Development of the Hypertension Index Model in General Adult using the Korea National Health and Nutritional Examination Survey and the Korean Genome and Epidemiology Study

Myung Jae Seo

The Graduate School
Yonsei University
Department of Medicine

Development of the Hypertension Index Model in General Adult using the Korea National Health and Nutritional Examination Survey and the Korean Genome and Epidemiology Study

A Dissertation
Submitted to the Department of Medicine
and the Graduate School of Yonsei University
in partial fulfillment of the
requirements for the degree of
Master of Medicine

Myung Jae Seo

December 2020

This certifies that the Dissertation of
Myung Jae Seo is approved.

_____

Thesis Supervisor: Jong Koo Kim

_____

Sung Gyun Ahn

_____

Yong Jae Lee

The Graduate School
Yonsei University
December 2020

Table of Contents

List of Tables

List of Figures

ABSTRACT

Development of the Hypertension Index Model in General Adult using the Korea National Health and Nutritional Examination Survey and the Korean Genome and Epidemiology Study

Myung Jae Seo

Department of Medicine

The Graduate School

Yonsei University

Along with emphasizing awareness and control of hypertension currently, the importance of predicting the incidence of hypertension raises. We constructed the gender-specific prediction model from Korean datasets, including socioeconomic status, medical history, life style-related variables, anthropometric status, and laboratory indices. We utilized the population of the Korea National Health and Nutrition Examination Survey (KNHANES) from 2011 to 2015 for deriving a hypertension prediction model. Participants with 40 years old or more were included, and 17 risk factors of hypertension were available based on literature review. We constructed the gender-specific hypertension classification model, and estimated the performance using the KNHANES dataset from 2016 to 2017 (AUC = 0.798 in men, AUC = 0.849 in women). The performance of our hypertension model was evaluated as significant by the cumulative incidence which was drawn from a longitudinal dataset, the community based Korean Genome and Epidemiology Study (KoGES) dataset. We suggested a hypertension prediction model using features which could be collected in clinical office without difficulty. The individualized result may alert a high-risk person to modify unhealthy life-styles.

Key words : Hypertension

1. INTRODUCTION

Approximately 1.4 billion population in worldwide are estimated to have hypertension in 2010, and more than 1.6 billion population are likely to have hypertension by 2025 [1]. In Korea, the prevalence of hypertension is approximately 30.5% in the 2007 – 15 Korea National Health and Nutrition Examination Survey (KNHANES) [2]. Hypertension is one of the leading risk factors for cardiovascular disease (CVD), chronic kidney disease, and all-cause & CVD-related mortality [3-5].

Major two issues for hypertension are the awareness and well-control of hypertension, which have known to vary according to region, sex, income, and education levels [1]. In worldwide, the awareness and treatment proportion of hypertension patients were less than half, and the hypertension control ratio was only 13.8% globally in 2010 [1]. A study by Kang et al. [2] reported that 32.7 % of patients with hypertension did not have the perception for their status. The early diagnosis or prediction of hypertension is one of the crucial challenges because uncontrolled or untreated hypertension have known to elevate mortality [5].

Two approaches have been proposed to early diagnosis of hypertension. The first approach is to identify socioeconomic or biological risk factors for hypertension. A lot of anthropometric, lifestyle, and laboratory indices have been proposed as the risk factor for hypertension [6]. Several studies have been conducted the intervention for subjects with several risk factors, such as body weight and alcohol consumption, which have significantly lowered blood pressure and reduced the incidence of hypertension [7, 8].

The other approach is the development of prediction model that measures the combinatory effect of risk factors [9-11]. Parikh et al. [9] presented a hypertension prediction model using about 1700 Americans that were enrolled in the Framingham Heart Study and the Framingham Offspring Study cohort of USA. This study identified seven risk factors through the Weibull regression, and made the prediction model using these seven variables as input features. A hypertension prediction model was proposed using

about 6700 UK populations, which used seven risk factors curated from the Weibull regression as input features for the model [10]. A study by Paynter et al. [11] introduced a prediction model using the Women's Health Study (WHS) cohort with about 14800 Americans. This model included nine risk factors; age, ethnicity, systolic blood pressure (BP), diastolic BP, body mass index (BMI) and laboratory indices.

Hypertension has the complicated pathophysiology interrelated to the cardiovascular system, renin-angiotensin system, autonomic nervous system, and other factors [12]. However, previous hypertension prediction models only used only finger-countable risk factors. Thereby, these models could not reflect the complex bio-signature of hypertension. This limitation, only selecting few candidate predictors occurred because the statistical method for feature selection had strict criteria when curating risk factors.

In the present study, we attempted to establish a gender-specific hypertension prediction model including a large number of candidate variables. Furthermore, we adopted weighting values estimated by the KNHANES to increase statistical power of backward-stepwise logistic regression model.

2. METHODS

2.1. Participants

The present study used the 2011 – 17 KNHANES and the community based Korean Genome and Epidemiology Study (KoGES). The KNHANES is the nationwide cross-sectional survey including a health status, chronic disease prevalence, and nutrition status, conducted by the Centers for Disease Control and Prevention in South Korea annually. The KoGES is an ongoing cohort study from 2001 established by the Korean National Institute of Health. It contains a health and lifestyle survey, laboratory results, and chronic disease incidence of Korean adults. This study used the KoGES data up to December 2018.

We included study population over 40 years old. We excluded subjects in both KNHANES and KoGES with incomplete data about socioeconomic status, medical history, life style-related variables, anthropometric, and laboratory indices. Specifically, participants in the KoGES were excluded when they were diagnosed as having hypertension or had high BP (systolic BP $\geq$ 140 or diastolic BP $\geq$ 90) at baseline. All participants were provided written informed consent prior to their participation in these surveys, and were processed anonymously. The KNHANES and KoGES were in the compliance with Institutional Review Board of Wonju Severance Christian Hospital (WSCH) with approval no. CR320310.

We arranged 2011 – 15 KNHANES and 2016 – 17 KNHANES as the derivation and the internal validation (IV) sets, respectively. The KoGES dataset was used as the external validation (EV) set. The final subjects to be analyzed were 15,395, 6,787, and 5,152 in the derivation, IV, and EV sets, respectively.

2.2. Definition of hypertension

Subjects in the KNHANES were categorized as hypertension group if previously diagnosed as having hypertension by a physician or received anti-hypertension medications due to the cross-sectional design. In the case of the KoGES, hypertension was defined according to

JNC 7: (1) systolic blood pressure of 140 mmHg or more, or diastolic blood pressure of 90 mmHg or more [13]; (2) previous diagnosis of hypertension by a physician; and (3) receiving anti-hypertension medications.

2.3. Predictor selection

We determined predictive risk factors by two methods, literature-based search and statistics-based selection as previous study performed [14]. First, two physicians who had experience for treating hypertension performed literature-based search to select candidate risk factors. As a result, approximately 40 risk factors were reviewed from the literature-based search, and detail information about individual studies is described in the Supplementary Table S1. Among 40 risk factors, we contained 17 variables, available in both KNHANES and KoGES.

We selected predictive candidates by backward-stepwise logistic regression (LR) after applying the weight values. The weight value is a number that indicates how many people a subject represents, and determined by a data constructor at the time that the data was generated. Note that the weight value was only used for the process of selecting the significant variables and constructing prediction model.

2.4. Variables measurement

We included education status classified by elementary, middle, high school and university graduation and income status categorized in quartile.

Diabetes mellitus was defined as follows: (1) serum fasting glucose level $\geq$ 126 mg/dL or (2) previously diagnosed by a physician or (3) receiving glucose-lowering medications [15]. Dyslipidemia was included with those receiving lipid-lowering medications and cancer was confirmed by previous diagnosis of a physician.

We defined heavy alcoholics when a subject drank over 70g in women and over 140g in men per week. Nutritional information such as daily energy intake was investigated via 24hr recall.

Blood pressure (BP) was measured three times by Baumanometer (WA Baum Co., Inc., Copiague, NY) in each subject, after resting for 30 minutes in a sitting position. The average value of the three times of measurement was selected for BP. The body mass index (BMI) was calculated as weight/height^2 (kg/m^2). Blood samples were collected after fasting for eight hours at least. These samples were refrigerated instantly, and analyzed in 24 hours. The levels of serum fasting glucose and lipid profiles were measured by the kinetic Jaffe method using a model 7600 auto-analyzer (Hitachi, Tokyo, Japan). The level of white blood cell (WBC) and hemoglobin (Hb) were measure with an XE-2100D (Sysmex, Tokyo, Japan).

2.5. Statistics

Distributions of candidate variables were determined to be normal via the Kolmogorov–Smirnov test. The continuous and categorical variables were analyzed based on hypertension status by means of student's t-test and Chi square test, respectively. LR was used to select risk factors for hypertension. The hypertension classification model was constructed from the training dataset using a generalized linear model (GLM). To validate performance of the prediction model, we used the area under the receiver operating characteristic curve (AUC) and F-measure. Cox regression was used for survival analysis. Statistical analysis was performed by R language (R packages ver.3.6.1). P-value < 0.05 was considered statistically significant.

3. Results

Table 1 and 2 present the general characteristics of the derivation and IV sets based on absence or presence of hypertension, respectively. The mean age of participants in the hypertension group was older than those of the non-hypertension group in both the derivation and IV datasets (Table 1, 2). The ratio of female did not significantly differ between non-hypertension and hypertension groups in the derivation set, besides those of hypertension group significantly differed in the IV set. Subjects with hypertension in both derivation and IV sets had more following diseases than those without hypertension: diabetes mellitus, dyslipidemia, and cancer.

The average levels of alcohol consumption, BMI, waist circumference, systolic BP, diastolic BP, serum fating glucose, triglyceride, and WBC in the hypertension group were significantly higher than those of the non-hypertension group in both datasets, except alcohol consumption in the development dataset and diastolic BP in the IV set. Inversely, the average levels of total energy intake and total cholesterol in the hypertension group were significantly lower than those in non-hypertension group in both datasets, besides Hb level were shown as insignificant result in IV set.

By means of the univariate LR (Table 3 and 4, Model 1), 17 features were selected as the candidate variables for the Model 2 in men and women, respectively. The Model 2 (multivariate LR) proposed 16 and 17 significant variables in men and women, respectively. These variables showed the significant association with the prevalence of hypertension in Model 3 (Table 3 and 4). Finally, we set the 16 and 17 risk factors for the hypertension as the input features for the classifying model in men and women, respectively (Appendices Table 1 and 2). We constructed the gender-specific hypertension classifying model by using the selected variables and LR.

We measured the probabilities that subjects in the IV set had hypertension by using the gender-specific model. Figure 1A and 1B present the receiver operating characteristic (ROC) curve, in men and women, respectively. The area under the receiver operating curve

(AUC) from the IV set was 0.798 in men and 0.849 in women, respectively. Furthermore, the cut-off value of our model was 0.268 (26.8%) and 0.373 (37.3%) in men and women, respectively, which presented the maximum performance in the f-measure.

The baseline characteristics of EV set of KoGES data is presented in Table 5. Most of differential characteristics of socioeconomic status, life style-related, anthropometric, and laboratory variables between normal and hypertension groups in the EV set were similar with those in the IV set. However, this table showed another results. Income status and diastolic BP were significant in EV set, and their basic characteristics presented similar to derivation set. In EV set, medical history of dyslipidemia and cancer, and total energy intake were not significantly differentiated between two groups in EV set. Furthermore, the serum levels of total cholesterol and Hb were higher in hypertension group than those in non-hypertension group in EV set, significantly.

We measured the probabilities that subjects in the KoGES (EV set) had hypertension by using the gender-specific model that was constructed from the derivation set (KNHANES). We divided subjects into five groups based on the ascending order of the probabilities. Figure 2 illustrates cumulative incidence of new-onset hypertension, and most groups had significant difference with other groups, except group 2 vs 3, 3 vs 4 in men and between 1 vs 2 in women.

Table 1. General characteristics of derivation dataset (KNHANES 2011 – 15).

	Derivation dataset (KNHANES 2011 – 15) n=15395		
	Non-hypertension	Hypertension	
	n = 10434	n = 4961	P-value
Age, years	56.4 (±0.1)	65.7 (±0.13)	<0.001
Sex (female), n	6146 (58.9)	2877 (58.0)	0.291
Income, n			0.005
1st quartile	2456 (23.5)	1204 (24.3)	
2nd quartile	2602 (24.9)	1336 (26.9)	
3rd quartile	2632 (25.2)	1227 (24.7)	
4th quartile	2744 (26.3)	1194 (24.1)	
Education, n			<0.001
Elementary school	2824 (27.1)	2482 (50.0)	
Middle school	1523 (14.6)	764 (15.4)	
High school	3519 (33.7)	1150 (23.2)	
University	2568 (24.6)	565 (11.4)	
Diabetes mellitus, n	1002 (9.6)	1391 (28.0)	<0.001
Dyslipidemia, n	593 (5.7)	1207 (24.3)	<0.001
Cancer, n	522 (5.0)	312 (6.3)	0.001
Alcohol consumption, g/week	56.0 (±1.2)	52.2 (±1.73)	0.068
Total energy intake, kcal	1976.8 (±8.39)	1799.0 (±10.69)	<0.001
BMI, kg/m^2	23.6 (±0.03)	25.0 (±0.05)	<0.001
Waist circumference, cm	81.2 (±0.09)	86.0 (±0.13)	<0.001
Systolic BP, mmHg	118.6 (±0.16)	130.8 (±0.24)	<0.001
Diastolic BP, mmHg	75.8 (±0.1)	76.6 (±0.16)	<0.001
Serum fasting glucose, mg/dL	99.7 (±0.22)	108.6 (±0.38)	<0.001
Total cholesterol, mg/dL	196.0 (±0.35)	186.5 (±0.53)	<0.001
Triglyceride, mg/dL	136.4 (±1.02)	150.0 (±1.41)	<0.001
WBC, thous/μL	5.9 (±0.02)	6.3 (±0.03)	<0.001
Hb, g/dL	14.0 (±0.02)	13.8 (±0.02)	<0.001

Continuous variables are presented as mean (± standard deviation), and categorical variables are presented as numbers (percentage, %).
Abbreviation: BMI, body mass index; BP, blood pressure; WBC, white blood cell; Hb, hemoglobin.

Table 2. General characteristics of internal validation dataset (KNHANES 2016 – 17).

	Internal validation dataset (KNHANES 2016 – 17) n=6787		
	Non-hypertension	Hypertension	
	n = 4476	n = 2311	P-value
Age, years	56.6 (±0.16)	66.4 (±0.21)	<0.001
Sex (female), n	2698 (60.3)	1230 (53.2)	<0.001
Income status, n			0.201
1st quartile	1054 (23.5)	591 (25.6)	
2nd quartile	1113 (24.9)	575 (24.9)	
3rd quartile	1130 (25.2)	579 (25.1)	
4th quartile	1179 (26.3)	566 (24.5)	
Education status, n			<0.001
Elementary school	987 (22.1)	1052 (45.5)	
Middle school	561 (12.5)	360 (15.6)	
High school	1446 (32.3)	528 (22.8)	
University	1482 (33.1)	371 (16.1)	
Diabetes mellitus, n	468 (10.5)	725 (31.4)	<0.001
Dyslipidemia, n	390 (8.7)	768 (33.2)	<0.001
Cancer, n	280 (6.3)	189 (8.2)	0.004
Alcohol consumption, g/week	54.5 (±1.79)	62.9 (±2.74)	0.01
Total energy intake, kcal	1872.8 (±11.95)	1737.9 (±16.43)	<0.001
BMI, kg/m^2	23.6 (±0.05)	25.1 (±0.07)	<0.001
Waist circumference, cm	81.7 (±0.14)	87.0 (±0.19)	<0.001
Systolic BP, mmHg	118.7 (±0.24)	129.6 (±0.35)	<0.001
Diastolic BP, mmHg	75.9 (±0.14)	75.7 (±0.23)	0.389
Serum fasting glucose, mg/dL	100.9 (±0.34)	112.0 (±0.67)	<0.001
Total cholesterol, mg/dL	200.0 (±0.56)	183.4 (±0.82)	<0.001
Triglyceride, mg/dL	136.1 (±1.62)	150.1 (±2.32)	<0.001
WBC, thous/μL	6.0 (±0.02)	6.5 (±0.04)	<0.001
Hb, g/dL	13.9 (±0.02)	13.9 (±0.03)	0.493

Continuous variables are presented as mean (± standard deviation), and categorical variables are presented as numbers (percentage, %).
Abbreviation: BMI, body mass index; BP, blood pressure; WBC, white blood cell; Hb, hemoglobin.

Table 3. Risk factors for hypertension selected by backward-stepwise logistic regression in Korean men.

	Model 1 (Univariate LR)	Model 2 (Multivariate LR)	Model 3 (Multivariate LR)
	OR (95% CI)	OR (95% CI)	OR (95% CI)
Age (years)	1.075 (1.074 - 1.075)	1.068 (1.068 - 1.069)	1.068 (1.068 - 1.069)
Income status (Ref: Q1)	1.033 (1.03 - 1.036)	1.032 (1.028 - 1.037)	1.032 (1.028 - 1.036)
Education status (Ref: Elementary)	0.739 (0.736 - 0.741)	0.999 (0.995 - 1.004)	-
Diabetes mellitus	2.963 (2.937 - 2.989)	1.531 (1.508 - 1.554)	1.531 (1.509 - 1.555)
Dyslipidemia	7.48 (7.383 - 7.579)	5.585 (5.503 - 5.669)	5.585 (5.502 - 5.669)
Cancer	1.43 (1.405 - 1.456)	0.968 (0.948 - 0.988)	0.968 (0.948 - 0.988)
Alcohol consumption (g/week)	1.003 (1.002 - 1.004)	1.038 (1.036 - 1.039)	1.038 (1.036 - 1.039)
Total energy intake (kcal)	0.637 (0.632 - 0.641)	0.882 (0.875 - 0.888)	0.881 (0.875 - 0.888)
BMI (kg/m^2)	1.116 (1.115 - 1.118)	1.12 (1.116 - 1.123)	1.12 (1.116 - 1.123)
Waist circumference (cm)	1.056 (1.056 - 1.057)	1.021 (1.02 - 1.022)	1.021 (1.02 - 1.022)
Systolic BP (mmHg)	1.036 (1.036 - 1.036)	1.031 (1.03 - 1.031)	1.031 (1.03 - 1.031)
Diastolic BP (mmHg)	1.001 (1 - 1.001)	0.998 (0.997 - 0.999)	0.998 (0.997 - 0.999)
Serum fasting glucose (mg/dL)	3.13 (3.092 - 3.168)	1.073 (1.051 - 1.095)	1.073 (1.051 - 1.095)
Total cholesterol (mg/dL)	0.991 (0.991 - 0.991)	0.994 (0.994 - 0.995)	0.994 (0.994 - 0.995)
Triglyceride (mg/dL)	1.053 (1.049 - 1.058)	1.07 (1.064 - 1.076)	1.07 (1.064 - 1.076)
WBC (thous/μL)	1.052 (1.05 - 1.054)	1.039 (1.037 - 1.042)	1.039 (1.037 - 1.042)
Hb (g/dL)	0.804 (0.801 - 0.806)	0.854 (0.851 - 0.858)	0.854 (0.851 - 0.858)
Abbreviation: LR, logistic regression; BMI, body mass index; BP, blood pressure; WBC, white blood cell; Hb, hemoglobin; OR, odds ratio; CI, confidence interval.			

Table 4. Risk factors for hypertension selected by backward-stepwise logistic regression in Korean women.

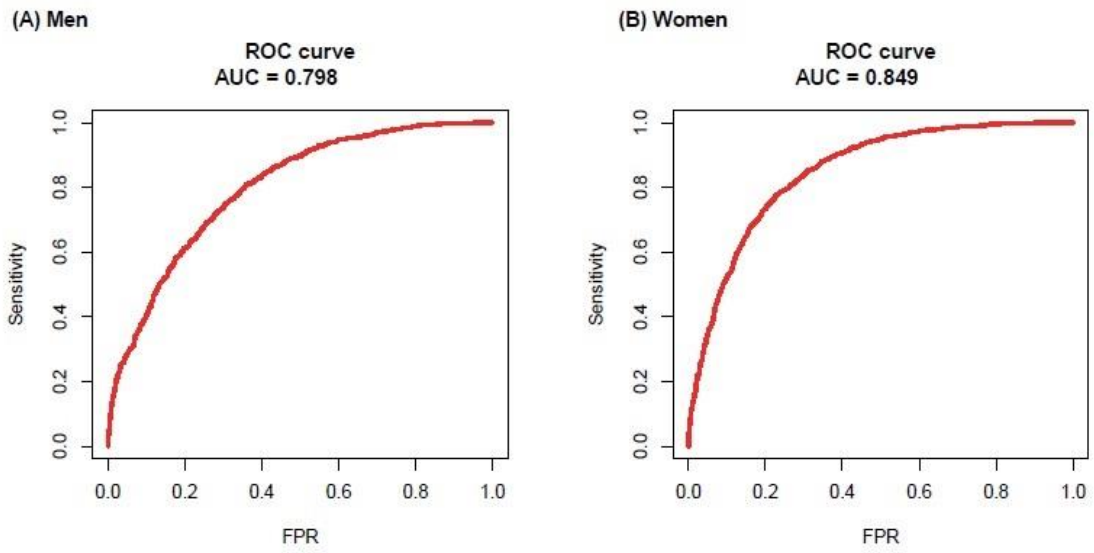
	Model 1 (Univariate LR)	Model 2 (Multivariate LR)	Model 3 (Multivariate LR)
	OR (95% CI)	OR (95% CI)	OR (95% CI)
Age (years)	1.103 (1.103 - 1.103)	1.078 (1.077 - 1.078)	1.078 (1.077 - 1.078)
Income status (Ref: Q1)	0.956 (0.953 - 0.959)	1.017 (1.013 - 1.021)	1.017 (1.013 - 1.021)
Education status (Ref: Elementary)	0.475 (0.473 - 0.476)	0.876 (0.872 - 0.881)	0.876 (0.872 - 0.881)
Diabetes mellitus	4.812 (4.765 - 4.86)	1.843 (1.814 - 1.873)	1.843 (1.814 - 1.873)
Dyslipidemia	5.324 (5.271 - 5.377)	2.87 (2.836 - 2.905)	2.87 (2.836 - 2.905)
Cancer	1.461 (1.44 - 1.482)	1.282 (1.26 - 1.304)	1.282 (1.26 - 1.304)
Alcohol consumption (g/week)	0.872 (0.871 - 0.874)	0.994 (0.992 - 0.996)	0.994 (0.992 - 0.996)
Total energy intake (kcal)	0.677 (0.673 - 0.681)	0.948 (0.942 - 0.955)	0.948 (0.942 - 0.955)
BMI (kg/m^2)	1.163 (1.161 - 1.164)	1.13 (1.127 - 1.133)	1.13 (1.127 - 1.133)
Waist circumference (cm)	1.069 (1.069 - 1.07)	0.993 (0.992 - 0.994)	0.993 (0.992 - 0.994)
Systolic BP (mmHg)	1.051 (1.051 - 1.051)	1.027 (1.027 - 1.027)	1.027 (1.027 - 1.027)
Diastolic BP (mmHg)	1.026 (1.026 - 1.026)	1.018 (1.018 - 1.019)	1.018 (1.018 - 1.019)
Serum fasting glucose (mg/dL)	6.632 (6.539 - 6.725)	1.202 (1.177 - 1.228)	1.202 (1.177 - 1.228)
Total cholesterol (mg/dL)	0.996 (0.996 - 0.997)	0.995 (0.995 - 0.995)	0.995 (0.995 - 0.995)
Triglyceride (mg/dL)	1.654 (1.646 - 1.661)	1.158 (1.151 - 1.165)	1.158 (1.151 - 1.165)
WBC (thous/μL)	1.179 (1.177 - 1.182)	1.054 (1.051 - 1.056)	1.054 (1.051 - 1.056)
Hb (g/dL)	1.076 (1.073 - 1.08)	0.986 (0.983 - 0.99)	0.986 (0.983 - 0.99)
Abbreviation: LR, logistic regression; BMI, body mass index; BP, blood pressure; WBC, white blood cell; Hb, hemoglobin; OR, odds ratio; CI, confidence interval.			

Table 5. Baseline characteristics of external validation dataset (KoGES).

	External validation dataset (KoGES) n=5152		P-value
	Non-hypertension	Hypertension	
	n = 3411	n = 1741	
Age, years	49.0 (±0.13)	52.7 (±0.21)	<0.001
Sex (female), n	1885 (55.3)	874 (50.2)	0.001
Income status, n			<0.001
1st quartile	397 (11.6)	345 (19.8)	
2nd quartile	933 (27.3)	581 (33.4)	
3rd quartile	1312 (38.5)	532 (30.6)	
4th quartile	769 (22.5)	283 (16.2)	
Education status, n			<0.001
Elementary school	766 (22.5)	627 (36.0)	
Middle school	768 (22.5)	410 (23.6)	
High school	1309 (38.4)	496 (28.5)	
University	568 (16.6)	208 (11.9)	
Diabetes mellitus, n	94 (2.8)	69 (4.0)	0.024
Dyslipidemia, n	68 (2.0)	35 (2.0)	0.999
Cancer, n	23 (0.7)	16 (0.9)	0.43
Alcohol consumption, g/week	55.6 (±2.32)	70.5 (±3.62)	0.001
Total energy intake, kcal	1942.0 (±10.38)	1968.9 (±16.57)	0.17
BMI, kg/m^2	23.8 (±0.05)	24.7 (±0.08)	<0.001
Waist circumference, cm	79.4 (±0.14)	83.4 (±0.2)	<0.001
Systolic BP, mmHg	109.2 (±0.19)	118.6 (±0.24)	<0.001
Diastolic BP, mmHg	72.9 (±0.14)	78.0 (±0.16)	<0.001
Serum fasting glucose, mg/dL	85.1 (±0.32)	86.5 (±0.47)	0.013
Total cholesterol, mg/dL	188.1 (±0.58)	190.5 (±0.83)	0.022
Triglyceride, mg/dL	142.4 (±1.46)	161.4 (±2.28)	<0.001
WBC, thous/μL	6.4 (±0.03)	6.5 (±0.04)	0.017
Hb, g/dL	13.4 (±0.03)	13.6 (±0.04)	<0.001

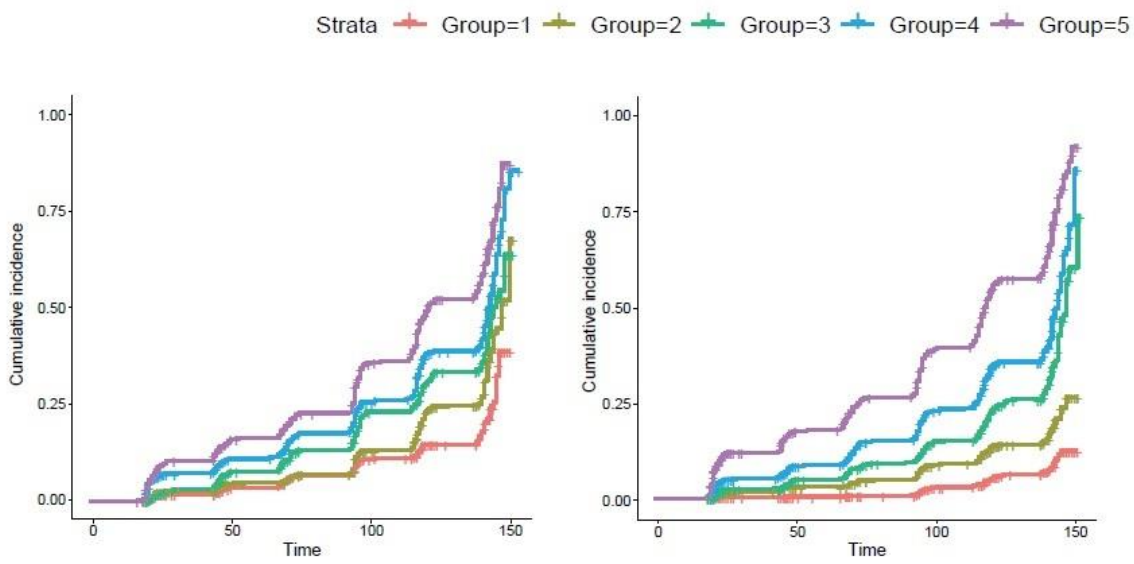
Continuous variables are presented as mean (± standard deviation), and categorical variables are presented as numbers (percentage, %).
Abbreviation: KoGES, Korean Genome and Epidemiology Study; BMI, body mass index; BP, blood pressure; WBC, white blood cell; Hb, hemoglobin.

Figure 1. Receiver operating characteristic curve for the present gender-specific hypertension prediction model using the KNHANES dataset from 2016 to 2017. (A) men (AUC = 0.798) (B) women (AUC = 0.849).



Abbreviations: KNHANES, Korea National Health and Nutrition Examination Survey; ROC, receiver operating characteristic; AUC, area under the receiver operative characteristic curves; FPR, false positive rate.

Figure 2. Cumulative incidence difference of new-onset hypertension between five groups, divided according to the expected probabilities of participants in the KoGES study.



Abbreviations: KoGES, Korean Genome and Epidemiology Study.

4. Discussion

We constructed the hypertension classification model using the 2011 – 15 KNHANES, which accurately classified subjects with hypertension in the 2016 – 17 KNHANES (IV set). Furthermore, we externally validated this model by predicting participants with the new onset of hypertension in the longitudinal set (KoGES). A hypertension risk score model from the Framingham Heart proposed an AUC of 0.788 [9]. The KoGES model by Lim et al. [16] resulted AUCs of 0.791 and 0.790 via coefficient- and point-based scores, respectively. A hypertension prediction model from the Atherosclerosis Risk in Communities study and the Cardiovascular Health Study proposed AUCs of 0.751 – 0.754 for 3 years follow-up and AUCs of 0.773 – 0.776 for 9 years follow-up [17]. Our model proposed AUCs of 0.798 and 0.849 in men and women, respectively, which were comparable with performances of other models.

For the input features of our model, we used the 17 variables in both men and women. The pathophysiology of hypertension is complicated [12], thereby, diverse risk factors for hypertension have been reported (Supplementary Table S1). However, in the previous studies, finger-countable clinical variables were used for the prediction model as a result of strict feature selection methods [9, 10, 16, 17]. In order to overcome this limitation, we used large-scale of cross-sectional study with weighting values instead of relative small size of longitudinal study. As a result, we included 17 various features of socioeconomic status, medical history, life style-related variables, anthropometric, and laboratory indices in the present study. Furthermore, these variables can be easily collected in the clinical setting, thereby, physicians can apply this classifying model in both the real hospital and epidemiologic researches.

Although several hypertension prediction models have been proposed, only a few studies were gender-specific models [6-8]. Epidemiologic studies have reported that gender difference for the prevalence or incidence of hypertension was significant [9-11]. Considering the underlying mechanism for hypertension, women might have different

hormonal effects from men for the development of hypertension, such as estrogen. Therefore, we constructed gender-specific hypertension prediction model.

In our final hypertension classification model of men, the positive association with hypertension is presented in following variables; age, income status, diabetes mellitus, dyslipidemia, alcohol consumption, BMI, waist circumference, systolic BP, serum fasting glucose, triglyceride, and WBC level. The negative relationship with hypertension is showed in cancer, total energy intake, diastolic BP, total cholesterol and Hb level. The result of women was revealed similarly to that of men except following several variables; education status, cancer, alcohol consumption, total energy intake, waist circumference and diastolic BP. Convergent results with previous studies were shown at systolic BP [1-3, 5] and triglyceride [12] in both men and women. However, inconsistent associations with other studies were proposed at total cholesterol [12-13] in both gender. In case of total cholesterol, age-adjusted LR showed positive correlation between the level of total cholesterol and risk of hypertension, but multivariate LR did not proposed significant result [12].

There were several limitations in the present studies. First, the dataset used for the prediction model was a cross-sectional study. We used this dataset due to two reasons, including large scale population and weight values that were estimated by the data constructors. To overcome this limitation, we attempted to validate this model using the longitudinal dataset. Second, we used the Korean dataset for the prediction model. Therefore, this model is difficult to apply in different countries, races, or cultures. Third, the LR algorithm was used for the hypertension prediction model. The LR is consisted of the weighted sum unit and the non-linear unit (sigmoid function), and has some crucial limitations, such as multicollinearity. We used the backward-stepwise regression to minimize the multicollinearity (Table 3 and 4). Recently, this like method was called as “shallow classifier” due to inability to represent the complex data. Consequently, deep learning can be used for the prediction model that consisted of a lot of input variables.

In conclusion, we proposed a hypertension risk index with Korean datasets,

including risk factors which could be collected in clinical office without difficulty. Physicians can utilize this model to predict the probability of hypertension individually, and have a caution the high-risk person to modify lifestyle and control their diseases.

APPENDICES

Table 1. Regression coefficients of the model for hypertension prediction in Korean men

	B	SE
Age (years)	0.066	0.000
Income status (Ref: Q1)	0.032	0.002
Education status (Ref: Elementary)	-	-
Diabetes mellitus	0.426	0.008
Dyslipidemia	1.720	0.008
Cancer	-0.032	0.010
Alcohol consumption (g/week)	0.037	0.001
Total energy intake (kcal)	-0.126	0.004
BMI ($kg/m^{2)}$	0.113	0.001
Waist circumference (cm)	0.021	0.001
Systolic BP (mmHg)	0.030	0.000
Diastolic BP (mmHg)	-0.002	0.000
Serum fasting glucose (mg/dL)	0.070	0.011
Total cholesterol (mg/dL)	-0.006	0.000
Triglyceride (mg/dL)	0.067	0.003
WBC (thous/μL)	0.038	0.001
Hb (g/dL)	-0.157	0.002
Constants	-9.892	0.087
Abbreviation: SE, standard error; BMI, body mass index; BP, blood pressure; WBC, white blood cell; Hb, hemoglobin;		

Table 2. Regression coefficients of the model for hypertension prediction in Korean women

	B	SE
Age (years)	0.075	0.000
Income status (Ref: Q1)	0.017	0.002
Education status (Ref: Elementary)	-0.132	0.003
Diabetes mellitus	0.612	0.008
Dyslipidemia	1.054	0.006
Cancer	0.248	0.009
Alcohol consumption (g/week)	-0.006	0.001
Total energy intake (kcal)	-0.053	0.004
BMI ($kg/m^{2)}$	0.122	0.001
Waist circumference (cm)	-0.007	0.000
Systolic BP (mmHg)	0.027	0.000
Diastolic BP (mmHg)	0.018	0.000
Serum fasting glucose (mg/dL)	0.184	0.011
Total cholesterol (mg/dL)	-0.005	0.000
Triglyceride (mg/dL)	0.147	0.003
WBC (thous/μL)	0.052	0.001
Hb (g/dL)	-0.014	0.002
Constants	-13.317	0.086
Abbreviation: SE, standard error; BMI, body mass index; BP, blood pressure; WBC, white blood cell; Hb, hemoglobin;		

Case study

Hypertension classification and an example from the generalized linear model (GLM).

The formula of the GLM:

$$\theta = \sum_{i=1}^{m} w_i x_i + w_0$$

$$\mu = f(\theta) = \frac{1}{1 + \exp(-\theta)}$$

θ indicates the combination of linear predictor; f represents a link function, which provides the relation between the linear predictor and the mean of the distribution function of the hypertension index (μ). We used logarithmic values in the following variables; alcohol consumption, total energy intake, serum fasting glucose and triglyceride. In an example case, the risk of a hypertension based on the GLM model is calculated, as follows:

Example case (woman)		
Variables	Parameters w_i	Values x_i (Score)
Age (years)	0.075	40
Income status (Ref: Q1)	0.017	2nd quartile (2)
Education status (Ref: Elementary)	-0.132	High school (3)
Diabetes mellitus	0.612	No (0)
Dyslipidemia	1.054	Yes (1)
Cancer	0.248	No (0)
Alcohol consumption (g/week)	-0.006	20
Total energy intake (kcal)	-0.053	2000
BMI (kg/m2)	0.122	23
Waist circumference (cm)	-0.007	70
Systolic BP (mmHg)	0.027	130
Diastolic BP (mmHg)	0.018	65
Serum fasting glucose (mg/dL)	0.184	85
Total cholesterol (mg/dL)	-0.005	180
Triglyceride (mg/dL)	0.147	115
WBC (thous/μL)	0.052	6
Hb (g/dL)	-0.014	13
Constants	-13.317	

$$\theta = \sum_{i=1}^{m} w_i x_i + w_0$$

$$= 0.075 \times 40 + 0.017 \times 2 + (-0.132) \times 3 + 0.612 \times 0 + 1.054 \times 1$$
$$+0.248 \times 0 + (-0.006) \times \log_2 20 + (-0.053) \times \log_2 2000 + 0.122 \times 23$$
$$+(-0.007) \times 70 + 0.027 \times 130 + 0.018 \times 65 + 0.184 \times \log_2 85$$
$$+(-0.005) \times 180 + 0.147 \times \log_2 115 + 0.052 \times 6$$
$$+(-0.014) \times 13 + (-13.317) = -2.305$$

$$\mu = f(\theta) = \frac{1}{1+\exp(-\theta)} = \frac{1}{1+\exp(2.305)} = 0.0907 = 9.07\%$$

Therefore, suggested hypertension probability of the example was reported as 9.07, which included in low-risk group ($\mu < 0.373$).

REFERENCE

1. Mills, K.T., et al., *Global Disparities of Hypertension Prevalence and Control: A Systematic Analysis of Population-Based Studies From 90 Countries.* Circulation, 2016. 134(6): p. 441-50.

2. Kang, S.H., et al., *Prevalence, Awareness, Treatment, and Control of Hypertension in Korea.* Sci Rep, 2019. 9(1): p. 10970.

3. Kjeldsen, S.E., *Hypertension and cardiovascular risk: General aspects.* Pharmacol Res, 2018. 129: p. 95-99.

4. Ku, E., et al., *Hypertension in CKD: Core Curriculum 2019.* Am J Kidney Dis, 2019. 74(1): p. 120-131.

5. Zhou, D., et al., *Uncontrolled hypertension increases risk of all-cause and cardiovascular disease mortality in US adults: the NHANES III Linked Mortality Study.* Sci Rep, 2018. 8(1): p. 9418.

6. Singh, S., R. Shankar, and G.P. Singh, *Prevalence and Associated Risk Factors of Hypertension: A Cross-Sectional Study in Urban Varanasi.* Int J Hypertens, 2017. 2017: p. 5491838.

7. Appel, L.J., et al., *Dietary approaches to prevent and treat hypertension: a scientific statement from the American Heart Association.* Hypertension, 2006. 47(2): p. 296-308.

8. Nakanishi, N., et al., *Alcohol consumption and risk for hypertension in middle-aged Japanese men.* J Hypertens, 2001. 19(5): p. 851-5.

9. Parikh, N.I., et al., *A risk score for predicting near-term incidence of hypertension: the Framingham Heart Study.* Ann Intern Med, 2008. 148(2): p. 102-10.

10. Kivimaki, M., et al., *Validating the Framingham Hypertension Risk Score: results from the Whitehall II study.* Hypertension, 2009. 54(3): p. 496-501.

11. Paynter, N.P., et al., *Prediction of incident hypertension risk in women with currently normal blood pressure.* Am J Med, 2009. 122(5): p. 464-71.

12. Beevers, G., G.Y. Lip, and E. O'Brien, *ABC of hypertension: The pathophysiology of hypertension.* Bmj, 2001. 322(7291): p. 912-6.

13. Chobanian, A.V., et al., *The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report.* Jama, 2003. 289(19): p. 2560-72.

14. Lee, J., et al., *Development of a suicide index model in general adolescents using the South Korea 2012-2016 national representative survey data.* Sci Rep, 2019. 9(1): p. 1846.

15. Pippitt, K., M. Li, and H.E. Gurgle, *Diabetes Mellitus: Screening and Diagnosis.* Am Fam Physician, 2016. 93(2): p. 103-9.

16. Lim, N.K., et al., *Predicting the risk of incident hypertension in a Korean middle-aged population: Korean genome and epidemiology study.* J Clin Hypertens (Greenwich), 2013. 15(5): p. 344-9.

17. Kshirsagar, A.V., et al., *A hypertension risk score for middle-aged and older adults.* J Clin Hypertens (Greenwich), 2010. 12(10): p. 800-8.

SUPPLEMENTARY

Table S1. References in the literature-based search for variables

Age, *years*	Hypertension and aging[1]	Buford TW	2016
Sex, *n*	Gender Differences in Epidemiology, Pathophysiology, and Treatment of Hypertension[2]	Di Giosia P.	2018
	Prevalence of hypertension in China[3]	Gao Y	2013
Ethnicity, *n*	Trends in prevalence, awareness, treatment, and control of hypertension in the United States	Hajjar I	2003
Education status, *n* Elementary school Middle school High school University	Education, literacy, and health: Mediating effects on hypertension knowledge and control[5]	Pandit A.U.	2009
	Socioeconomic status and hypertension[6]	Leng B	2015
Income status, *n* 1st quartile 2nd quartile 3rd quartile 4th quartile	Socioeconomic status and hypertension[6]	Leng B	2015
Diabetes mellitus, *n*	Hypertension and diabetes mellitus : coprediction and time trajectories[7]	Tsimihodimos V	2018
Dyslipidemia, *n*	Dyslipidemia and the Risk of Developing Hypertension in a Working-Age Male Population[8]	Otsuka T	2016
Cancer, *n*	Hypertension and breast cancer risk[9]	Han H	2017
	Hypertension and risk of prostate cancer[10]	Liang Z	2016
	Blood pressure and kidney cancer risk[11]	Hidayat K	2017
Alzheimer disease, *n*	Association between blood pressure and Alzheimer disease[12]	Gabin, J.M.	2017

Parkinson disease, *n*	Association of blood pressure and hypertension with the risk of Parkinson disease[13]	Qiu C	2011
Non-alcoholic fatty liver disease, *n*	Hypertension and Nonalcoholic Fatty Liver Disease Proven by Transient Elastography[14]	Wang Y	2016
Smoking, *PY*	Association between smoking and blood pressure[15]	Primatesta P	2001
Alcohol consumption, *g/week*	Effects of alcohol reduction on blood pressure[16]	Xin X	2001
	Alcohol consumption and risk for hypertension in middle-aged Japanese men[17]	Nakanishi N.	2001
Exercise, *n*	Progressive resistance exercise and resting blood pressure[18]	Kelley GA	2000
	Effect of aerobic exercise on blood pressure[19]	Whelton SP	2002
Diet Total energy intake, *kcal* Sodium intake, *mg*	Effects on blood pressure of reduced dietary sodium and the Dietary Approaches to Stop Hypertension (DASH) diet[20]	Sacks FM	2001
Potassium intake, *mg*	Effects of diet and sodium intake on blood pressure[21]	Vollmer WM	2001
	Dietary approaches to prevent and treat hypertension[22]	Appel L.J.	2006
Sleep duration, *hours*	Short sleep duration is associated with hypertension risk among adults[23]	Wang Q	2012
Menopause, *n*	The influence of menopause on blood pressure[24]	Staessen J	1989
	Menopause-related blood pressure increase and its relationship to age and body mass index[25]	Zanchetti A	2005
Oral contraceptive, *n*	Prospective study of oral contraceptives and hypertension among women in the United States[26]	Chasan-Taber L.	1996

BMI, *kg/m^2*	Influence of weight reduction on blood pressure[27]	Neter JE	2003
	Long-term effects of weight loss and dietary sodium reduction on incidence of hypertension[28]	He J	2000
Waist circumference, *cm*	The relationship of waist circumference to blood pressure[29]	Siani A	2002
Systolic BP, *mmHg*	Assessment of frequency of progression to hypertension in non-hypertensive participants in the Framingham Heart Study[30]	Vasan RS	2001
Diastolic BP, *mmHg*	Assessment of frequency of progression to hypertension in non-hypertensive participants in the Framingham Heart Study[30]	Vasan RS	2001
	High-normal blood pressure progression to hypertension in the Framingham Heart Study[31]	Leitschuh M	1991
White blood cell, *thous/μL*	Relationship between white blood cell count and incident hypertension[32]	Shankar A	2004
Hemoglobin, *g/dL*	Association of hematocrit with blood pressure and hypertension[33]	EmamianM	2017
Serum fasting glucose, *mg/dL*	The association between fasting plasma glucose and glycated hemoglobin in the prediabetes range and future development of hypertension[34]	Geva M	2019
Total cholesterol, *mg/dL*	A prospective study of plasma lipid levels and hypertension in women[35]	Sesso HD	2005
Triglyceride, *mg/dL*	Triglycerides and triglycerides to high-density lipoprotein cholesterol ratio are strong predictors of incident hypertension[36]	Tohidi M.	2012
Creatinine, *mg/dL*	A prospective study of blood pressure and serum creatinine[37]	Perneger TV	1993

Liver enzyme, *IU/L* Aspartate aminotransferase Alanine aminotransferase Gammaglutamyl trasferase	Gamma-glutamyltransferase is a predictor of incident diabetes and hypertension[38]	Lee DH	2003
C-reactive protein, *mg/dL*	C-reactive protein and the risk of developing hypertension[39]	Sesso HD	2003
	Multiple biomarkers and the risk of incident hypertension[40]	Wang TJ	2007
Interleukin-6, *ng/mL*	Inflammation and hypertension: the interplay of interleukin-6, dietary sodium and the renin-angiotensin system in humans[41]	Chamarthi B	2011
Aldosterone, *ng/dL*	Serum aldosterone and the incidence of hypertension in nonhypertensive persons[42]	Vasan RS	2004
Follicle-stimulating hormone, *IU/L*	Follicle-stimulating hormone, its association with cardiometabolic risk factors[43]	Wang N	2017
Estrogen, *pg/mL*	Serum estrogen metabolites and systolic blood pressure[44]	Masi CM	2009
Testosterone, *ng/dL*	Serum Testosterone Levels and Arterial Blood Pressure[45]	Fogari R	2005
Forced vital capacity, *L*	Rate of decline of forced vital capacity predicts future arterial hypertension[46]	Jacobs D.R	2012
Forced expiratory volume during 1 s, *L* Other	Blood pressure increase is inversely related to lung function[47]	Engstrom G	2001

REFERENCE OF SUPPLEMENTARY

1. Buford TW. Hypertension and aging. Ageing Res Rev. 2016;26:96–111.
2. Di Giosia P., et al. Gender Differences in Epidemiology, Pathophysiology, and Treatment of Hypertension. Curr Atheroscler Rep. 2018 14;20(3):13
3. Gao Y., Chen G., Tian H., et al. Prevalence of hypertension in China: a cross-sectional study. PLoS ONE. 2013;8(6)
4. Hajjar I, Kotchen TA. Trends in prevalence, awareness, treatment, and control of hypertension in the United States, 1988–2000. JAMA : the journal of the American Medical Association. 2003;290(2):199–206.
5. Pandit, A.U., et al., Education, literacy, and health: Mediating effects on hypertension knowledge and control. Patient Educ Couns, 2009. 75(3): p. 381-5.
6. Leng B, Jin Y, Li G. et al. Socioeconomic status and hypertension: a meta-analysis. Journal of hypertension. 2015;33(2):221–229
7. Tsimihodimos V, Gonzalez-Villalpando C, Meigs JB, Ferrannini E. Hypertension and diabetes mellitus: coprediction and time trajectories.Hypertension. 2018; 71:422–428.
8. Otsuka, T., et al., Dyslipidemia and the Risk of Developing Hypertension in a Working-Age Male Population. J Am Heart Assoc, 2016. 5(3): p. e003053.
9. Han H. et al. Hypertension and breast cancer risk: a systematic review and meta-analysis. Sci. Rep. 7, 44877 (2017)
10. Liang Z. et al. Hypertension and risk of prostate cancer: a systematic review and meta-analysis. Sci. Rep. 6, 31358 (2016)
11. Hidayat K. et al. Blood pressure and kidney cancer risk: meta-analysis of prospective studies. J Hypertens. 2017 Jul;35(7):1333-1344.
12. Gabin, J.M., Tambs, K., Saltvedt, I. et al. Association between blood pressure and Alzheimer disease measured up to 27 years prior to diagnosis: the HUNT Study. Alz Res Therapy 9, 37 (2017)

13. Qiu C, Hu G, Kivipelto M, et al. Association of blood pressure and hypertension with the risk of Parkinson disease: the National FINRISK Study. Hypertension, 57 (2011), pp. 1094-1100
14. Wang Y, Zeng Y, Lin C, Chen Z. Hypertension and Nonalcoholic Fatty Liver Disease Proven by Transient Elastography. Hepatol Res. 2016;46:1304–1310
15. Primatesta P, Falaschetti E, Gupta S, Marmot MG, Poulter NR. Association between smoking and blood pressure: evidence from the Health Survey for England. Hypertension. 2001; 37: 187–193.
16. Xin X, He J, Frontini MG. et al. Effects of alcohol reduction on blood pressure. Hypertension.2001;38:1112-1117.
17. Nakanishi, N., et al., Alcohol consumption and risk for hypertension in middle-aged Japanese men. J Hypertens, 2001. 19(5): p. 851-5.
18. Kelley GA, Kelley KS. Progressive resistance exercise and resting blood pressure. Hypertension.2000;35:838-843.
19. Whelton SP, Chin A, Xin X, He J. Effect of aerobic exercise on blood pressure. Ann Intern Med.2002;136:493-503.
20. Sacks FM, Svetkey LP, Vollmer WM. et al. for the DASH-Sodium Collaborative Research Group. Effects on blood pressure of reduced dietary sodium and the Dietary Approaches to Stop Hypertension (DASH) diet. N Engl J Med.2001;344:3-10.
21. Vollmer WM, Sacks FM, Ard J. et al. Effects of diet and sodium intake on blood pressure. Ann Intern Med.2001;135:1019-1028.
22. Appel, L.J., et al., Dietary approaches to prevent and treat hypertension: a scientific statement from the American Heart Association. Hypertension, 2006. 47(2): p. 296-308.
23. Wang Q, Xi B, Liu M. et al. Short sleep duration is associated with hypertension risk among adults: a systematic review and meta-analysis. Hypertension Research. 2012;35(10):1012–1018.

24. Staessen J, Bulpitt CJ, Fagard R, Lijnen P, Amery A. The influence of menopause on blood pressure. J Hum Hypertens. 1989;3:427–33.
25. Zanchetti A, Facchetti R, Cesana GC, Modena MG, Pirrelli A, Sega R. Menopause-related blood pressure increase and its relationship to age and body mass index: the SIMONA epidemiological study. J Hypertens. 2005;23:2269–76.
26. Chasan-Taber L., Wilett W.C., Manson J.E., et al. Prospective study of oral contraceptives and hypertension among women in the United States, Circulation, 1996, vol. 94 (pg. 483-489)
27. He J, Whelton PK, Appel LJ, Charleston J, Klag MJ. Long-term effects of weight loss and dietary sodium reduction on incidence of hypertension. Hypertension.2000;35:544-549.
28. Neter JE, Stam BE, Kok FJ, Grobbee DE, Geleijnse JM. Influence of weight reduction on blood pressure: a meta-analysis of randomized controlled trials. Hypertension. 2003; 42: 878–884
29. Siani A, Cappuccio FP, Barba G, Trevisan M, Farinaro E, Lacone R, Russo O, Russo P, Mancini M, and Strazzullo P. The relationship of waist circumference to blood pressure: the Olivetti Heart Study. Am J Hypertens 15: 780-786, 2002.
30. Vasan RS, Larson MG, Leip EP, Kannel WB, Levy D. Assessment of frequency of progression to hypertension in non-hypertensive participants in the Framingham Heart Study: a cohort study. Lancet. 2001;358:1682–1686.
31. Leitschuh M, Cupples LA, Kannel W, Gagnon D, Chobanian A. High-normal blood pressure progression to hypertension in the Framingham Heart Study. Hypertension. 1991;17:22–27.
32. Shankar A, Klein BE, Klein R . Relationship between white blood cell count and incident hypertension. Am J Hypertens 2004; 17: 233–239.
33. Emamian, M., et al., Association of hematocrit with blood pressure and hypertension. J Clin Lab Anal, 2017. 31(6).
34. Geva, M., Shlomai, G., Berkovich, A. et al. The association between fasting

plasma glucose and glycated hemoglobin in the prediabetes range and future development of hypertension. Cardiovasc Diabetol 18, 53 (2019).

35. Sesso HD, Buring JE, Chown MJ, et al. A prospective study of plasma lipid levels and hypertension in women. Archives of Internal Medicine. 2005;165(20):2420–2427.
36. Tohidi, M., Hatami, M., Hadaegh, F. et al. Triglycerides and triglycerides to high-density lipoprotein cholesterol ratio are strong predictors of incident hypertension in Middle Eastern women. J Hum Hypertens 2012; 26: 525–532.
37. Perneger TV, Nieto FJ, A prospective study of blood pressure and serum creatinine:results for the “Clue” Study and the ARIC Study, JAMA. 269 (1993) 488-493
38. Lee DH, Jacobs DR, Jr, Gross M, Kiefe CI, Roseman J, Lewis CE, Steffes M. Gamma-glutamyltransferase is a predictor of incident diabetes and hypertension: the Coronary Artery Risk Development in Young Adults (CARDIA) Study. Clin Chem. 2003;49:1358–1366.
39. Sesso HD, Buring JE, Rifai N, et al. C-reactive protein and the risk of developing hypertension. JAMA: The Journal of the American Medical Association. 2003;290(22):2945–2951.
40. Wang TJ, Gona P, Larson MG, et al. Multiple biomarkers and the risk of incident hypertension. Hypertension. 2007;49(3):432–438.
41. Chamarthi B, Williams GH, Ricchiuti V, Srikumar N, Paul N. Inflammation and hypertension: the interplay of interleukin-6, dietary sodium and the renin-angiotensin system in humans. Am J Hypertens. 2011;24(10):1–13.
42. Wang N, Shao H, Chen Y, Xia F, Chi C, Li Q, Han B, Teng Y, Lu Y. Follicle-stimulating hormone, its association with cardiometabolic risk factors, and 10-year risk of cardiovascular disease in postmenopausal women. J Am Heart Assoc. 2017;6(9):e005918.
43. Masi CM, Hawkley LC, Xu X, Veenstra TD, Cacioppo JT. Serum estrogen

metabolites and systolic blood pressure among middle-aged and older women and men. Am J Hypertens. 2009; 22:1148–1153

44. Fogari, R., Preti, P., Zoppi, A. et al. Serum Testosterone Levels and Arterial Blood Pressure in the Elderly. Hypertens Res 28, 625–630 (2005)
45. Vasan RS, Evans JC, Larson MG, et al. Serum aldosterone and the incidence of hypertension in nonhypertensive persons. N Engl J Med, 351 (2004), pp. 33-41
46. Jacobs, D.R., Jr., et al., Rate of decline of forced vital capacity predicts future arterial hypertension: the Coronary Artery Risk Development in Young Adults Study. Hypertension, 2012. 59(2): p. 219-25.
47. Engstrom G, Wollmer P, Valind S, Hedblad B, Janzon L. Blood pressure increase between 55 and 68 years of age is inversely related to lung function: longitudinal results from the cohort study ‘Men born in 1914.'J Hypertens. 2001; 19:1203–1208.

국 문 초 록

국민 건강 영양 조사와 한국 유전체 역학 연구를 활용한 성인 고혈압 지수 모델 개발

최근 고혈압에 대한 인식과 조절에 대한 중요성을 강조하면서, 고혈압 발병률을 예측하는 것 또한 중요하게 여겨진다. 본 연구는 2011년부터 2015년까지의 국민건강영양조사 인구 중 40세 이상의 대상자를 선정하였으며, 이들의 사회경제적 상태, 병력, 생활양식, 인체측정상태, 실험실 지표를 사용하여 각 성별에 따른 고혈압 예측 모델을 구성하였다. 2016년부터 2017년까지 국민건강영양조사의 40세 이상의 대상자를 선별하여 성인 고혈압 지수 모델의 수행력을 평가하였다. (남성 AUC = 0.798, 여성 AUC = 0.849). 또한 고혈압 모델의 성능은 종단 연구인 한국인유전체역학연구에서 추출한 누적 발생률과 비교하였으며, 본 연구의 고혈압 모델이 유의미하다는 것을 입증하였다. 본 연구는 진료실에서 쉽게 수집 할 수 있는 자료를 이용하여 고혈압 예측 모델을 만들었으며, 이를 통해 고혈압 위험도가 높은 사람을 선별해낼 수 있다.

핵심되는 말 : 고혈압