are normal or not divided into four groups, and have not applied the multivariate and different algorithms of machine learning tools.

*2.2. Proposed Multi-Stage Machine Learning Algorithm-Based Scheme*

This study developed a multi-stage hypertension prediction framework based on the machine learning algorithms for the four subgroups (G1, G2, G3, and G4) to identify, integrate, and examine the key risk factors for hypertension prediction in each subgroup. The overall procedure of the multi-stage machine learning algorithm-based scheme is shown in Figure 2. In the prediction framework, the first step was to collect the subjects’ health examination data from the MJPD database for analysis. The second step was to define the risk variables, identify subjects, and distinguish between the subjects in the four subgroups. The third step was to use the five learning algorithms (SGB, MARS, Lasso, Ridge, and CatBoost) to develop the prediction model for each subgroup in Table 1 using 24 risk factors as predictor variables (excluding HDL-C, LDL-C, and HTN) and the HTN as the target variable.

SGB implementation process sequentially generates many decision trees that are weak learners through multiple iterations so that each tree is trained based on the residual of the previous iteration [27,49]. The iterative process continues until the guideline of the maximum number of iterations or the convergence condition is reached. Finally, the cumulative results of many trees are obtained by weighed summation, and then the final robust model is determined.

MARS uses multiple piecewise linear segments (splines) with differing gradients. Its concept considers each sample as a knot and divides it into several sections for successive linear regression of the data within each section [28]. In the process for determining knots, a forward algorithm is used to select all possible basic functions and their corresponding knots, and a backward algorithm eliminates all basic functions to generate the best combinations of existing knots.

The Lasso Ridge methods share the same basic concept. The Lasso principle integrates the least absolute selection and shrinkage operator with L1 regularization, which can force compression of the coefficients of covariates making a minor contribution to the model to exactly zero to attain lower variance to reduce the problem of overfitting [50,51]. The main difference is the use of the L2 regularization technique to shrink model coefficients in Ridge. L2 regularization does not eliminate the coefficients or encourage sparse models. The addition of appropriate L2 penalties to the model shrinks all the coefficients to a nonzero

value or a value approaching zero, and then minimizes the sum of squared error, and further controls the trade-off between bias and variance to reduce overfitting [52].

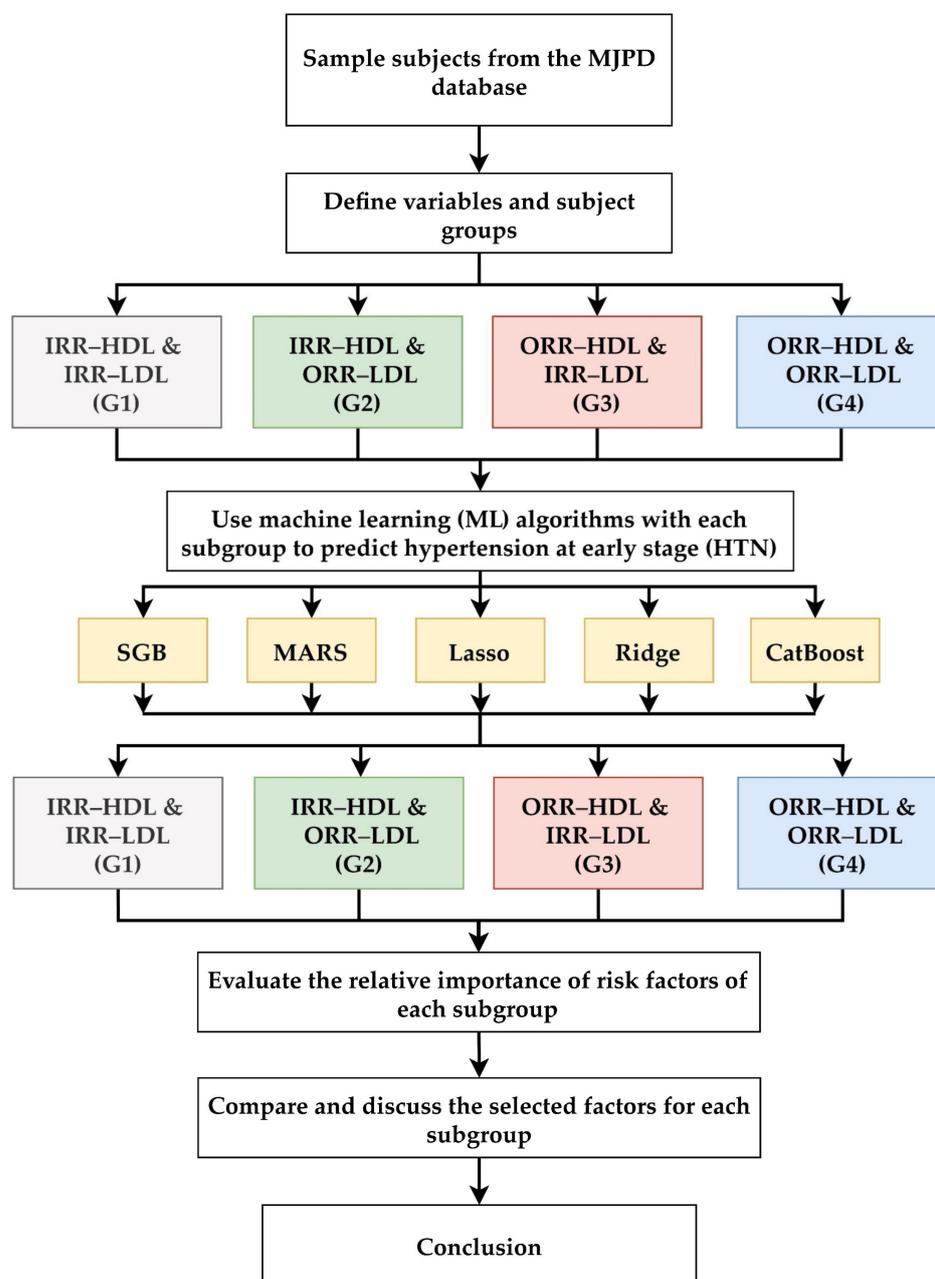


Figure 2. Proposed multi-stage machine learning algorithm-based scheme.

The CatBoost process is constructed using random multiple permutations generated to obtain gradients and correlations with the category variable [30]. As decision trees are weak learners, gradient boosting is successively fitted to each decision tree, where each tree is developed with a smaller loss compared to the previous one. Finally, it integrates all combinations and classification variables of the current tree into a sequence to generate the final model. CatBoost uses the ordered method of gradient boosting, which overcomes the prediction shift of the gradient estimation, and thus improves the algorithm’s accuracy and generalization [30,53].

When constructing each model, the data set was randomly partitioned 80% for the training data set and 20% for the testing data set. Model hyperparameter tuning and validation were executed using a 10-fold cross-validation approach on all the samples

available in the training data set. The model with the best hyperparameter was chosen as the final model. The predictive performance of the models was assessed using the following measures: sensitivity, specificity, and area under the receiver operating characteristic (ROC) curve (AUC). However, the models were highly influenced by data class distribution of these measures. Therefore, we also computed balanced accuracy (BA) and g-mean (GM) because they can be excellent measures for evaluating skewed data in any data class [54–57].

The used five ML models were implemented using R software of version 3.6.2 and RStudio software of version 1.1.453 (<http://www.R-project.org>; accessed on 25 May 2022; <https://www.rstudio.com/products/rstudio/>; accessed on 25 May 2022). Each algorithm was performed based on the related R packages. SGB, MARS, Lasso, Ridge, and CatBoost were implemented in “gbm” R package version 2.1.8 [58], “earth” R package version 5.3.1 [59], “glmnet” R package version 4.1-4 [60], and “catboost” R package version 1.0.6 [61], respectively. For developing efficient and exact SGB, MARS, Lasso, Ridge, and CatBoost models, the “caret” R package version 6.0-92 was used for all models to estimate the best hyperparameters [62].

In the fifth step, after obtaining the effective prediction models for all four subgroups as derived by SGB, MARS, Lasso, Ridge, and CatBoost, the relative importance of variables generated by each algorithm for each risk factor was also obtained. The variable importance of the most and least important risk factors was 100 and 0, respectively. Values can be repeated, that is, two or more variables can have similar variable importance. Because different machine learning algorithms have different model development principles and features, the variable importance values generated by the five algorithms for a single risk factor can differ. Within the same subgroup, a single robust and complete value for variable importance can be generated for each risk factor in order to facilitate subsequent comparison of variable rankings and identification of important risk factors. We generated a single consolidated value of variable importance based on the mean value of variable importance derived from the five machine learning models.

In the sixth step, we compared the important variables in G1 to G4 in order to examine and discuss their similarities and differences. The seventh and final step was to propose the conclusions of this study.

### 3. Results

Table 3 presents the model prediction performance of SGB, MARS, Lasso, Ridge, and CatBoost in relation to the four subgroups (HDL–IRR & LDL–IRR [G1], HDL–IRR & LDL–ORR [G2], HDL–ORR & LDL–IRR [G3], and HDL–ORR & LDL–ORR [G4]). Figure 3 presents the subgroup performance of the five models using ROC curves. To compare the predictive performance of the five methods in each of four subgroups, the DeLong’s test was used in this study to compare AUC values between the five ML models. DeLong’s test is one of the useful methods to determine if there is a statistically significant difference between the performances of the methods based on AUC values [63]. Table 4 shows pairwise comparisons of AUC values of the five used ML methods in all subgroups using DeLong’s test. It can be observed that the performance of any two ML methods is not significant different as all *p*-values in the table are above 0.05. That is, the prediction performances of the models were similar for each of the four subgroups

**Table 3.** Model performance in predicting hypertension for HDL and LDL IRR/ORR subgroups.

Subgroup, Total N = 30,255	Method	Sensitivity	Specificity	AUC	BA	GM
IRR–HDL & IRR–LDL (G1) <i>n</i> = 17,327 (57.27%)	SGB	0.625	0.770	0.762	0.698	0.694
	MARS	0.659	0.732	0.762	0.695	0.694
	Lasso	0.645	0.755	0.762	0.700	0.698
	Ridge	0.668	0.725	0.761	0.696	0.696
	CatBoost	0.605	0.791	0.764	0.698	0.692

Table 3. Cont.

Subgroup, Total N = 30,255	Method	Sensitivity	Specificity	AUC	BA	GM
IRR-HDL & ORR-LDL (G2) n = 9492 (31.37%)	SGB	0.595	0.715	0.705	0.655	0.652
	MARS	0.567	0.735	0.707	0.651	0.645
	Lasso	0.691	0.617	0.705	0.654	0.653
	Ridge	0.713	0.594	0.705	0.653	0.650
	CatBoost	0.682	0.613	0.703	0.648	0.647
ORR-HDL & IRR-LDL (G3) n = 2525 (8.35%)	SGB	0.642	0.695	0.702	0.668	0.668
	MARS	0.660	0.649	0.685	0.655	0.655
	Lasso	0.572	0.733	0.688	0.653	0.648
	Ridge	0.583	0.718	0.687	0.650	0.647
	CatBoost	0.575	0.741	0.693	0.658	0.652
ORR-HDL & ORR-LDL (G4) n = 911 (3.01%)	SGB	0.581	0.702	0.658	0.642	0.639
	MARS	0.728	0.575	0.649	0.651	0.647
	Lasso	0.706	0.596	0.653	0.651	0.649
	Ridge	0.478	0.787	0.653	0.633	0.613
	CatBoost	0.456	0.809	0.650	0.632	0.607

Note: SGB: stochastic gradient boosting; MARS: multivariate adaptive regression splines; Lasso: least absolute shrinkage and selection operator; Ridge: ridge regression; CatBoost: gradient boosting with categorical features support. IRR-HDL & IRR-LDL: subjects whose HDL and LDL data were within the reference range; IRR-HDL & ORR-LDL: subjects whose HDL data were within the reference range and LDL data were out of the reference range; ORR-HDL & IRR-LDL: subjects whose HDL data were out of the reference range and LDL data were within the reference range; ORR-HDL & ORR-LDL: subjects whose HDL and LDL data were both out of the reference range.

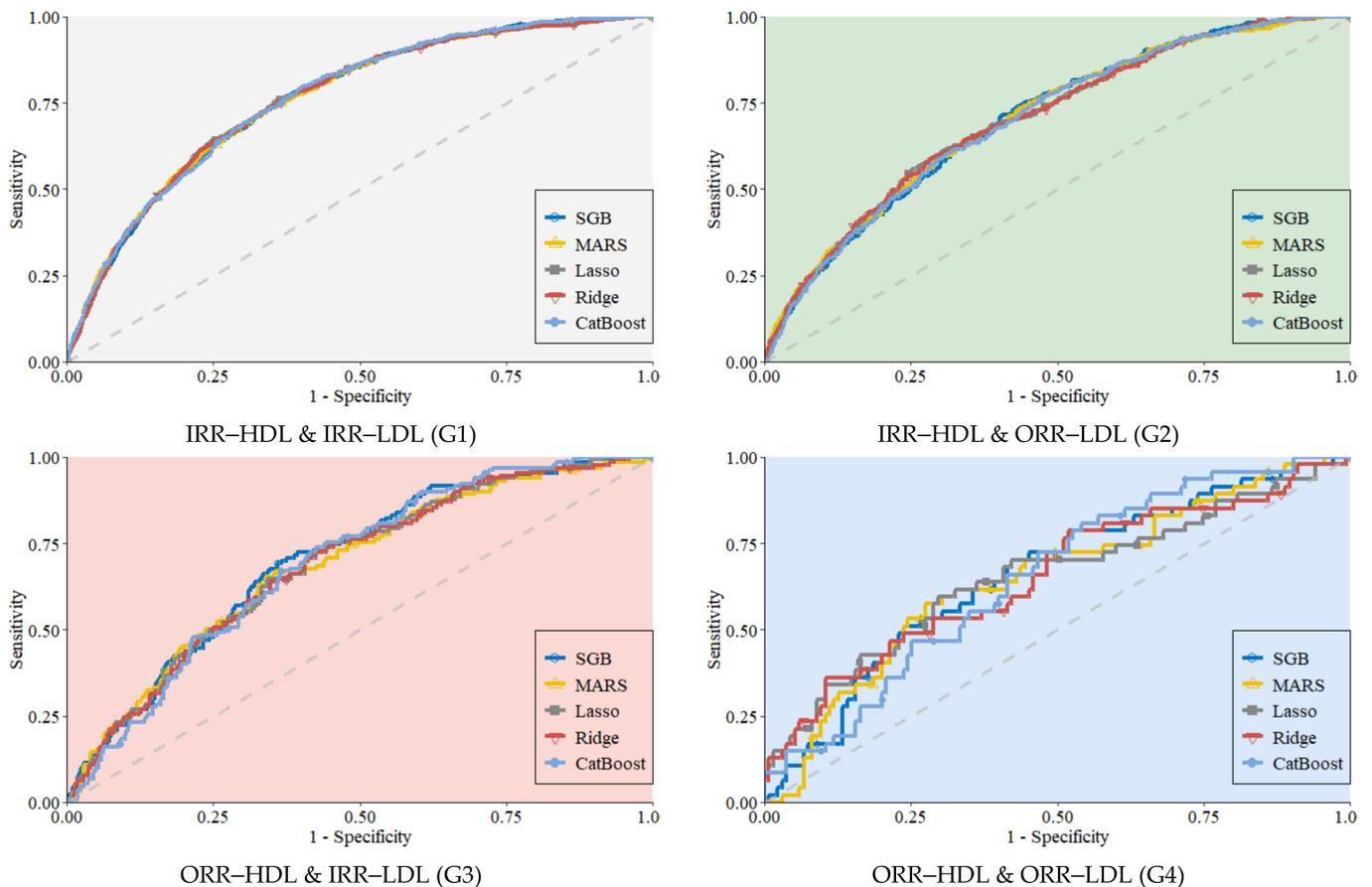


Figure 3. ROC curves of the five algorithms for each subgroup.

**Table 4.** Pairwise comparisons of AUC values of the five used ML methods in all subgroup using DeLong’s test.

Subgroup	Methods	SGB	MARS	Lasso	Ridge
IRR–HDL & IRR–LDL (G1)	SGB	–			
	MARS	0.467	–		
	Lasso	0.286	0.716	–	
	Ridge	0.164	0.517	0.085	–
	CatBoost	0.068	0.350	0.647	0.912
IRR–HDL & ORR–LDL (G2)	SGB	–			
	MARS	0.643	–		
	Lasso	0.874	0.778	–	
	Ridge	0.957	0.711	0.494	–
	CatBoost	0.589	0.410	0.588	0.664
ORR–HDL & IRR–LDL (G3)	SGB	–			
	MARS	0.273	–		
	Lasso	0.319	0.857	–	
	Ridge	0.288	0.933	0.477	–
	CatBoost	0.436	0.653	0.742	0.933
ORR–HDL & ORR–LDL (G4)	SGB	–			
	MARS	0.774	–		
	Lasso	0.899	0.904	–	
	Ridge	0.906	0.910	0.992	–
	CatBoost	0.865	0.967	0.960	0.957

Note: The numbers in table are the corresponding *p*-values.

However, the results differed between subgroups. For G1, the AUC of each algorithm was greater than 0.761 and was the highest among the subgroups. This shows that model prediction accuracy was highest for G1. Specifically, the Ridge algorithm had the highest sensitivity at 0.668, CatBoost had the highest specificity at 0.791 and the highest AUC at 0.764, while Lasso generated the highest values for BA and GM at 0.700 and 0.698, respectively. Lasso was comparatively the best prediction algorithm for G1.

For G2, the AUC of each algorithm was greater than 0.703 and was the second highest among the subgroups. Specifically, the Ridge algorithm had the highest sensitivity at 0.713, MARS had the highest specificity at 0.735 and the highest AUC at 0.707, SGB had a high BA at 0.655, and Lasso had the highest GM at 0.653. MARS was comparatively the best prediction algorithm for G2.

For G3, the AUC of each algorithm was greater than 0.68 and was the third highest among the subgroups. Specifically, the MARS algorithm had the highest sensitivity at 0.660, CatBoost the highest specificity at 0.741, and SGB the highest AUC, BA, and GM values at 0.702, 0.668, and 0.668, respectively. SGB was comparatively the best prediction algorithm for G3.

For G4, the AUC of each algorithm was greater than 0.649 and was the lowest among the subgroups. This shows that prediction of G4 was more difficult compared to the other subgroups. Specifically, the MARS algorithm had the highest sensitivity and BA values at 0.728 and 0.651, respectively; CatBoost had the highest specificity at 0.809; SGB had the highest AUC at 0.658; and Lasso had the highest BA and GM values at 0.651 and 0.649, respectively. Lasso was comparatively the best prediction algorithm for G4.

In general, even though the overall prediction performance differed between subgroups, all five machine learning algorithms had promising and similar performance in hypertension prediction.

The variable importance generated by the five machine learning algorithms provides high reference value because of the similarity in prediction performance of the models. However, the variable importance of the same risk factor differed between algorithms. To account for the variable importance generated by every algorithm, we derived the mean importance of each risk factor based on the five variable importance values.

Figure 4 shows the individual variable importance values generated by the five algorithms for each risk factor in the four subgroups. The 10 risk factors with the highest variable importance are presented for each subgroup in decreasing order of mean importance. For example, in G1, SGB, MARS, Ridge, and Lasso chose waist-to-hip ratio (WHR) as the most important variable, with an importance value of 100. On the other hand, CatBoost determined WHR to be of moderate importance at 41.7, although it is the most important risk factor in G1, with a mean importance of 88.3. Similarly, the second most important risk factor in G1 was age, with a mean variable importance of 46.4. Among the five models, age was chosen by CatBoost as the most important variable, with an importance value of 100, while SGB and MARS determined it to be of moderate importance at 67.7 and 64.5, respectively. However, Ridge and Lasso determined age to be the least important variable, with an importance of 0. In general, age remained the second most important variable in G1. Using the same concept and method, we were able to derive the ranking of variable importance for each of the four subgroups, as shown in Table 5.

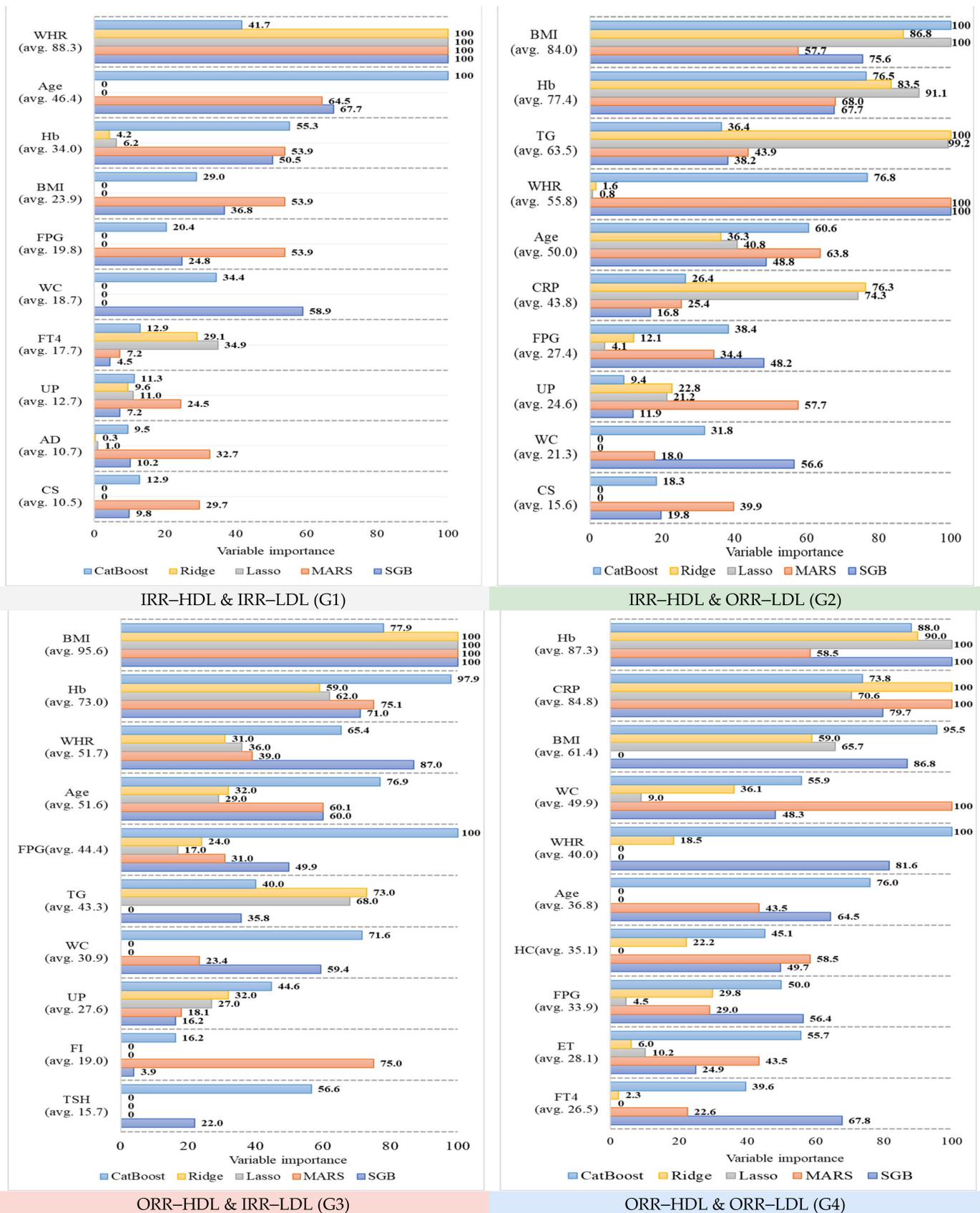
Table 5 shows that the ranking of variable importance differs between subgroups. For example, the three most important variables (in decreasing order of mean variable importance) in G1 were WHR, age, and hemoglobin (Hb), whereas in G2 they were body mass index (BMI), Hb, and triglycerides (TG), in G3 they were BMI, Hb, and WHR, and in G4 they were Hb, C-reactive protein (CRP), and BMI. The similarities and differences between the 10 most important variables in the four subgroups will be elaborated in the Discussion section.

To distinguish between subgroups, the overall degree of similarity between the importance rankings of all 24 prediction variables in the four subgroups was represented by the correlation coefficients (*R*) of variable importance ranking, as shown in Table 6. The closer the *R* value was to 1, the more similar were the variable importance rankings of two subgroups and the more distant *R* was from 1, the less similar were the variable importance rankings of the subgroups.

**Table 5.** Ranking of the top ten most important variables of the four subgroups.

Rank \ Subgroup	IRR-HDL & IRR-LDL (G1)	IRR-HDL & ORR-LDL(G2)	ORR-HDL & IRR-LDL (G3)	ORR-HDL & ORR-LDL (G4)
1	WHR	BMI	BMI	Hb
2	Age	Hb	Hb	CRP
3	Hb	TG	WHR	BMI
4	BMI	WHR	Age	WC
5	FPG	Age	FPG	WHR
6	WC	CRP	TG	Age
7	FT4	FPG	WC	HC
8	UP	UP	UP	FPG
9	AD	WC	FI	ET
10	CS	CS	TSH	FT4

Note: AD: alcohol drinker; BMI: body mass index; CRP: C-reactive protein; CS: current smoker; ET: excise time; FI: family income; FPG: fasting plasma glucose; FT4: free thyroxine 4; Hb: hemoglobin; HC: hip circumference; TG: triglycerides; TSH: thyroid-stimulating hormone; UP: urine protein; WC: waist circumference; WHR: waist-to-hip ratio. Note: G1: group whose LDL-C and HDL-C were all within the reference range; G2: group whose LDL-C started to rise but HDL-C was still within the reference range; G3: group whose HDL-C started to decrease but LDL-C was still within the reference range; G4: group whose HDL-C and LDL-C values were out of the reference range.



**Figure 4.** The variable importance generated by the generated by the five algorithms for each risk factor in the four subgroups.

**Table 6.** Correlation coefficients (*R*) of the variable importance ranking orders of the four groups.

Subgroup	IRR-HDL & IRR-LDL (G1)	IRR-HDL & ORR-LDL(G2)	ORR-HDL & IRR-LDL (G3)	ORR-HDL & ORR-LDL (G4)
IRR-HDL & IRR-LDL (G1)	1			
IRR-HDL & ORR-LDL (G2)	0.622	1		
ORR-HDL & IRR-LDL (G3)	0.633	0.899	1	
ORR-HDL & ORR-LDL (G4)	0.371	0.707	0.602	1

Note: G1: group whose LDL-C and HDL-C were all within the reference range; G2: group whose LDL-C started to rise but HDL-C was still within the reference range; G3: group whose HDL-C started to decrease but LDL-C was still within the reference range; G4: group whose HDL-C and LDL-C values were out of the reference range.

The results of the analysis showed that the group whose LDL-C began to rise but whose HDL-C was still within the reference range (IRR-HDL & ORR-LDL [G2]) had a risk factor ranking similar to that of the group whose HDL-C began to decrease but whose LDL-C was still within the reference range (ORR-HDL & IRR-LDL [G3]), and the types of data were also similar ( $R = 0.899$ ). However, in the comparison of G2 and G3 with the group whose HDL-C and LDL-C values were outside the reference range (ORR-HDL & ORR-LDL [G4]) shows that the correlation coefficient for G2 vs. G4 ( $R = 0.707$ ) was higher than that for G3 vs. G4 ( $R = 0.602$ ). This means that the group with abnormal LDL-C (G2) was more similar to the group with abnormal HDL-C and LDL-C (G4) than the group with abnormal HDL-C (G3); therefore, the beginning of abnormal LDL-C is a leading indicator and the start of the rise in LDL-C has reference value for prediction of prehypertension.

#### 4. Discussion

This study identified risk factors that have utility in the prediction of hypertension in different dyslipidemia groups. We applied several predictive models using machine learning algorithms, and the results obtained with the different models were similar.

Hypertension is a worldwide health burden, with high prevalence in those with cardiovascular disease. According to the World Health Organization, about 17 million people die from CVD worldwide, and about 9.4 million die from hypertension. The prevalence of hypertension is about 29% worldwide and is expected to increase from 26% in 2000 to 29.2% in 2025. Several studies have advanced an association between early stage hypertension (or prehypertension) and CVD, but the relationship between blood pressure and mortality is controversial [64–67]. The inconsistent results may be related to the age of participants, associated metabolic risk factors (e.g., abnormal lipid profiles), and the definition of early stage hypertension (prehypertension). As there is a positive association between blood pressure and cardiovascular morbidity or mortality [2], the definition and management of hypertension should be more aggressive [4–7]. It is well known that positive associations between CVD and high blood pressure (BP) or dyslipidemia were identified as early as half a century ago in the Framingham Heart Study. Dyslipidemia, either an increase in LDL-C or a decrease in HDL-C, also plays an important role in the development of CVD. Meanwhile, the biological interrelation between hypertension and LDL-C [13,68] or HDL-C [12,69,70] has been documented. The structural and functional change in LDL-C and HDL-C, inflammation, and oxidative stress may be associated with vascular atherosclerotic processes, and lead to elevation of blood pressure [71].

The relationships between HDL-C concentration categories and blood pressure are U- or J-shaped [69,70,72–74]. In the Kanagawa Investigation of Total Checkup Data from the National Database-9 study, Nakajima et al. [70] found inverted J-shaped relationships between HDL-C and odds ratios for hypertension ( $\geq 140/90$  mmHg) using the logistic regression analysis method, and both low and extremely high HDL-C concentrations are associated with high blood pressure within both sexes. In the South-West Seoul (SWS) Study, the elderly population with prehypertension combined with low HDL-C showed a twofold higher risk of all-cause mortality (HR: 2.01; 95% CI: 1.11–3.64) [69]. These studies showed that low HDL-C concentration is positively related to high blood pressure, but a linear relationship under extremely high HDL-C is not found. In clinical trials with cholesteryl ester transfer protein (CETP) inhibitors, a substantial increase in HDL-C concentration did

not show a protective effect against CVD events. In addition, a slight increase in systolic blood pressure of 1.2–5.4 mmHg has been shown after intervention [72,73]. The tendency to develop hypertension is correlated with HDL-C subfraction (HDL-3 concentration) and total HDL-C concentration [74]. As mentioned above, it is difficult to predict early stage hypertension with only HDL-C. In health screening data for sub-healthy adults, the group in which HDL-C was beginning to decrease but LDL-C remained within the reference range was smaller than the group in which LDL-C was beginning to rise but HDL-C remained within the reference range. In this study (Table 6), we found that LDL-C is a leading indicator, and rising LDL-C is a reference for predicting prehypertension. This finding agrees with the results of previous studies that showed that the relationship between HDL-C concentration and blood pressure is U- or J-shaped.

It is basic knowledge that the higher the LDL-C level will increase the risk of developing cardiovascular disease (CVD). With intervention medical trials, using lipid-lowering agents to reduce LDL-C had shown consistent reductions in major CVDs [4–6]. Otsuka et al. [75] showed the development of hypertension according to LDL-C quintiles in Asian populations. Their results indicated the risk of hypertension was 1.27 times higher in the highest quintile compared to the lowest quintile [75]. Most previous studies identified a relationship between LDL-C and CVD, but really did not make sure the incidence of hypertension.

This is the first study demonstrating an association between dyslipidemia and the risk of incident hypertension. Otsuka et al. proposed several mechanisms for dyslipidemia and the increased risk of hypertension. First, dyslipidemia, may impair endothelial function and regulation of blood pressure by disrupting the production of nitric oxide. Second, by reducing baroreflex sensitivity, dyslipidemia may predispose individuals to the development of hypertension. Third, dyslipidemia decreases the distensibility of large elastic arteries. This decrease may reduce the wind vessel effect, and then increase systolic blood pressure. Fourth, a lack of physical activity or regular exercise and a high-fat daily diet promotes obesity. The adipose tissue excessively secretes adipocytokines, and the cytokines result in insulin resistance and subsequent activation of the sympathetic nervous system and the renin-angiotensin system in obese individuals. Those biological changes have been confirmed and reported to increase blood pressure and raise incident hypertension.

The presence of dyslipidemia in subjects with early stage hypertension (prehypertension) can significantly increase the risk of cardiovascular mortality [69]. Individuals with dyslipidemia and elevated blood pressure have metabolic syndrome. In previous studies, metabolic syndrome was associated with a higher risk of CVD mortality in middle-aged or elderly populations [76,77]. In subjects with type 2 diabetes, a target blood pressure of 120 mmHg, compared to 140 mmHg, did not reduce the rate of fatal and nonfatal CVD events in the ACCORD study [78]. However, in subjects without diabetes, the prognostic benefit of blood pressure control was clarified in the SPRINT trial [79]. In the same study, with a target SBP of less than 120 mm Hg, compared to less than 140 mm Hg, the results showed that in lower rates of fatal and nonfatal major cardiovascular events and all-cause death. During the follow-up period (median, 3.26 years) of this clinical trial, 25% of subjects showed a lower relative risk of cardiovascular-related outcomes, including the composite outcomes of myocardial infarction, stroke, acute coronary syndrome not resulting in acute or chronic myocardial infarction, acute decompensated heart failure, death from cardiovascular causes, etc. Additionally, the rates of lots other important outcomes in the treatment group, including death from cardiovascular causes (reduce 43% relative risk), heart failure (reduce 38% relative risk), and death from any cause (reduce 27% relative risk), was lower than the control group [79]. These results indicate that it is valuable to aggressively treat individuals with early stage hypertension. In addition, subjects with early stage hypertension and dyslipidemia are at a greater risk of mortality, suggesting that it is reasonable to treat this specific group to improve their prognosis.

Several factors (Table 5) are associated with early stage hypertension, including age, BMI, waist circumference (WC), Hb, CRP, etc. The correlation between hypertension and

age, BMI, WC, or WHR is well known. The Hb level is also a positive contributor to blood pressure elevation and was one of the top three important risk factors in all four groups. In a large cohort study, Atsma et al. reported that systolic blood pressure increased by 0.7 mm Hg for every 0.9 mm Hg per millimole per liter increase in the hemoglobin level, and the results for diastolic blood pressure were comparable [80]. There was no gender difference in the study. Several mechanisms for the association between hemoglobin and blood pressure have been proposed. Hemoglobin is positively associated with pulse-wave velocity, an indicator of arterial stiffness, and increased systolic and diastolic blood pressure [81]. Nitroxide (NO), produced in the blood vessel endothelial cells, relaxes vascular smooth muscles, and thereby controls blood pressure. Acellular Hb may bind to NO and cause vessel constriction and elevation of blood pressure [82]. Increased Hb levels may lead to increased blood viscosity, and increased blood viscosity may worsen cardiovascular function, but the production of NO may also increase. In this study, blood viscosity was not measured; therefore, we do not know the influence of viscosity on blood pressure.

C-reactive protein is a biomarker of systemic inflammation. In hypertensive individuals, CRP levels are associated with cardiovascular events and end-organ damage because CRP is correlated with vascular stiffness and severity of atherosclerosis [83]. CRP appeared in the out-of-reference-range LDL-C groups (G2 and G4) and was one of the top two important variables in the ORR-HDL & ORR-LDL group. However, in normotensive individuals, genetic variability may influence circulating levels of CRP. A predictive association between changes in blood pressure and the development of hypertension remains controversial [83,84]. In this study, a residual contribution to blood pressure elevation is found in groups with increased LDL-C, which suggests that LDL-C levels are associated with CPR levels and that the LDL-C level may be a more important factor for predicting the development of hypertension.

Raised blood pressure is the leading cause of death globally [8]. The association between CVD and early stage hypertension is documented in several studies [9–11]. An individual with hypertension or dyslipidemia is predicted to be at lower risk for CVD than one with both of them. Predicting the probability of hypertension in dyslipidemia individuals with normal blood pressure is an important clinical issue. Because non-pharmacological methods, such as body weight control, aerobic exercise, salt restriction, and the DASH diet, are recommended to effectively prevent the development of hypertension. The machine learning model provided the possibility for early detection of the individual with early stage hypertension. In order to prevent future CVD, it would be valuable to suggest they modify their lifestyle aggressively. In addition, several other cofactors of early stage hypertension are also found in the ML model. It suggested that correcting those factors may be important for the development of hypertension. Meanwhile, the application of the ML model could be another method to establish a new direction for future studies to detect early stage hypertension. The implications of the model synthesized to clinical should be helpful and predictable in the public health practice.

The correlation between metabolic syndrome-related variables, including age, BMI, WC, WHR, fasting plasma glucose, and hypertension, is well known. The hemoglobin level is also a positive contributor to blood pressure elevation and it was one of the top three important factors in all four LDL-C/HDL-C groups in this study; therefore, it may be an important variable that affects blood pressure in the early stage of hypertension. A residual contribution to blood pressure elevation is found in groups with increased LDL-C. This suggests that LDL-C levels are associated with CPR levels, and that the LDL-C level may be a more important factor for predicting the development of hypertension. Even though this project is not a longitudinal study design, it may have confounding effects, but the cause of the huge amount of data, and the results were in line with clinical manifestations, so it still had application value in preventive medicine. Using directed acyclic graph to discuss the logical connective or confounding effects of the identified important risk factors is worth of further research.

## 5. Limitations

The main limitations of this study were the use of a single data set without comparing it to data from other countries and the lack of continuity of data analysis. In addition, this study was similar to previous studies that used cross-sectional data in that we estimated the influencing factors and speculated on the possible effects without providing a causal inference. To avoid the selected variables having logical connectors or confounding effects, the following research should be used the prospective or retrospective cohort study to prove the clinical significance. Another limitation of our study is that our inferences may not be suitable for outpatients or inpatients who are already ill. In addition, older patients, those with abnormal extreme values, or those using a physician's prescription for an extended period were excluded from this study.

## 6. Conclusions

The five prediction models (SGB, MARS, Lasso, Ridge, and CatBoost) provided a similar classification of risk factors in this study. Based on the results of this study, we suggest that BMI, WHR, Hb, and CRP should be the important indicators of early stage hypertension in sub-healthy adults. A rise in the LDL-C level appears to be a signal and is more important than the start of a decrease in HDL-C. Raising awareness of hypertension is crucial in government health promotion activities, and the findings of this study should be of value for further discussions and follow-up research.

**Author Contributions:** Made an equal contribution to the study, P.-C.L. and M.-S.C.; provided the study idea and resources, P.-C.L. and M.-S.C.; the corresponding author, conceived and designed the experiments, C.-J.L.; performed the experiments, M.-J.J., M.-S.C. and C.-J.L.; analyzed the data, C.-J.L., C.-T.Y., T.-C.C. and M.-J.J.; were supervisors, C.-J.L. and M.-S.C.; were project administrators, P.-C.L., M.-S.C. and C.-J.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Far Eastern Memorial Hospital (NSC-RD-110-1-10-503 & NSC-RD-110-1-10-510) and the Ministry of Science and Technology (MOST-110-2221-E-030-010 & MOST-110-2622-H-161-001). This study did not receive a specific grant from any funding agency in the public, commercial, or not-for-profit sector.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki and was approved by the Research Ethics Review Committee of Far Eastern Memorial Hospital (No: IRB-110027-E10-43) and ClinicalTrials.gov (ID: NCT05225454).

**Informed Consent Statement:** Informed consent was not required.

**Data Availability Statement:** Authorization is required for the use of all data sets collected from the MJ Health Research Foundation. The application procedures are accessed via this link. <http://www.mjhrf.org/main/page/release1/en/#release01> (accessed on 18 April 2022).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Kannel, W.B.; Gordon, T.; Schwartz, M.J. Systolic versus diastolic blood pressure and risk of coronary heart disease: The Framingham study. *Am. J. Cardiol.* **1971**, *27*, 335–346. [[CrossRef](#)]
2. Ettehad, D.; Emdin, C.A.; Kiran, A.; Anderson, S.G.; Callender, T.; Emberson, J.; Chalmers, J.; Rodgers, A.; Rahimi, K. Blood pressure lowering for prevention of cardiovascular disease and death: A systematic review and meta-analysis. *Lancet* **2016**, *387*, 957–967. [[CrossRef](#)]
3. Neaton, J.D. Serum cholesterol, blood pressure, cigarette smoking, and death from coronary heart disease. Overall findings and differences by age for 316,099 white men. Multiple Risk Factor Intervention Trial Research Group. *Arch. Intern. Med.* **1992**, *152*, 56–64. [[CrossRef](#)] [[PubMed](#)]
4. Knuuti, J.; Wijns, W.; Saraste, A.; Capodanno, D.; Barbato, E.; Funck-Brentano, C.; Prescott, E.; Storey, R.F.; Deaton, C.; Cuisset, T.; et al. 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes. *Eur. Heart J.* **2020**, *41*, 407–477. [[CrossRef](#)]
5. Members, W.C.; Lawton, J.S.; Tamis-Holland, J.E.; Bangalore, S.; Bates, E.R.; Beckie, T.M.; Zwischenberger, B.A. 2021 ACC/AHA/SCAI guideline for coronary artery revascularization: A report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *J. Am. Coll. Cardiol.* **2022**, *79*, e21–e129. [[CrossRef](#)]

6. Kleindorfer, D.O.; Towfighi, A.; Chaturvedi, S.; Cockcroft, K.M.; Gutierrez, J.; Lombardi-Hill, D.; Kamel, H.; Kernan, W.N.; Kittner, S.J.; Leira, E.C.; et al. Guideline for the Prevention of Stroke in Patients with Stroke and Transient Ischemic Attack: A Guideline from the American Heart Association/American Stroke Association. *Stroke* **2021**, *52*, e364–e467. [[CrossRef](#)]
7. Unger, T.; Borghi, C.; Charchar, F.; Khan, N.A.; Poulter, N.R.; Prabhakaran, D.; Ramirez, A.; Schlaich, M.; Stergiou, G.S.; Tomaszewski, M.; et al. 2020 International Society of Hypertension Global Hypertension Practice Guidelines. *Hypertension* **2020**, *75*, 1334–1357. [[CrossRef](#)]
8. GBD 2016 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *Lancet* **2017**, *390*, 1345–1422. [[CrossRef](#)]
9. Ferguson, T.S.; Younger, N.O.; Tulloch-Reid, M.K.; Wright, M.B.; Ward, E.M.; Ashley, D.E.; Wilks, R.J. Prevalence of prehypertension and its relationship to risk factors for cardiovascular disease in Jamaica: Analysis from a cross-sectional survey. *BMC Cardiovasc. Disord.* **2008**, *8*, 20. [[CrossRef](#)]
10. Grotto, I.; Grossman, E.; Huerta, M.; Sharabi, Y. Prevalence of prehypertension and associated cardiovascular risk profiles among young Israeli adults. *Hypertension* **2006**, *48*, 254–259. [[CrossRef](#)]
11. Yu, D.H.; Huang, J.F.; Hu, D.S.; Chen, J.C.; Cao, J.; Li, J.X.; Gu, D.F. Association between prehypertension and clustering of cardiovascular disease risk factors among Chinese adults. *J. Cardiovasc. Pharm.* **2009**, *53*, 388–400. [[CrossRef](#)]
12. Bønaa, K.H.; Thelle, D.S. Association between blood pressure and serum lipids in a population. The Tromsø Study. *Circulation* **1991**, *83*, 1305–1314. [[CrossRef](#)]
13. Cicero, A.F.; D’Addato, S.; Veronesi, M.; Rosticci, M.; Santi, F.; Dormi, A.; Borghi, C.; Brisighella Heart Study Group. Relationship between blood pressure, cholesterolemia and serum apolipoprotein B in a large population sample: The Brisighella Heart Study. *J. Hypertens.* **2012**, *30*, 492–496. [[CrossRef](#)]
14. Chobanian, A.V.; Bakris, G.L.; Black, H.R.; Cushman, W.C.; Green, L.A.; Izzo, L., Jr.; Materson, B.J.; Oparil, S.; Wright, J.T., Jr.; Roccella, E.J. Seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure. *Hypertension* **2003**, *42*, 1206–1252. [[CrossRef](#)]
15. Briasoulis, A.; Agarwal, V.; Valachis, A.; Messerli, F.H. Antihypertensive effects of statins: A meta-analysis of prospective controlled studies. *J. Clin. Hypertens.* **2013**, *15*, 310–320. [[CrossRef](#)]
16. Borghi, C.; Cicero, A.F.; Saragoni, S.; Buda, S.; Cristofori, C.; Lilli, P.; Degli Esposti, L. Rate of control of LDL cholesterol and incident hypertension requiring antihypertensive treatment in hypercholesterolemic subjects in daily clinical practice. *Ann. Med.* **2014**, *46*, 97–102. [[CrossRef](#)]
17. Chen, C.-L.; Liu, X.-C.; Liu, L.; Lo, K.; Yu, Y.-L.; Huang, J.-Y.; Huang, Y.-Q.; Chen, J.-Y. U-Shaped Association of High-Density Lipoprotein Cholesterol with All-Cause and Cardiovascular Mortality in Hypertensive Population. *Risk Manag. Healthc. Policy* **2020**, *13*, 2013–2025. [[CrossRef](#)]
18. Ko, D.T.; Alter, D.A.; Guo, H.; Koh, M.; Lau, G.; Austin, P.C.; Booth, G.L.; Hogg, W.; Jackevicius, C.A.; Lee, D.S.; et al. High-density lipoprotein cholesterol and cause-specific mortality in individuals without previous cardiovascular conditions: The CANHEART study. *J. Am. Coll. Cardiol.* **2016**, *68*, 2073–2083. [[CrossRef](#)]
19. Hirata, A.; Sugiyama, D.; Watanabe, M.; Kotani, K.; Ueshima, H.; Imai, Y.; Ohkubo, T.; Irie, F.; Iso, H.; Kitamura, A.; et al. Association of extremely high levels of high-density lipoprotein cholesterol with cardiovascular mortality in a pooled analysis of 9 cohort studies including 43,407 individuals: The EPOCH-JAPAN study. *J. Clin. Lipidol.* **2018**, *12*, 674–684. [[CrossRef](#)]
20. DeGregory, K.W.; Kuiper, P.; DeSilvio, T.; Pleuss, J.D.; Miller, R.; Roginski, J.W.; Fisher, C.B.; Harness, D.; Viswanath, S.; Heymsfield, S.B.; et al. A review of machine learning in obesity. *Obes. Rev.* **2018**, *19*, 668–685. [[CrossRef](#)]
21. Liu, Y.; Chen, P.-H.C.; Krause, J.; Peng, L. How to Read Articles That Use Machine Learning: Users’ Guides to the Medical Literature. *JAMA* **2019**, *322*, 1806–1816. [[CrossRef](#)] [[PubMed](#)]
22. Triantafyllidis, A.K.; Tsanas, A. Applications of Machine Learning in Real-Life Digital Health Interventions: Review of the Literature. *J. Med. Internet Res.* **2019**, *21*, e12286. [[CrossRef](#)] [[PubMed](#)]
23. Peiffer-Smadja, N.; Rawson, T.M.; Ahmad, R.; Buchard, A.; Georgiou, P.; Lescure, F.-X.; Birgand, G.; Holmes, A.H. Machine learning for clinical decision support in infectious diseases: A narrative review of current applications. *Clin. Microbiol. Infect.* **2020**, *26*, 584–595. [[CrossRef](#)] [[PubMed](#)]
24. Chiu, Y.-L.; Jhou, M.-J.; Lee, T.-S.; Lu, C.-J.; Chen, M.-S. Health Data-Driven Machine Learning Algorithms Applied to Risk Indicators Assessment for Chronic Kidney Disease. *Risk Manag. Healthc. Policy* **2021**, *14*, 4401–4412. [[CrossRef](#)] [[PubMed](#)]
25. Chang, C.-C.; Yeh, J.-H.; Chiu, H.-C.; Chen, Y.-M.; Jhou, M.-J.; Liu, T.-C.; Lu, C.-J. Utilization of Decision Tree Algorithms for Supporting the Prediction of Intensive Care Unit Admission of Myasthenia Gravis: A Machine Learning-Based Approach. *J. Pers. Med.* **2022**, *12*, 32. [[CrossRef](#)] [[PubMed](#)]
26. Huang, Y.-C.; Cheng, Y.-C.; Jhou, M.-J.; Chen, M.; Lu, C.-J. Important Risk Factors in Patients with Nonvalvular Atrial Fibrillation Taking Dabigatran Using Integrated Machine Learning Scheme—A Post Hoc Analysis. *J. Pers. Med.* **2022**, *12*, 756. [[CrossRef](#)]
27. Friedman, J. Stochastic gradient boosting. *Comput. Stat. Data Anal.* **2002**, *38*, 367–378. [[CrossRef](#)]
28. Friedman, J.H. Multivariate adaptive regression splines. *Ann. Stat.* **1991**, *19*, 1–67. [[CrossRef](#)]
29. Guindo, M.L.; Kabir, M.H.; Chen, R.; Liu, F. Particle Swarm Optimization and Multiple Stacked Generalizations to Detect Nitrogen and Organic-Matter in Organic-Fertilizer Using Vis-NIR. *Sensors* **2021**, *21*, 4882. [[CrossRef](#)]

30. Dorogush, A.V.; Ershov, V.; Gulin, A. CatBoost: Gradient boosting with categorical features support. *arXiv* **2018**, arXiv:1810.11363. [[CrossRef](#)]
31. Gu, D.; Su, S.; Ge, D.; Chen, S.; Huang, J.; Li, B.; Chen, R.; Qiang, B. Association study with 33 single-nucleotide polymorphisms in 11 candidate genes for hypertension in Chinese. *Hypertension* **2006**, *47*, 1147–1154. [[CrossRef](#)]
32. Ge, D.; Zhu, H.; Huang, Y.; Treiber, F.A.; Harshfield, G.A.; Snieder, H.; Dong, Y. Multilocus analyses of Renin-Angiotensin-aldosterone system gene variants on blood pressure at rest and during behavioral stress in young normotensive subjects. *Hypertension* **2007**, *49*, 107–112. [[CrossRef](#)]
33. Hsu, C.-C.; Gu, M.; Lee, T.-S.; Lu, C.-J. The Effects of Daily Sleep Condition on Performances of Physical Fitness among Taiwanese Adults: A Cross-Sectional Study. *Int. J. Environ. Res. Public Health* **2020**, *17*, 1907. [[CrossRef](#)]
34. Ong, M.S.; Klann, J.G.; Lin, K.J.; Maron, B.A.; Murphy, S.N.; Natter, M.D.; Mandl, K.D. Claims-Based Algorithms for Identifying Patients with Pulmonary Hypertension: A Comparison of Decision Rules and Machine-Learning Approaches. *J. Am. Heart. Assoc.* **2020**, *9*, e016648. [[CrossRef](#)]
35. Chang, C.-C.; Yeh, J.-H.; Chen, Y.-M.; Jhou, M.-J.; Lu, C.-J. Clinical Predictors of Prolonged Hospital Stay in Patients with Myasthenia Gravis: A Study Using Machine Learning Algorithms. *J. Clin. Med.* **2021**, *10*, 4393. [[CrossRef](#)]
36. Liu, X.; Zhang, Y.; Fu, C.; Zhang, R.; Zhou, F. EnRank: An Ensemble Method to Detect Pulmonary Hypertension Biomarkers Based on Feature Selection and Machine Learning Models. *Front. Genet.* **2021**, *12*, 636429. [[CrossRef](#)]
37. Zhao, H.; Zhang, X.; Xu, Y.; Gao, L.; Ma, Z.; Sun, Y.; Wang, W. Predicting the Risk of Hypertension Based on Several Easy-to-Collect Risk Factors: A Machine Learning Method. *Front. Public Health* **2021**, *9*, 619429. [[CrossRef](#)]
38. Lee, H.J.; Kim, H.; Kim, Y.T.; Won, K.; Czosnyka, M.; Kim, D.J. Prediction of Life-Threatening Intracranial Hypertension During the Acute Phase of Traumatic Brain Injury Using Machine Learning. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 3967–3976. [[CrossRef](#)]
39. Islam, S.; Talukder, A.; Awal, M.A.; Siddiqui, M.; Ahamad, M.M.; Ahammed, B.; Rawal, L.B.; 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. [[CrossRef](#)]
40. Hernandez-Pacheco, J.A.; Torres-Torres, J.; Martinez-Portilla, R.J.; Solis-Paredes, J.M.; Estrada-Gutierrez, G.; Mateu-Rogell, P.; Nares-Torices, M.A.; Lopez-Marenco, M.E.; Escobedo-Segura, K.R.; Posadas-Nava, A.; et al. sFlt-1 Is an Independent Predictor of Adverse Maternal Outcomes in Women With SARS-CoV-2 Infection and Hypertensive Disorders of Pregnancy. *Front. Med.* **2022**, *9*, 894633. [[CrossRef](#)]
41. Chang, C.D.; Wang, C.C.; Jiang, B.C. Using data mining techniques for multi-diseases prediction modeling of hypertension and hyperlipidemia by common risk factors. *Expert Syst. Appl.* **2011**, *38*, 5507–5513. [[CrossRef](#)]
42. Ang, L.; Yim, M.H.; Do, J.-H.; Lee, S. A Novel Method in Predicting Hypertension Using Facial Images. *Appl. Sci.* **2021**, *11*, 2414. [[CrossRef](#)]
43. Shan, Y.; Li, Y.; Xu, X.; Feng, J.; Wu, X.; Gao, G. Evaluation of Intracranial Hypertension in Traumatic Brain Injury Patient: A Noninvasive Approach Based on Cranial Computed Tomography Features. *J. Clin. Med.* **2021**, *10*, 2524. [[CrossRef](#)]
44. Chai, S.S.; Goh, K.L.; Cheah, W.L.; Chang, Y.H.R.; Ng, G.W. Hypertension Prediction in Adolescents Using Anthropometric Measurements: Do Machine Learning Models Perform Equally Well? *Appl. Sci.* **2022**, *12*, 1600. [[CrossRef](#)]
45. Chen, M.S.; Chen, S.H. A data-driven assessment of the metabolic syndrome criteria for adult health management in Taiwan. *Int. J. Environ. Res. Public Health* **2019**, *16*, 92. [[CrossRef](#)]
46. Liao, C.-M.; Lin, C.-M. Life course effects of socioeconomic and lifestyle factors on metabolic syndrome and 10-year risk of cardiovascular disease: A longitudinal study in taiwan adults. *Int. J. Environ. Res. Public Health* **2018**, *15*, 2178. [[CrossRef](#)]
47. Chen, M.S.; Chiu, C.H.; Chen, S.H. Risk assessment of metabolic syndrome prevalence involving sedentary occupations and socioeconomic status. *BMJ Open* **2021**, *11*, e042802. [[CrossRef](#)]
48. Chan, T.C.; Zhang, Z.; Lin, B.C.; Lin, C.; Deng, H.B.; Chuang, Y.C.; Chan, J.W.M.; Jiang, W.K.; Tam, T.; Chang, L.Y.; et al. Long-Term Exposure to Ambient Fine Particulate Matter and Chronic Kidney Disease: A Cohort Study. *Environ. Health Perspect.* **2018**, *126*, 107002. [[CrossRef](#)]
49. Friedman, J. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [[CrossRef](#)]
50. Hastie, T.; Tibshirani, R.; Wainwright, M. *Statistical Learning with Sparsity: The Lasso and Generalizations*; CRC Press: Boca Raton, FL, USA, 2015. [[CrossRef](#)]
51. Kwon, S.; Lee, S.; Na, O. Tuning parameter selection for the adaptive Lasso in the autoregressive model. *J. Korean Stat. Soc.* **2017**, *46*, 285–297. [[CrossRef](#)]
52. Hoerl, A.E.; Kennard, R.W. Ridge regression: Biased estimation for non-orthogonal problems. *Technometrics* **1970**, *12*, 55–67. [[CrossRef](#)]
53. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. Catboost: Unbiased boosting with categorical features. In *Advances in Neural Information Processing Systems 31, Proceedings of the 32nd International Conference on Neural Information Processing Systems, Montréal, QC, Canada, 3–8 December 2018*; Curran Associates Inc.: Red Hook, NY, USA, 2018; pp. 6639–6649.
54. Meng, J.; Hao, H.; Luan, Y. Classifier ensemble selection based on affinity propagation clustering. *J. Biomed. Inform.* **2016**, *60*, 234–242. [[CrossRef](#)] [[PubMed](#)]
55. Tripathi, D.; Edla, D.R.; Cheruku, R.; Kuppili, V. A novel hybrid credit scoring model based on ensemble feature selection and multilayer ensemble classification. *Comput. Intell.* **2019**, *35*, 371–394. [[CrossRef](#)]

56. Woloshuk, A.; Khochare, S.; Almulhim, A.F.; McNutt, A.T.; Dean, D.; Barwinska, D.; Ferkowicz, M.J.; Eadon, M.T.; Kelly, K.J.; Dunn, K.W.; et al. In Situ Classification of Cell Types in Human Kidney Tissue Using 3D Nuclear Staining. *Cytom. Part A* **2021**, *99*, 707–721. [[CrossRef](#)] [[PubMed](#)]
57. Larner, A.J. Accuracy of cognitive screening instruments reconsidered: Overall, balanced or unbiased accuracy? *Neurodegener. Dis. Manag.* **2022**, *12*, 67–76. [[CrossRef](#)] [[PubMed](#)]
58. Greenwell, B.; Boehmke, B.; Cunningham, J. Gbm: Generalized Boosted Regression Models. R Package Version, 2.1.8. 2020. Available online: <https://CRAN.R-project.org/package=gbm> (accessed on 25 May 2022).
59. Milborrow, S. Derived from Mda: MARS by T. Hastie and R. Tibshirani. Earth: Multivariate Adaptive Regression Splines. R Package Version, 5.3.1. 2021. Available online: <http://CRAN.R-project.org/package=earth> (accessed on 25 May 2022).
60. Friedman, J.; Hastie, T.; Tibshirani, R.; Narasimhan, B.; Tay, K.; Simon, N.; Qian, J.; Yang, J. Glmnet: Lasso and Elastic-Net Regularized Generalized Linear Models. R Package Version, 4.1-4. 2022. Available online: <https://CRAN.R-project.org/package=glmnet> (accessed on 25 May 2022).
61. Yandex Technologies. CatBoost: Unbiased Boosting with Categorical Features. R Package Version, 1.0.6. 2022. Available online: <https://github.com/catboost/catboost/> (accessed on 25 May 2022).
62. Kuhn, M. Caret: Classification and Regression Training. R Package Version, 6.0-92. 2022. Available online: <https://CRAN.R-project.org/package=caret> (accessed on 25 May 2022).
63. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* **1988**, *44*, 837–845. [[CrossRef](#)] [[PubMed](#)]
64. Mainous, A.G., III; Everett, C.J.; Liszka, H.; King, D.E.; Egan, B.M. Prehypertension and mortality in a nationally representative cohort. *Am. J. Cardiol.* **2004**, *94*, 1496–1500. [[CrossRef](#)]
65. Hsia, J.; Margolis, K.L.; Eaton, C.B.; Wenger, N.K.; Allison, M.; Wu, L.; LaCroix, A.Z.; Black, H.R.; Women’s Health Initiative Investigators. Prehypertension and cardiovascular disease risk in the Women’s Health Initiative. *Circulation* **2007**, *115*, 855–860. [[CrossRef](#)]
66. Gu, D.; Chen, J.; Wu, X.; Duan, X.; Jones, D.W.; Huang, J.F.; Chen, C.S.; Chen, J.C.; Kelly, T.N.; Whelton, P.K.; et al. Prehypertension and risk of cardiovascular disease in Chinese adults. *J. Hypertens.* **2009**, *27*, 721–729. [[CrossRef](#)]
67. Hozawa, A.; Kuriyama, S.; Kakizaki, M.; Ohmori-Matsuda, K.; Ohkubo, T.; Tsuji, I. Attributable risk fraction of prehypertension on cardiovascular disease mortality in the Japanese population: The Ohsaki Study. *Am. J. Hypertens.* **2009**, *22*, 267–272. [[CrossRef](#)]
68. Tsai, P.S.; Ke, T.L.; Huang, C.J.; Tsai, J.C.; Chen, P.L.; Wang, S.Y.; Shyu, Y.K. Prevalence and determinants of prehypertension status in the Taiwanese general population. *J. Hypertens.* **2005**, *23*, 1355–1360. [[CrossRef](#)]
69. Kim, N.H.; Cho, H.J.; Kim, Y.J.; Cho, M.J.; Choi, H.Y.; Eun, C.R.; Kim, J.H.; Yang, S.J.; Yoo, H.J.; Kim, H.Y.; et al. Combined effect of high-normal blood pressure and low HDL cholesterol on mortality in an elderly Korean population: The South-West Seoul (SWS) study. *Am. J. Hypertens.* **2011**, *24*, 918–923. [[CrossRef](#)]
70. Nakajima, K.; Igata, M.; Higuchi, R.; Tanaka, K.; Mizusawa, K.; Nakamura, T. Association of Serum High-Density Lipoprotein Cholesterol with High Blood Pressures at Checkup: Results of Kanagawa Investigation of Total Checkup Data from the National Database-9 (KITCHEN-9). *J. Clin. Med.* **2021**, *10*, 5118. [[CrossRef](#)]
71. Otocka-Kmieciak, A.; Mikhailidis, D.P.; Nicholls, S.J.; Davidson, M.; Rysz, J.; Banach, M. Dysfunctional HDL: A novel important diagnostic and therapeutic target in cardiovascular disease? *Prog. Lipid Res.* **2012**, *51*, 314–324. [[CrossRef](#)]
72. Barter, P.J.; Caulfield, M.; Eriksson, M.; Grundy, S.M.; Kastelein, J.J.P.; Komajda, M.; Lopez-Sendon, J.; Mosca, L.; Tardif, J.-C.; Waters, D.D.; et al. Effects of torcetrapib in patients at high risk for coronary events. *N. Engl. J. Med.* **2007**, *357*, 2109–2122. [[CrossRef](#)]
73. Lincoff, A.M.; Nicholls, S.J.; Riesmeyer, J.S.; Barter, P.J.; Brewer, H.B.; Fox, K.A.A.; Gibson, C.M.; Granger, C.; Menon, V.; Montalescot, G.; et al. Evacetrapib and Cardiovascular Outcomes in High-Risk Vascular Disease. *N. Engl. J. Med.* **2017**, *376*, 1933–1942. [[CrossRef](#)]
74. Chruściel, P.; Stemplewska, P.; Stemplewski, A.; Wattad, M.; Bielecka-Dąbrowa, A.; Maciejewski, M.; Penson, P.; Bartłomiejczyk, M.A.; Banach, M. Associations between the lipid profile and the development of hypertension in young individuals—the preliminary study. *Arch. Med. Sci.* **2019**, *18*, 25–35. [[CrossRef](#)]
75. Otsuka, T.; Takada, H.; Nishiyama, Y.; Kodani, E.; Saiki, Y.; Kato, K.; Kawada, T. Dyslipidemia and the Risk of Developing Hypertension in a Working-Age Male Population. *J. Am. Heart Assoc.* **2016**, *5*, e003053. [[CrossRef](#)]
76. Wen, C.J.; Lee, Y.S.; Lin, W.Y.; Huang, H.L.; Yao, C.A.; Sung, P.K.; Huang, K.C. The metabolic syndrome increases cardiovascular mortality in Taiwanese elderly. *Eur. J. Clin. Investig.* **2008**, *38*, 469–475. [[CrossRef](#)]
77. Shin, C.Y.; Yun, K.E.; Park, H.S. Blood pressure has a greater impact on cardiovascular mortality than other components of metabolic syndrome in Koreans. *Atherosclerosis* **2009**, *205*, 614–619. [[CrossRef](#)]
78. ACCORD Study Group; Cushman, W.C.; Evans, G.W.; Byington, R.P.; Goff, D.C.; Grimm, R.H.; Cutler, J.A.; Simons-Morton, D.G.; Basile, J.N.; Corson, M.A.; et al. Effects of intensive blood-pressure control in type 2 diabetes mellitus. *N. Engl. J. Med.* **2010**, *362*, 1575–1585. [[CrossRef](#)] [[PubMed](#)]
79. SPRINT Research Group; Wright, J.T.; Williamson, J.D.; Whelton, P.K.; Snyder, J.K.; Sink, K.M.; Rocco, M.V.; Reboussin, D.M.; Rahman, M.; Oparil, S.; et al. A Randomized Trial of Intensive versus Standard Blood-Pressure Control. *N. Engl. J. Med.* **2015**, *373*, 2103–2116. [[CrossRef](#)] [[PubMed](#)]

80. Atsma, F.; Veldhuizen, I.; de Kort, W.; van Kraaij, M.; Pasker-de Jong, P.; Deinum, J. Hemoglobin level is positively associated with blood pressure in a large cohort of healthy individuals. *Hypertension* **2012**, *60*, 936–941. [[CrossRef](#)] [[PubMed](#)]
81. Kawamoto, R.; Tabara, Y.; Kohara, K.; Miki, T.; Kusunoki, T.; Katoh, T.; Ohtsuka, N.; Takayama, S.; Abe, M. A slightly low hemoglobin level is beneficially associated with arterial stiffness in Japanese community-dwelling women. *Clin. Exp. Hypertens.* **2012**, *34*, 92–98. [[CrossRef](#)]
82. Cabrales, P.; Han, G.; Nacharaju, P.; Friedman, A.J.; Friedman, J.M. Reversal of hemoglobin-induced vasoconstriction with sustained release of nitric oxide. *Am. J. Physiol.-Heart Circ. Physiol.* **2011**, *300*, H49–H56. [[CrossRef](#)]
83. Hage, F.G. C-reactive protein and hypertension. *J. Hum. Hypertens.* **2014**, *28*, 410–415. [[CrossRef](#)]
84. Davey, S.G.; Lawlor, D.A.; Harbord, R.; Timpson, N.; Rumley, A.; Lowe, G.D.; Day, I.N.; Ebrahim, S. Association of C-reactive protein with blood pressure and hypertension: Life course confounding and mendelian randomization tests of causality. *Arterioscler. Thromb. Vasc. Biol.* **2005**, *25*, 1051–1056. [[CrossRef](#)]