



Published in final edited form as:

J Clin Hypertens (Greenwich). 2010 October ; 12(10): 800–808. doi:10.1111/j.1751-7176.2010.00343.x.

A Hypertension Risk Score for Middle-Aged and Older Adults

Abhijit V. Kshirsagar, MD, MPH¹, Ya-lin Chiu, MS², Andrew S. Bombback, MD, MPH³, Phyllis A. August, MD⁴, Anthony J. Viera, MD, MPH⁵, Romulo E. Colindres, MD, MPSH¹, and Heejung Bang, PhD²

¹Department of Medicine, UNC Kidney Center and Division of Nephrology and Hypertension, School of Medicine, Chapel Hill, NC

²Department of Public Health, Division of Biostatistics and Epidemiology, Weill Medical College of Cornell University, New York, NY

³Division of Nephrology, Columbia University School of Medicine, New York, NY

⁴Department of Medicine, Division of Nephrology and Hypertension, Weill Medical College of Cornell University, New York, NY

⁵Department of Family Medicine, UNC School of Medicine, Chapel Hill, NC

Abstract

Determining which demographic and medical variables predict the development of hypertension could help clinicians stratify risk in both prehypertensive and nonhypertensive persons. Subject-level data from 2 community-based biracial cohorts were combined to ascertain the relationship between baseline characteristics and incident hypertension. Hypertension, defined as diastolic blood pressure ≥ 90 mm Hg, systolic blood pressure ≥ 140 mm Hg, or reported use of medication known to treat hypertension, was assessed prospectively at 3, 6, and 9 years. Internal validation was performed by the split-sample method with a 2:1 ratio for training and testing samples, respectively. A scoring algorithm was developed by converting the multivariable regression coefficients to integer values. Age, level of systolic or diastolic blood pressure, smoking, family history of hypertension, diabetes mellitus, high body mass index, female sex, and lack of exercise were associated with the development of hypertension in the training sample. Regression models showed moderate to high capabilities of discrimination between hypertension vs nonhypertension (area under the receiver operating characteristic curve 0.75–0.78) in the testing sample at 3, 6, and 9 years of follow-up. This risk calculator may aid health care providers in guiding discussions with patients about the risk for progression to hypertension.

The designation of prehypertension^{1,2} has two important goals.³ The first goal is to focus attention on a segment of the population with higher-than-normal cardiovascular disease risk.⁴⁻⁷ The second goal, often forgotten, is to identify individuals for whom targeted approaches to prevent or delay the onset of hypertension might be valuable.

©2010 Wiley Periodicals, Inc.

Address for correspondence: Abhijit V. Kshirsagar, MD, MPH, Department of Medicine, UNC Kidney Center and Division of Nephrology and Hypertension, School of Medicine, CB 7155 7017 Burnett-Womack Hall, Chapel Hill, NC 27599, sagar@med.unc.edu.

Author contributions: Study concept and design: Kshirsagar, Bang, and Bombback; Analysis and Interpretation of data: Bang, Chiu, Kshirsagar, and Bombback; Drafting of the manuscript: Kshirsagar, Bombback, and Bang; Critical revisions of manuscript and important intellectual content: August, Colindres, and Viera; Statistical expertise: Bang and Chiu; Obtaining funding: none; Study supervision: Kshirsagar and Bang.

In the United States alone, 70 million adults are estimated to have blood pressure (BP) within the broad range of prehypertension (systolic BP [SBP] from 120 to 139 mm Hg and/or diastolic BP [DBP] from 80 to 89 mm Hg).⁸ These individuals have a variety of comorbid medical conditions that may independently affect BP.^{7,9} Consequently, the risk of progression to overt hypertension likely varies, and clinicians face the challenge of correctly identifying individuals at high risk.

Determining which demographic and medical variables predict the development of hypertension could help clinicians risk-stratify nonhypertensive individuals, including those with prehypertension. Contemporary studies examining risk factors for progression may have limited generalizability because of the reliance on a single cohort¹⁰⁻¹² that lacks minority representation and has relatively short follow-up periods.

We therefore developed a simple method to help clinicians determine future risk for hypertension. We had two requirements for this model-based system: (1) the use of routinely available and minimally intrusive demographic and medical variables that are easily understood by lay persons and available to health care providers; and (2) the development of a model that determines the cumulative effect of multiple coexistent variables on incident hypertension.

METHODS

Study Design

We analyzed subject-level data from 2 community-based, prospective, public-use datasets to ascertain the relationship between baseline characteristics and incident hypertension. Risk prediction models were developed based on data available in nonhypertensive patients at baseline.

Study Population

Data from 2 nonconcurrent cohort studies were combined: the Atherosclerosis Risk in Communities (ARIC) study and the Cardiovascular Health Study (CHS) (<https://biolincc.nhlbi.nih.gov/home/>). Detailed descriptions of these two studies have been published previously.^{13,14}

Briefly, ARIC enrolled 15,732 participants aged 45 to 64 years between 1987 and 1989 (visit 1) from 4 communities and followed them for a maximum of 4 visits, approximately 3 years apart, for a maximum follow-up of 9 years. CHS recruited 5201 participants 65 years and older between 1989 and 1990 from 4 communities. Both studies recruited from 2 communities in common: Forsyth County, NC, and Washington County, MD. The two distinct recruiting regions selected by ARIC are suburban Minneapolis, MN, and Jackson, MS, whereas CHS recruited from Sacramento, CA, and Pittsburg, PA. Between 1992 and 1993, CHS enrolled an additional 687 black persons to increase minority participation. The CHS participants were followed annually for up to 10 years.

The ARIC and CHS cohorts are felt to be highly complementary in many aspects: (1) both were designed and conducted by the National Heart, Lung, and Blood Institute (NHLBI); (2) age ranges of the study participants and the study periods are continuous but without redundancy; (3) both studies enrolled black and white participants; and (4) data collection procedures, measurement processes, and study protocols were highly consistent.

In this analysis, we assessed the development of hypertension status at 3, 6, and 9 years of follow-up. Hypertension was defined as DBP \geq 90 mm Hg or SBP \geq 140 mm Hg or reporting the use of medication known to treat hypertension.

Measurements

We chose our candidate covariates among the ones that were validated from the literature and several new ones that are suspected of playing important roles in the development of hypertension.¹⁵⁻²⁹ As such, our covariate selection can be regarded as being guided by scientific as well as numeric evidence. The following variables served as standard candidate risk factors: age, sex, body mass index (BMI), waist circumference, exercise, diabetes, SBP, DBP, alcohol intake, smoking, and family history of hypertension. Additionally, we examined the predictability of various laboratory variables such as total, high-density lipoprotein, and low-density lipoprotein cholesterol and creatinine, and nutritional or diet-related variables such as the consumption of soft drinks and intake of sodium, total calories, carbohydrates, and fat.

BMI was computed using body weight and height measured under standardized conditions by trained and certified staff. Physical activity level was estimated by a “sport index,”³⁰ incorporating frequency, duration, and intensity of reported sports activities, information on leisure time activities, and sweating. Caloric intake was assessed through a semiquantitative food frequency interview.^{31,32} All medications taken during the 2 weeks prior to each study examination were recorded based on bottle labels if provided (otherwise by participant report). Smoking status, educational level, alcohol consumption, and parental history of hypertension were self-reported.

While missing data are minimal for most variables, this was not the case for family history of hypertension, a variable that was not measured in the CHS. Imputation of family history of hypertension was performed in the CHS cohort (detailed in the next subsection).

Statistical Analyses

We identified a set of risk factors that were associated with the outcome of interest using continuous covariates (for maximal use of information and power) and then categorized the continuous covariates using well-accepted cut-points whenever available or otherwise using intuitive, user-friendly cut-points. As such, continuous variables were used to select significant risk factors and categorical variables were used to develop a risk scoring algorithm. Specifically, we categorized BP variables at baseline using a 5-mm Hg interval for SBP and a 10-mm Hg interval for DBP. Since those are the strongest predictors of future hypertension, the use of multiple categories would capture the risk gradient properly. For the obesity measure, we used clinical guidelines (<http://www.nhlbi.nih.gov/guidelines/obesity/>). Age was also categorized using 10-year intervals starting from 45. These variables were used for the risk scoring algorithm.

The definitions of some variables in the two separate cohorts, ARIC and CHS, differed. For example, exercise in ARIC was defined as a binary variable (yes vs no), while in CHS it was defined as a 3-level exposure (low, moderate, and high). Thus, we dichotomized the variable in CHS by “no” vs all others. Family history of hypertension, which was collected in ARIC but not in CHS, was imputed using a statistical technique for missing data (by MI procedure in SAS [SAS Institute, Cary, NC]).³³ This method is regarded as a standard method for handling partially missing data in statistics, and imputed data are created using the information contained in the observed data (here in ARIC).

Baseline characteristics of study participants are summarized using mean and standard deviation for continuous variables and percentage for categorical variables. We randomly split the study cohort into training sample and testing sample using a 2:1 ratio for internal validation.

In the training sample, multiple logistic regression was used to identify independent risk factors for hypertension events. In the regression modeling, the main effects of individual risk factors and their interaction effects with age were tested. In order to generate a simple risk scoring algorithm, we intentionally did not consider/test other interaction effects because interaction effects can increase the complexity of the algorithm greatly without resulting in sufficient improvements in numeric and/or clinical performance characteristics. We employed backward elimination (deleting a covariate with the largest P value at a time) until we reached a parsimonious model that includes a set of covariates with all P values $< .05$ (a conventional P value threshold) in the “ever” model, in which the outcome was defined as the ever occurrence of hypertension during 9-year follow-up. We fitted the resulting parsimonious model with the final set of the risk factors for 3-, 6-, and 9-year models, in which the outcome was defined as the occurrence of hypertension at 3-, 6-, and 9-year follow-ups, respectively. We also fitted the same sets of models after categorizing the continuous covariates.

The magnitude of association between a risk factor and the outcome, controlling for other covariates, was assessed by odds ratios along with 95% confidence intervals; statistical significance was assessed by P value. We adopted a discrimination statistic, the area under the receiver operating characteristic curve (AUC) to quantify the capability to distinguish the events vs no events of hypertension.³⁴ We computed the AUCs in the training as well as testing samples in order to understand the discrimination capability of the final model in different samples and with continuous vs categoric variables.

For constructing a risk assessment algorithm, we created a weighted risk scoring algorithm by rounding the odds ratios in the final model to the nearest integer while preserving monotonicity from the “training” sample. The 3-, 6-, and 9-year risks for total scores were estimated from the model (by averaging subject-specific predicted risks derived from the final model) as well as empirically (by counting how many people developed hypertension among those who had the same total score at base-line) from the “testing” sample.

Finally, we re-generated the risk scoring algorithm and the risk table from the combined cohort (training + testing samples) and suggest that these to be used in practice. Rationales are that: (1) the similar results were obtained and we judged that the difference observed is not likely to have clinical relevance, and (2) we hoped to come up with the most accurate scoring algorithm and risk estimates by using the maximal sample size/information in the given dataset (ARIC/CHS), while waiting for true external validation to be conducted in the future.

Statistical analysis was performed using SAS version 9.1 software (SAS Institute, Cary, NC). Two-sided hypotheses and tests were adopted for all statistical inferences.

RESULTS

The study cohort consisted of 11,407 men and women after exclusion of individuals with prevalent hypertension. The training cohort consisted of 7683 participants, while the testing cohort consisted of 3724 participants (based on a 2:1 ratio for model development and internal validation). After excluding the patients with missing covariate data, the actual sample sizes used for analyses were 7610 for the training sample and 3692 for the testing sample.

Several baseline characteristics of the cohort are presented in Table I. The average age of the cohort was 56 years, while the average BP was 115 mm Hg (systolic) and 70 mm Hg (diastolic). The majority of participants (6709, or 59%) had BP below the prehypertensive range. Other demographic and medical characteristics are also available in Table II.

Progression to Hypertension

Among all participants included in the analysis, 33% (3795 of 11,407) progressed to hypertension during the study period. Of the prehypertensive patients, 52% (2450 of 4693) developed hypertension by study end, while 20% (1342 of 6709) of participants without baseline prehypertension developed incident hypertension.

Risk Factors Associated With Incident Hypertension

Risk factors associated with incident hypertension at 3, 6, and 9 years of follow-up are presented in Table II. Several characteristics were associated with developing hypertension throughout the study period—age, SBP, DBP, smoking, family history of hypertension, presence of diabetes mellitus, BMI, and the age–DBP interaction. Statistical significance for female sex, family history of diabetes, and exercise did not reach a conventional threshold (ie, $P < .05$) in some models but the directions of the association with the outcomes were consistent.

Internal Validation Evaluation of the final models (using continuous or categorical variables) in the testing sample revealed the regression coefficients to be highly similar to the ones obtained from the training sample. There was some attenuation of the statistical significance of some of the variables, primarily due to smaller sample size in the testing sample (results not shown). The regression models showed moderate to high capabilities of discrimination between hypertension vs nonhypertension (AUC 0.74–0.80) in the training sample (AUC 0.74–0.77) in the testing sample.

Variable categorization did not result in considerable loss in the discrimination capability, further justifying the use of an integer-based scoring system (Table III). We calculated the hypertension risks for each individual in the testing sample using a scoring algorithm derived from the training sample. To illustrate, Figure 1 displays 9-year risk per total score at baseline, where the estimated risk from the model and the empirically estimated risk were presented together.

Risk Scoring Algorithm

After finalizing the model, we fitted the final model in the combined sample (ie, training + testing) to derive the ultimate scoring algorithm. Figure 2 provides a risk assessment questionnaire that consists of 9 questions—age, sex, smoking, exercise, family history of hypertension, BMI, diabetes, SBP, and DBP. Based on the total score from this questionnaire, an individual's risk for developing incident hypertension can then be estimated at 3, 6, and 9 years (Table IV). For example, a man aged 45 with optimal BP (SBP <110 mm Hg and DBP <70 mm hg) and no other risk factors would have a 3% to 5% risk of developing hypertension in the next 3 to 9 years. At the other extreme, an older prehypertensive woman who has many of the listed risk factors (eg, obesity, smoking, and lack of exercise) has >70% chance of developing hypertension in the future.

DISCUSSION

We have developed a scoring algorithm to stratify persons at risk for developing hypertension. This prediction rule translates a parsimonious set of medical and demographic characteristics into an average like-lihood of developing hypertension among middle-aged and older adults with up to 9 years of follow-up. These characteristics are often present together and cumulatively affect the risk of hypertension. Most of the characteristics (eg, age, sex, BP level) are easily identified by both health care providers and the general public.

We found that the strongest predictors of incident hypertension were: (1) baseline SBP in the pre-hypertension range; (2) advanced age; and (3) obesity. The treatment of hypertension in the elderly can be complicated, and there still remains uncertainty as to whether the BP goals we set for young individuals pertain to older individuals.^{35,36} Based on the risk score that integrates and quantifies comorbid conditions, it may be possible to individualize the approach to elderly persons. For example, a 75-year-old who is nondiabetic, regularly exercises, and has never smoked could be viewed (and managed) differently from a 75-year-old with identical systolic and diastolic recordings but a history of tobacco abuse, inactivity, and impaired glucose tolerance. With regards to the potential effect of obesity on hypertension risk, this tool may help fuel discussions between obese patients and their primary care providers about the need for dietary and exercise lifestyle modifications, and could feasibly even provide the context for discussions about weight loss reduction medication or surgery.

We were surprised to find that certain variables did not independently predict onset of hypertension after multivariate adjustment, including cholesterol, kidney function, and caloric intake. We would stress that these are still important covariates for assessment of cardiovascular risk, but maybe not as necessary for development of hypertension. Rather, this study finds evidence that the 8 identified risk factors capture the overall risk very well. While clinicians still continue to provide standard recommendations to patients, including healthy eating and cholesterol lowering, the 8 risk factors may offer an improved and more targeted strategy to prevent hypertension.

Risk scores are practical tools to help identify individuals at an increased risk for adverse health outcomes. Clinicians can use our scoring system to communicate expected risk with patients and to facilitate discussions about possible preventive strategies. The algorithm may also motivate individuals to engage lifestyle change to lower the risk of hypertension and to bring this topic to their health care providers. It also has potential public health applications. It can be posted on medical Web sites for the public to access or may be used in community settings to identify individuals who may wish to be referred to health care providers. The identification of high-risk individuals with this scoring system can also be used to optimize the benefit and cost-effectiveness of targeted screenings. At the very least, the prediction rule can be used in concert with other public health initiatives to increase the awareness of hypertension and hypertension-associated diseases.

LIMITATIONS AND STRENGTHS

This study should be evaluated in the context of the following limitations. Given the inherent variability of BP, some participants may have been mis-classified even after following a standard protocol. Yet, we believe that this would likely have been nondifferential misclassification bias and thus would move any point estimates toward the null. Second, the BP measurements were office- and clinic-based, whereas recent evidence suggests that 24-hour BP recordings may bear more weight on clinical outcomes. These results must therefore be interpreted as solely predicting the risk for progression to clinic-based hypertension and may not reflect 24-hour ambulatory hypertension. Third, it was necessary to impute data on family history for participants in the CHS cohort. The validity of imputation is contingent on the data being missing at random, given a set of available information. Finally, we did not validate our model in an external or independent dataset. External validation on various populations/cohorts will help clarify the potential generalizability of the algorithm to larger populations. In the future, we plan to assess the performance of the risk calculator in independent datasets and community screening efforts.

Still, the study has some notable strengths. We used data from a large, community-based sample drawn from multiple locations, and participants were followed for 9 years. Both the demographic composition of our dataset and the extended follow-up period of this dataset add to previous work¹⁰⁻¹² performed primarily in the Framingham cohort. The Framingham cohort lacks sufficient data on blacks, and the follow-up period was confined to 4 years. Our sample included representative data from both blacks and whites, and the nearly decade-long follow-up period in this data gives clinicians an idea about not only which patients will likely progress to frank hypertension, but also when these patients will do so. In addition, the complementary age of the participants of the two cohorts, ARIC and CHS, provides an age range that mirrors the age range of most individuals who are at risk for developing hypertension. The study, however, does not include young adults, who may be more commonly affected by obesity, and its subsequent effect on the development of hypertension.

In addition to the set of risk factors identified by the Framingham cohort, we found that exercise/physical activity significantly contributes to our new model. This is an important finding because exercise is a scientifically supported protective factor that often fails to reach statistical significance due to various issues (eg, difficulty in quantifying, measurement error/misclassification). Moreover, it is a highly modifiable and dynamic behavioral/life-style factor; change in exercise patterns can modify the risk score/status, in contrast to demographic and health history variables, which are not modifiable. We focused on scientifically validated risk factors instead of identifying novel risk factors and elucidating complicated relationships among the risk factors and/or the outcome (eg, interaction or nonlinear relationship) in order to develop a user-friendly risk scoring algorithm.

CONCLUSIONS

We have developed a tool for predicting progression to overt hypertension that relies on readily available demographic and medical risk factors. The risk calculator may be most judiciously applied to individuals with prehypertension to better delineate subsequent risk for progression to hypertension. It may furthermore facilitate discussions on how to most effectively modify this risk.³⁷⁻⁴⁰

Acknowledgments

Disclosures: The ARIC and CHS studies are conducted and supported by the NHLBI in collaboration with ARIC and CHS investigators, respectively. This manuscript was prepared using limited access datasets obtained from the NHLBI and does not necessarily reflect the opinions or views of the ARIC/CHS or NHLBI. The authors thank the staff and participants of the ARIC/CHS study for their important contributions and valuable information for health research. Dr Bang and Ms Chiu were partially supported by CSTC grant at Weill Cornell Medical College (UL1-RR024996). Dr Viera was supported by a career development award from the National Institutes of Health (KL2RR025746). Dr Bang and Ms Chiu had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Chobanian AV, Bakris GL, Black HR, et al. Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. *Hypertens*. 2003; 42:1206–1252.
2. Chobanian AV, Bakris GL, Black HR, 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; 289:2560–2572. [PubMed: 12748199]
3. Chobanian AV. Prehypertension revisited. *Hypertens*. 2006; 48:812–814.

4. Vasan RS, Larson MG, Leip EP, et al. Impact of high-normal blood pressure on the risk of cardiovascular disease. *N Engl J Med.* 2001; 345:1291–1297. [PubMed: 11794147]
5. Qureshi AI, Suri MF, Kirmani JF, et al. Is prehypertension a risk factor for cardiovascular diseases? *Stroke.* 2005; 36:1859–1863. [PubMed: 16081866]
6. Liszka HA, Mainous AG III, King DE, et al. Prehypertension and cardiovascular mortality. *Ann Fam Med.* 2005; 3:294–299. [PubMed: 16046560]
7. Kshirsagar AV, Carpenter M, Bang H, et al. Blood pressure usually considered normal is associated with an elevated risk of cardiovascular disease. *Am J Med.* 2006; 119:133–141. [PubMed: 16443415]
8. Wang Y, Wang QJ. The prevalence of prehypertension and hypertension among US adults according to the new joint national committee guidelines: new challenges of the old problem. *Arch Intern Med.* 2004; 164:2126–2134. [PubMed: 15505126]
9. Greenlund KJ, Croft JB, Mensah GA. Prevalence of heart disease and stroke risk factors in persons with prehypertension in the United States, 1999–2000. *Arch Intern Med.* 2004; 164:2113–2118. [PubMed: 15505124]
10. Vasan RS, Larson MG, Leip EP, et al. Assessment of frequency of progression to hypertension in non-hypertensive participants in the Framingham Heart Study: a cohort study. *Lancet.* 2001; 358:1682–1686. [PubMed: 11728544]
11. Vasan RS, Beiser A, Seshadri S, et al. Residual lifetime risk for developing hypertension in middle-aged women and men: the Framingham Heart Study. *JAMA.* 2002; 287:1003–1010. [PubMed: 11866648]
12. Parikh NI, Pencina MJ, Wang TJ, et al. A risk score for predicting near-term incidence of hypertension: the Framingham Heart Study. *Ann Intern Med.* 2008; 148:102–110. [PubMed: 18195335]
13. ARIC Investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. *Am J Epidemiol.* 1989; 129:687–702. [PubMed: 2646917]
14. Fried LP, Borhani NO, Enright P, et al. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol.* 1991; 1:263–276. [PubMed: 1669507]
15. Burke GL, Savage PJ, Sprafka JM, et al. Relation of risk factor levels in young adulthood to parental history of disease. The CARDIA study. *Circulation.* 1991; 84:1176–1187. [PubMed: 1884448]
16. Jones DW. Body weight and blood pressure. Effects of weight reduction on hypertension. *Am J Hypertens.* 1996; 9:50s–54s. [PubMed: 8862237]
17. Rebbeck TR, Turner ST, Sing CF. Probability of having hypertension: effects of sex, history of hypertension in parents, and other risk factors. *J Clin Epidemiol.* 1996; 49:727–734. [PubMed: 8691221]
18. Zheng ZJ, Folsom AR, Ma J, et al. Plasma fatty acid composition and 6-year incidence of hypertension in middle-aged adults: the Atherosclerosis Risk in Communities (ARIC) Study. *Am J Epidemiol.* 1999; 150:492–500. [PubMed: 10472949]
19. Dyer AR, Liu K, Walsh M, et al. Ten-year incidence of elevated blood pressure and its predictors: the CARDIA study. *Coronary Artery Risk Development in (Young) Adults.* *J Hum Hypertens.* 1999; 13:13–21. [PubMed: 9928747]
20. Fuchs FD, Chambless LE, Whelton PK, et al. Alcohol consumption and the incidence of hypertension: the Atherosclerosis Risk in Communities Study. *Hypertens.* 2001; 37:1242–1250.
21. van der Sande MA, Walraven GE, Milligan PJ, et al. Family history: an opportunity for early interventions and improved control of hypertension, obesity and diabetes. *Bull World Health Organ.* 2001; 79:321–328. [PubMed: 11357211]
22. Diez Roux AV, Chambless L, Merkin SS. Socioeconomic disadvantage and change in blood pressure associated with aging. *Circulation.* 2002; 106:703–710. [PubMed: 12163431]
23. Sinclair AM, Isles CG, Brown I, et al. Secondary hypertension in a blood pressure clinic. *Arch Intern Med.* 1987; 147:1289–1293. [PubMed: 3606286]
24. Coresh J, Wei GL, McQuillan G, et al. Prevalence of high blood pressure and elevated serum creatinine level in the United States: findings from the third National Health and Nutritional Examination Survey (1988–1994). *Arch Intern Med.* 2001; 161:1207–1216. [PubMed: 11343443]

25. Juhaeri, Stevens J, Chambless LE, et al. Associations between weight gain and incident hypertension in a bi-ethnic cohort: the Atherosclerosis Risk in Communities Study. *Int J Obes Relat Metab Disord.* 2002; 26:58–64. [PubMed: 11791147]
26. Rose KM, Holme I, Light KC, et al. Association between the blood pressure response to a change in posture and the 6-year incidence of hypertension: prospective findings from the ARIC study. *J Hum Hypertens.* 2002; 16:771–777. [PubMed: 12444538]
27. Juhaeri, Stevens J, Chambless LE, et al. Associations of weight loss and changes in fat distribution with the remission of hypertension in a bi-ethnic cohort: the Atherosclerosis Risk in Communities Study. *Prev Med.* 2003; 36:330–339. [PubMed: 12634024]
28. Yu D, Huang J, Hu D, et al. Prevalence and risk factors for prehypertension among Chinese adults. *J Cardiovasc Pharmacol.* 2008; 52:363–368. [PubMed: 18841073]
29. Erem C, Hacıhasanoglu A, Kocak M, et al. Prevalence of prehypertension and hypertension and associated risk factors among Turkish adults: Trabzon Hypertension Study. *J Public Health (Oxf).* 2009; 31:42–58.
30. Baecke JA, Burema J, Frijters JE. A short questionnaire for the measurement of habitual physical activity in epidemiologic studies. *Am J Clin Nutr.* 1982; 36:939–942.
31. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol.* 1985; 122:51–65. [PubMed: 4014201]
32. Shimakawa T, Sorlie P, Carpenter MA, et al. Dietary intake patterns and sociodemographic factors in the atherosclerosis risk in communities study. ARIC Study Investigators. *Prev Med.* 1994; 23:769–780. [PubMed: 7855109]
33. SAS Institute Inc. SAS/STAT User's Guide Version 8.2. SAS Institute Inc; Cary, NC: 2000.
34. Gonen, M. Analyzing Receiver Operating Characteristic Curves with SAS. SAS Publishing; Cary, NC: 2007.
35. Port S, Demer L, Jennrich R, et al. Systolic blood pressure and mortality. *Lancet.* 2000; 355:175–180. [PubMed: 10675116]
36. Pastor-Barriuso R, Banegas JR, Damián J, et al. Systolic blood pressure, diastolic blood pressure, and pulse pressure: an evaluation of their joint effect on mortality. *Ann Intern Med.* 2003; 139:731–739. [PubMed: 14597457]
37. Appel LJ, Sacks FM, Carey VJ, et al. for the OmniHeart Collaborative Research Group. Effects of protein, mono-unsaturated fat, and carbohydrate intake on blood pressure and serum lipids: results of the OmniHeart randomized trial. *JAMA.* 2005; 294:2455–2464. [PubMed: 16287956]
38. Elmer PJ, Obarzanek E, Vollmer WM, et al. for the PREMIER Collaborative Research Group. Effects of comprehensive lifestyle modification on diet, weight, physical fitness, and blood pressure control: 18-month results of a randomized trial. *Ann Intern Med.* 2006; 144:485–495. [PubMed: 16585662]
39. Julius S, Nesbitt SD, Egan BM, et al. for the Trial of Preventing Hypertension (TROPHY) Study Investigators. Feasibility of treating prehypertension with an angioten-sin-receptor blocker. *N Engl J Med.* 2006; 354:1685–1697. [PubMed: 16537662]
40. Chobanian AV. The hypertension paradox- more uncontrolled disease despite improved therapy. *N Engl J Med.* 2009; 361:878–887. [PubMed: 19710486]

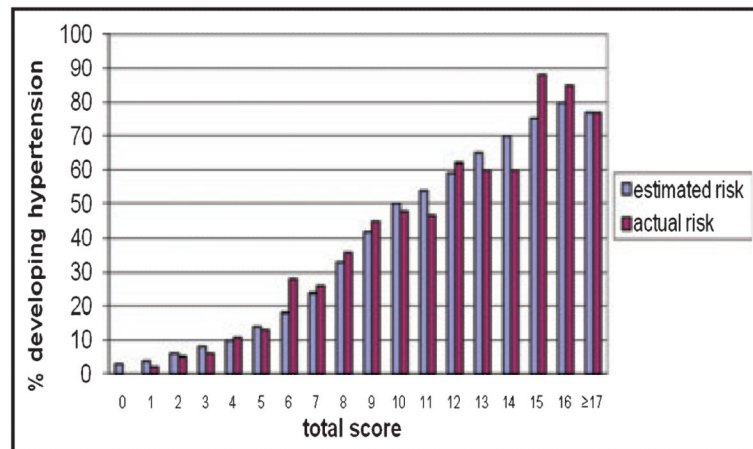


Figure 1.

Nine-year risk of developing hypertension in testing sample (N=3692) based on risk score derived from training sample. (Scores 17 or greater were combined due to small sample sizes in the testing sample for reliable estimation of risk. Risk score derived from the training sample can be found in Figure 2. Blue bars denote risk estimated from the model and purple bars denote risk estimated by counting [the number of events divided by the total number of individuals at risk at baseline]).

	VARIABLE	POINTS	
Step 1	Age (years) • 45-54 • 55-64 • 65-74 • ≥ 75	0 2 3 4	Subtotal:
Step 2	Sex • Male • Female	0 1	Subtotal:
Step 3	Smoking • No • Yes	0 1	Subtotal:
Step 4	Exercise • No • Yes	1 0	Subtotal:
Step 5	Family History of Hypertension • No • Yes	0 1	Subtotal:
Step 6	Body Mass Index • < 25 (normal weight) • 25-29 (overweight) • 30-39 (obese) • ≥ 40 (extremely obese)	0 1 2 3	Subtotal:
Step 7	Diabetes • No • Yes	0 2	Subtotal:
Step 8	Systolic Blood Pressure (mm Hg) • < 110 • 110-114 • 115-119 • 120-124 • 125-129 • 130-134 • 135-139	0 2 3 4 6 8 14	Subtotal:
Step 9	Diastolic Blood Pressure (mm Hg) If age < 55 years • < 70 • 70-79 • >80 If age 55-64 years • 70-79 • >80 If age 65-74 years • 70-79 • >80 If age ≥ 75 years • 70-79 • >80	0 2 3 -1 -1 -2 -3 -1 -2	Subtotal: Total Score:

Figure 2. Risk scoring algorithm. (The final algorithm was derived from the combined sample [training + testing sample]).

Table I

Baseline Characteristics of the Original Study Cohort (N=11,407)

Variable	Mean (SD) or Percentage
Age, y	56 (9.0)
Sex, % female	54
Race, % white	83
Education, % <high school	48
Systolic BP, mm Hg	115 (12.5)
Diastolic BP, mm Hg	70 (8.8)
Total cholesterol, mmol/L	5.4 (1.0)
LDL cholesterol, mmol/L	3.5 (1.0)
HDL cholesterol, mmol/L	1.4 (0.4)
Creatinine, μ mol/L	88.4 (26.5)
BMI, kg ² /cm	26 (4.6)
Current exercise, %	71
Current alcohol use, %	74
Diabetes mellitus, %	7
Current smoker, %	25
Former smoker, %	34
Parental history of hypertension, ^a %	46

Abbreviations: BMI, body mass index; BP, blood pressure; HDL, high-density lipoprotein; LDL, low-density lipoprotein; SD, standard deviation.

^aData was imputed for the Cardiovascular Health Study (CHS) cohort by a statistical technique using the information available in the Atherosclerosis Risk in Communities (ARIC) cohort as this variable was not collected in the CHS.

Table II

Multiple Logistic Regression Models for Incident Hypertension

COVARIATE	3 Years (AUC=0.742)		6 Years (AUC=0.750)		9 Years (AUC=0.791)		Ever (AUC=0.775)	
	Odds Ratio (95% CI)	P Value	Odds Ratio (95% CI)	P Value	Odds Ratio (95% CI)	P Value	Odds Ratio (95% CI)	P Value
Age (1-y increment)	1.14 (1.08-1.19)	<.0001	1.07 (1.02-1.12)	.0093	1.23 (1.17-1.29)	<.0001	1.17 (1.12-1.22)	<.0001
Female	1.01 (0.89-1.15)	.86	1.09 (0.97-1.22)	.141	1.50 (1.34-1.68)	<.0001	1.34 (1.21-1.48)	<.0001
Systolic BP (1-mm Hg increment)	1.06 (1.06-1.07)	<.0001	1.07 (1.03-1.12)	<.001	1.09 (1.08-1.09)	<.0001	1.08 (1.07-1.08)	<.0001
Diastolic BP (1-mm Hg increment)	1.13 (1.08-1.18)	<.0001	1.07 (1.03-1.12)	.0007	1.18 (1.13-1.24)	<.0001	1.13 (1.09-1.17)	<.0001
Current smoker	1.24 (1.09-1.40)	.0011	1.23 (1.10-1.38)	.0003	1.27 (1.14-1.42)	<.0001	1.34 (1.22-1.48)	<.0001
Family history of hypertension	1.14 (1.01-1.28)	.038	1.19 (1.07-1.33)	.0015	1.38 (1.23-1.54)	<.0001	1.26 (1.14-1.39)	<.0001
Body mass index (1-kg ² /m increment)	1.03 (1.02-1.04)	<.0001	1.05 (1.03-1.06)	<.0001	1.03 (1.02-1.05)	<.0001	1.04 (1.03-1.05)	<.0001
Age-diastolic BP interaction	1.00 (1.00-1.00)	<.0001	1.19 (1.07-1.33)	.0098	1.38 (1.23-1.54)	<.0001	1.00 (1.00-1.00)	<.0001
Diabetes	1.48 (1.20-1.81)	.0002	1.30 (1.06-1.60)	.0106	1.28 (1.02-1.61)	.0334	1.59 (1.30-1.93)	<.0001
No exercise	1.14 (0.99-1.30)	.0650	1.17 (1.03-1.32)	.0125	1.12 (0.99-1.27)	.0750	1.17 (1.05-1.31)	.0044

Abbreviations: AUC, area under the receiver operating characteristic curve; BP, blood pressure; CI, confidence interval.

Table III

Discrimination Statistics in Training Sample (N=7610) and Testing (N=3692) Sample

DATASET, COVARIATE	Area Under the Receiver Operating Characteristic Curve			
	3 Years	6 Years	9 Years	Ever
Training, continuous	0.739	0.755	0.800	0.782
Training, categorical	0.747	0.760	0.799	0.783
Testing, continuous	0.751	0.743	0.773	0.761
Testing, categorical	0.754	0.747	0.776	0.763

Table IV
 Three-, 6-, and 9-Year Risk of Incident Hypertension Based on Baseline Risk Score^a

Total Points	3-Year No.	3-Year Risk, %	6-Year No.	6-Year Risk, %	9-Year No.	9-Year Risk, %	Total Points	3-Year No.	3-Year Risk, %	6-Year No.	6-Year Risk, %	9-Year No.	9-Year Risk, %
0	97	3.58	5.44	7.89	12	482	23.67	40.21	61.03				
1	535	3.74	5.63	8.02	13	424	27.52	43.73	64.50				
2	806	4.14	6.52	9.21	14	322	31.17	47.93	67.74				
3	830	4.61	8.05	11.40	15	185	35.95	52.93	70.57				
4	905	5.15	9.50	14.05	16	84	36.39	53.36	72.59				
5	898	5.93	11.21	17.53	17	116	31.98	48.56	75.53				
6	873	7.34	14.41	22.71	18	193	33.51	50.67	76.77				
7	953	8.91	17.94	28.92	19	177	36.03	54.23	77.80				
8	936	11.10	22.29	36.18	20	128	40.22	59.11	79.86				
9	916	13.72	26.83	42.56	21	54	47.07	64.18	82.24				
10	761	16.81	31.73	49.64	22	17	52.53	65.97	83.23				
11	601	20.72	36.63	55.80	23	9	53.31	70.51	87.07				

^aScore can be computed from the algorithm in Figure 2 and risk was computed from the combined sample (training + testing sample, N=11,302).