Journal of Human Hypertension (2010), 1–10

www.nature.com/jhh

ORIGINAL ARTICLE

Prediction models for the risk of new-onset hypertension in ethnic Chinese in Taiwan

K-L Chien[1,2], H-C Hsu[2], T-C Su[2], W-T Chang[3], F-C Sung[4], M-F Chen[2] and Y-T Lee[2,4]
[1]*Institute of Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan;* [2]*Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan;* [3]*Department of Emergency Medicine, National Taiwan University Hospital, Taipei, Taiwan and* [4]*China Medical University Hospital, Taichung, Taiwan*

Prediction model for hypertension risk in Chinese is still lacking. We aimed to propose prediction models for new-onset hypertension for ethnic Chinese based on a prospective cohort design on community, which recruited 2506 individuals (50.8% women) who were not hypertensive at the baseline (1990–91). Total 1029 cases of new-onset hypertension developed during a median of 6.15 (interquartile range, 4.04–9.02) years of follow-up. In the clinical model, gender (2 points), age (8 points), body mass index (10 points), systolic blood pressure (19 points) and diastolic blood pressure (7 points) were assigned. The biochemical measures, including white blood count (3 points), fasting glucose (1 point), uric acid (3 points), additional to above clinical variables, were constructed. The areas under the receiver operative characteristic curves (AUCs) were 0.732 (95% confidence interval (CI), 0.712–0.752) for the point-based clinical model and 0.735 (95% CI, 0.715–0.755) for the point-based biochemical model. The coefficient-based models had a good performance (AUC, 0.737–0.741). The point-based clinical model had a similar net reclassification improvement as the coefficient-based clinical model ($P=0.30$), and had a higher improvement than the point-based biochemical model ($P=0.015$). We concluded that the point-based clinical model could be considered as the first step to identify high-risk populations for hypertension among Chinese.
Journal of Human Hypertension advance online publication, 8 July 2010; doi:10.1038/jhh.2010.63

Keywords: cohort study; prediction model; community screening

Introduction

Identifying individuals who are at high risk of hypertension will improve the efficiency of primary prevention strategies. Recent clinical trials have demonstrated that body weight control and lifestyle intervention in individuals with pre-hypertensive status can substantially delay hypertension development,[1] providing a rationale for the identification of high-risk individuals so as to implement early lifestyle intervention strategies to prevent hypertension. Routinely available and easily collected clinical information and lifestyle-related factors have been found to be effective in identifying hypertension risk in prevalent and incident cases.[2–8] However, the evidence on prediction models providing absolute risk for hypertension risk is relatively scanty and these prediction models have also been developed, primarily in Caucasians.[2,4–5] Moreover, previous studies based on hypertension prediction models were limited because of short follow-up periods,[3–4,9] an inability to incorporate laboratory data,[4] multiple biomarkers,[10] limited validation[11] and a lack of simple algorithm usage (Supplementary Table S1). Furthermore, the self-reporting of hypertension incidence may invalidate the accuracy of incidence rates.[4] Therefore, we constructed the prediction models for hypertension risk using a community-based cohort of middle-aged and elderly ethnic Chinese in Taiwan as the following strategies. First, we incorporated gender, age, body mass index (BMI), systolic (SBP) and diastolic blood pressures (DBP) as the clinical model and included white blood cell count, fasting glucose and uric acid[12–13] as the biochemical models. Second, we proposed two different scoring systems: regression coefficient-based scores[14] and point-based scores.[15] Finally, we tested the performance measures of these prediction models and compared the available models.

Materials and methods

Study design and study participants

Details of this cohort study have been previously published.[16–17] Briefly, the Chin-Shan Community Cardiovascular Cohort study began in 1990 by

Correspondence: Dr Y-T Lee, Department of Internal Medicine, National Taiwan University Hospital, Taipei 100, Taiwan.
E-mail: ytlee@ntu.edu.tw
Received 29 December 2009; revised 8 April 2010; accepted 24 April 2010

recruiting 1703 men and 1899 women of Chinese ethnicity aged 35 years old and above from Chin-Shan township. Information about anthropometry, lifestyle and medical conditions was assessed by the interview questionnaires in 2-year cycles for the initial 6 years and the validity and reproducibility of the collected data and measurements have been reported in detail elsewhere.[17] The response rate of the cohort participants was 85.7% at the end of the study.

BMI was calculated as weight (in kilograms)/height (in metres).[2] Family history was defined by first-degree hypertension. Blood pressure was measured twice in the right arm by a mercury sphygmomanometer with the subject seated comfortably and the arms supported and positioned at the level of the heart. The average of the blood pressure measurements was used as previously described.[18–19] Family history of hypertension was coded as the prevalent hypertension among first relatives.[2] Smoking habit was defined by current smoking status. Drinking history was defined as a binary variable using the frequency of drinking habits. Regular physical activity was coded as daily exercise habits.

Measurement of biochemical markers

The procedure for blood collection has been reported elsewhere.[20–21] Briefly, all venous blood samples drawn after a 12-h overnight fast were immediately refrigerated and transported within 6 h to the National Taiwan University Hospital. Serum samples were then stored at −70 °C before batch assay for levels of total cholesterol, triglycerides and high-density lipoprotein cholesterol. Standard enzymatic tests for serum cholesterol and triglycerides were used (Merck 14354 and 14366, Merck KGaA, Darmstadt, Germany). Glucose levels were measured in the supernatant by enzymatic assay (Merck 3389) in an Eppendorf 5060 autoanalyzer (Eppendorf Corp., Hamburg, Germany). The peripheral blood cell analysis was measured using a blood cell counter (Sysmex Cell Counter NE-8000, TOA Medical Electronics Co. Ltd, Kobe, Japan). Plasma uric acid concentrations were assayed with commercial kits (Merck Chem. Co, Darmstadt, Germany) placed in an Eppendorf 5060 autoanalyzer (Eppendorf Corp.).[22]

Follow-up strategy

We collected events and blood samples from the participants at baseline (1990–1991), the first follow-up period (1992–1993) and the fourth follow-up period (1997–1998).[16] We measured blood pressure and collected data on anti-hypertensive medication in the serial follow-up visit periods. Participants with baseline hypertension (defined by SBP or DBP $\geqslant$140/90 mm Hg or a history of anti-hypertensive medication use in the baseline period, 1990–91) were excluded from this investigation. A total of 2506 participants were included. We calculated the cumulative incidence rates of hypertension biennially in the first three periods (1992–93, 94–95 and 96–97) and the 2000–2001 period was the end of the study. The response rates in all periods were relatively high, from 86 to 96%.

Definition of hypertension and associated risk factors

We defined the incident hypertension categories according to the criteria established by the Seventh Joint National Committee. Normotensive was defined as SBP $<$120 mm Hg and DBP $<$80 mm Hg. Hypertension was defined as SBP $\geqslant$140 mm Hg or DBP $\geqslant$90 mm Hg, and individuals on anti-hypertensive medications in the follow-up periods were also included as incident cases. Individuals with a fasting blood sugar level $>$126 mg per 100 ml and/or use of oral hypoglycaemia agents or insulin injections were defined as diabetes mellitus.[15,23]

Statistical analysis

The basic clinical and biochemical measures were listed according to the status of developing hypertension or not. We used the multivariate Weibull model to construct the prediction models because the Weibull model is suitable for interval-censored data.[4] We specified the stepwise method for best subset selection, by choosing variables entering in or removing from the model using a significant level as 0.05. We forced gender into the biochemical model for completeness. We constructed the parsimonious model for predicting the risk of hypertension according to two categories of covariate. First, the clinical model included gender, age, BMI, SBP and DBP, which were obtained from questionnaires and physical examinations and were statistically associated with the risk of hypertension. Second, the biochemical model included white blood cell count, fasting glucose and uric acid additionally. The lifestyle factors, such as smoking, drinking alcohol and physical activity, were also tested but the likelihood ratio tests showed that adding these variables into the model did not improve prediction beyond the parsimonious models. Furthermore, adding family history of hypertension did not increase appreciably the prediction measures; therefore, we decided to exclude family history in the model. We still incorporated the family history in constructing the models from the John Hopkins[8] and Framingham cohorts[4] when comparing their performance.

We constructed coefficient-based[24] and point-based models[15,25] for predicting the risk of hypertension using the clinical and biochemical variables. With regard to the coefficient-based model, the risk scores were derived from the estimated coefficients and were calculated to absolute risk in the Weibull model.[4] For the point-based model, the absolute risk was summed by the derived point scores, which was

acquired by categorizing the covariates.[25] The details were described in the Supplementary Materials.

To enhance the comparability of our models with those from other studies, we compared the prediction models with available prediction models, including John Hopkins[8] and Framingham cohorts[4] (Supplementary Table S2), and tested the prediction performance using calibration and discrimination ability.

First, we assessed the goodness of fit for all models based on the Hosmer–Lemeshow test,[26] which was a calibration measure to calculate how close the predicted risks were to the actual observed risks,[27] and the results showed the calibration was good (Supplementary Figure S1), except John Hopkins model (Supplementary Table S3). Second, we compared the discrimination ability using the area under receiver operative characteristic curve (AUC). An AUC curve is a graph of sensitivity vs 1-specificity (or false-positive rate) for various cutoff definitions of a positive diagnostic test result.[28] Statistical differences in the AUCs were compared using the method of DeLong *et al*.[29] The AUC was a global summary measure for discrimination between individuals developing hypertension and those who did not.[30] We also conducted an internal validation of the simple points model and obtained a bias-corrected estimate of AUC using a fivefold cross-validation procedure,[31] and the overall performance by averaging the AUC estimates obtained from the five different partitions were similar to those of all data sets. Third, we compared the models by using the net reclassification improvement (NRI) and integrated discrimination improvement statistics.[32] The NRI statistic was based on the reclassification tables and was calculated from a sum of differences between the 'upward' movement in categories for event subjects and the 'downward' movement in those for non-event subjects. We presented the NRI according to the *a priori* risk categories of hypertension risk as 20, 40 and 60% risk. The integrated discrimination improvement can be interpreted as a difference between improvement in average sensitivity and any potential increase in average 'one minus specificity', and the statistic was a difference in Yates discrimination slopes between the new and old models. Finally, we plotted the Bland–Altman plot of the difference between point-based and coefficient-based risks vs the average of these two risks to compare the patterns between the point-based and coefficient-based risks.[33]

All statistical tests were two sided and *P*-values <0.05 were considered statistically significant. Analyses were performed with SAS version 9.1 (SAS Institute, Cary, NC, USA) and Stata version 9.1 (Stata Corporation, College Station, TX, USA).

Results

Patient characteristics

The mean and proportions of various clinical and biochemical measures are listed in Table 1. Compared with those who did not develop hypertension, participants with new-onset hypertension were likely to have a family history of hypertension, to be alcohol drinkers and older, and to have diabetes and a higher BMI, blood pressure, white blood cell count, fasting glucose and uric acid levels. A total of 2506 individuals (50.8% women) who were not hypertensive at the baseline (1990) were followed up and 1029 cases of new-onset hypertension developed during a median 6.15 (interquartile range, 4.04–9.02) years of period. Table 2 shows the parsimonious models using multivariate Weibull models for the prediction models.

Table 1 The means and proportions of various clinical and biochemical measures in the study participants according to the status of developing hypertension or not ($n=2506$)

Characteristic	*Unit*	*New-onset HT (−)* n = *1477*		*New-onset HT (+)* n = *1029*		P-*value*
Gender	Men	48.8		50.0		0.55
	Women	51.3		50.1		
Family history of hypertension	—	22.9		27.6		0.007
Smoking history	—	37.6		37.0		0.78
Drinking history	—	28.2		34.4		0.001
Regular physical activity habit	—	13.3		14.1		0.59
Type II diabetes history	—	8.3		13.8		<0.0001
Age	Year	51.5	12.1	54.0	11.7	<0.0001
Body mass index	$kg\,m^{-2}$	22.4	3.0	23.9	3.4	<0.0001
SBP	mm Hg	112.2	10.6	120.4	10.1	<0.0001
DBP	mm Hg	70.9	7.7	75.5	7.4	<0.0001
White blood cell count	$1000\,\mu l^{-1}$	6.1	1.6	6.4	1.8	<0.0001
Fasting glucose	mg per 100 ml	105.4	24.8	111.0	31.8	<0.0001
Uric acid	mg per 100 ml	5.3	1.6	5.7	1.6	<0.0001

Abbreviations: DBP, diastolic blood pressure; HT, hypertension; SBP, systolic blood pressure.

Table 2 Estimated coefficient, s.e., RR, 95% CI and significant levels in the clinical and biochemical models for the risk of hypertension, based on the Weibull regression model

Covariate	*Coefficient*	*s.e.*	*RR*	*95% CI*		P-*value*
Clinical model						
Sex, women vs men	0.124	0.038	0.88	0.82	0.95	0.001
Age, +1 years	−0.011	0.002	1.01	1.01	1.01	<0.0001
Body mass index ($kg\,m^{-2}$)	−0.043	0.006	1.04	1.03	1.06	<0.0001
SBP (mm Hg)	−0.029	0.002	1.03	1.02	1.03	<0.0001
DBP (mm Hg)	−0.014	0.003	1.01	1.01	1.02	<0.0001
Biochemical model						
Sex, women vs men	0.037	0.043	0.964	1.048	0.887	0.39
Age, +1 years	−0.010	0.002	1.010	1.014	1.007	<0.0001
Body mass index ($kg\,m^{-2}$)	−0.036	0.006	1.036	1.049	1.024	<0.0001
SBP (mmHg)	−0.028	0.002	1.029	1.034	1.024	<0.0001
DBP (mm Hg)	−0.013	0.003	1.013	1.019	1.007	<0.0001
White blood cell count ($1000\,\mu l^{-1}$)	−0.035	0.011	1.036	1.059	1.013	0.002
Fasting glucose (mg per 100 ml)	−0.001	0.001	1.001	1.003	1.000	0.030
Uric acid (mg per 100 ml)	−0.038	0.013	1.039	1.065	1.014	0.002

Abbreviations: CI, confidence interval; DBP, diastolic blood pressure; RR, relative risk; SBP, systolic blood pressure.
The Weibull regression uses an opposite metric to other proportional hazard models and results in opposite signs and interpretation of regression coefficients. The Weibull scale parameters are 0.592 and 0.589, and intercepts 8.173 and 8.604 for the clinical and biochemical models.

Point-based prediction models and performance measures

The clinical and biochemical point-based score chart to estimate the risk of hypertension is shown in Tables 3 and 4. In the clinical model, gender (2 points), age (8 points), BMI (10 points), SBP (19 points) and DBP (7 points) were assigned. The biochemical measures, including white blood count (3 points), fasting glucose (1 point) and uric acid (3 points), in addition to the above clinical variables, were constructed. These risk charts allowed a manual estimation of the annual and cumulative risk of developing hypertension for each individual, as shown in Tables 3 and 4 for 1-, 4-, 5- and 10-year predicted risk. By using the clinical point-based risk chart, we determined that 50% of the sample had a <20% risk, 33% had a 20–40% risk, 13% had a 20–60% risk and 4% had a >60% risk of incident hypertension during a 5-year follow-up interval. The AUCs were 0.732 (95% confidence interval (CI), 0.712–0.752) for the point-based clinical model and 0.735 (95% CI, 0.715–0.755) for the point-based biochemical model, indicating a good discrimination ability (Figure 1 and Supplementary Table S4). The AUC difference between the clinical and biochemical models was not significantly different ($P=0.17$). Comparing reclassification measures between the point-based and coefficient-based models (Table 5), we found that the point-based clinical model had a similar NRI as the coefficient-based clinical model (NRI, 2.0%, $P=0.30$), and had a higher improvement than the point-based biochemical model (NRI, 3.7%, $P=0.015$). Finally, the Bland–Altman plot showed that compared with the coefficient-based model, the point-based prediction model overestimated the clinical risk (estimated coefficient, 0.0579 ± 0.0037, $P<0.001$), yet underestimated the biochemical risk (estimated coefficient, -0.0430 ± 0.0044, $P<0.001$) (Figure 2).

Discussion

In the ethnic Chinese cohort data, we have constructed the point-based prediction models using clinical and biochemical measures. The clinical model, which contained age, gender, BMI, SBP and DBP and had a better prediction performance, was suggested for further application in mass screening for the risk of hypertension. The availability of the manual risk charts to predict future risk of hypertension, as has been the case for the prediction of coronary heart disease,[34] would improve the prediction of hypertension risk, identify high-risk populations and enhance preventive strategies.

Clinical risk factors

To our knowledge, this is the first hypertension prediction model specifically developed for an ethnic Chinese population. Several hypertension prediction models have previously been developed in various populations. Among 3202 Iranian diabetic patients with 2.9 years of follow-up,[3] gender, age at diagnosis of diabetes, BMI, fasting glucose and glycosated haemoglobin concentrations were associated with hypertension risk. Restriction to only type II diabetes patients might limit the generalizability to primary prevention in a general population. In the Framingham Heart Study cohort, Parikh *et al*.[4] proposed a prediction model on 1717 adult Caucasians without diabetes nor hypertension on a median 3.8-year follow-up period. The proposed prediction model included gender, age,

Table 3 The simple points system according to the clinical model and the total points (left) and predicted risk (%) (right) for hypertension in the study participants

Risk factor	*Categories*	*Points*
Sex	Men	2
	Women	0
Age (year)	35–39	0
	40–44	1
	45–49	2
	50–54	3
	55–59	4
	60–64	5
	65–69	6
	70–74	7
	⩾75	8
Body mass index (kg m^{-2})	<18	0
	18–19.9	2
	20–21.9	3
	22–23.9	5
	24–25.9	6
	26–27.9	8
	⩾28	10
SBP (mm Hg)	<105	0
	105–109	3
	110–114	5
	115–119	10
	120–124	11
	125–129	14
	130–134	16
	135–139	19
DBP (mm Hg)	<65	0
	65–69	2
	70–74	3
	75–79	4
	80–84	5
	85–89	7

Point total	*1-year risk (%)*	*4-year risk (%)*	*5-year risk (%)*	*10-year risk (%)*
0	0.3	3.0	4.4	13.4
1	0.3	3.3	4.8	14.6
2	0.4	3.6	5.2	15.9
3	0.4	3.9	5.7	17.2
4	0.4	4.3	6.2	18.7
5	0.5	4.7	6.8	20.3
6	0.5	5.1	7.4	22.0
7	0.6	5.6	8.1	23.8
8	0.6	6.1	8.8	25.7
9	0.7	6.7	9.6	27.8
10	0.7	7.3	10.4	29.9
11	0.8	7.9	11.4	32.3
12	0.9	8.7	12.4	34.7
13	0.9	9.4	13.5	37.3
14	1.0	10.3	14.6	40.0
15	1.1	11.2	15.9	42.8
16	1.2	12.2	17.3	45.8
17	1.4	13.3	18.8	48.8
18	1.5	14.4	20.3	52.0
19	1.6	15.7	22.0	55.2
20	1.8	17.0	23.8	58.5
21	1.9	18.5	25.8	61.8
22	2.1	20.1	27.8	65.1
23	2.3	21.7	30.0	68.4
24	2.5	23.5	32.4	71.7
25	2.8	25.4	34.8	74.9
26	3.0	27.5	37.4	77.9
27	3.3	29.7	40.1	80.9
28	3.6	32.0	43.0	83.7
29	4.0	34.4	45.9	86.2
30	4.3	37.0	49.0	88.6
31	4.7	39.6	52.1	90.7
32	5.2	42.5	55.3	92.6
33	5.6	45.4	58.6	94.2
34	6.2	48.4	61.9	95.6
35	6.7	51.6	65.3	96.7
36	7.3	54.8	68.6	97.6
37	8.0	58.0	71.8	98.3
38	8.7	61.4	75.0	98.9
39	9.5	64.7	78.1	99.3
40	10.4	68.0	81.0	99.5
41	11.3	71.3	83.8	99.7
42	12.3	74.5	86.3	99.8
43	13.4	77.6	88.7	99.9
44	14.5	80.5	90.8	100
45	15.8	83.3	92.6	100
46	17.2	85.9	94.3	100

Abbreviations: DBP, diastolic blood pressure; SBP, systolic blood pressure.

SBP, DBP, family history of hypertension, BMI and the interaction between age and DBP for the prediction model, with an AUC of 0.788 (95% CI, 0.733–0.803). Our clinical model was similar to this Framingham data; however, the interaction between age and DBP did not reach a significant level ($P = 0.74$). Accordingly, we did not include the interaction items of age and DBP into our prediction model. In addition, we examined the role of body weight change in the first 2 years, and found that the 2-year one-unit BMI change increased risk of hypertension by 10%.[35] These results are consistent with available evidence on weight change.[36] However, we did not include serial BMI change in the prediction model because serial measures increased the difficulty for the applicability of the prediction model.

Table 4 The simple points system according to the biochemical model (left) and predicted risk (%) (right) for hypertension in the study participants

Risk factor	*Categories*	*Points*
Sex	Men	1
	Women	0
Age (year)	35–39	0
	40–44	1
	45–49	2
	50–54	3
	55–59	4
	60–64	5
	65–69	6
	70–74	7
	⩾75	8
Body mass index (kg m^{-2})	<18	0
	18–19.9	1
	20–21.9	3
	22–23.9	4
	24–25.9	5
	26–27.9	7
	⩾28	9
SBP (mm Hg)	<105	0
	105–109	3
	110–114	5
	115–119	10
	120–124	11
	125–129	14
	130–134	16
	135–139	20
DBP (mm Hg)	<65	0
	65–69	2
	70–74	3
	75–79	4
	80–84	5
	85–89	7
WBC (1000 μl^{-1})	<5.1	0
	5.1–5.9	1
	6.0–7.0	1
	⩾7.1	3
Fasting glucose (mg per 100 ml)	<95	0
	95–101	0
	102–110	0
	⩾111	1
Uric acid (mg per 100 ml)	<4.4	0
	4.4–5.2	1
	5.3–6.4	2
	⩾6.5	3

Table 4 Continued

Point total	*1-year risk (%)*	*4-year risk (%)*	*5-year risk (%)*	*10-year risk (%)*
0	0.2	2.2	3.2	10.0
1	0.2	2.4	3.5	10.9
2	0.3	2.6	3.8	11.8
3	0.3	2.9	4.2	12.9
4	0.3	3.1	4.5	14.0
5	0.3	3.4	4.9	15.1
6	0.4	3.7	5.4	16.4
7	0.4	4.1	5.9	17.8
8	0.4	4.4	6.4	19.3
9	0.5	4.8	7.0	20.8
10	0.5	5.2	7.6	22.5
11	0.6	5.7	8.2	24.3
12	0.6	6.2	9.0	26.2
13	0.7	6.8	9.7	28.3
14	0.7	7.4	10.6	30.5
15	0.8	8.0	11.5	32.8
16	0.9	8.7	12.5	35.2
17	0.9	9.5	13.6	37.7
18	1.0	10.3	14.7	40.4
19	1.1	11.2	16.0	43.1
20	1.2	12.2	17.3	46.0
21	1.3	13.3	18.8	49.0
22	1.5	14.4	20.3	52.1
23	1.6	15.6	22.0	55.2
24	1.7	16.9	23.7	58.4
25	1.9	18.3	25.6	61.7
26	2.1	19.8	27.6	64.9
27	2.3	21.5	29.7	68.2
28	2.5	23.2	32.0	71.3
29	2.7	25.0	34.4	74.5
30	2.9	27.0	36.9	77.5
31	3.2	29.1	39.5	80.4
32	3.5	31.3	42.2	83.1
33	3.8	33.7	45.1	85.7
34	4.2	36.1	48.0	88.0
35	4.5	38.7	51.1	90.2
36	5.0	41.4	54.2	92.0
37	5.4	44.2	57.4	93.7
38	5.9	47.2	60.6	95.1
39	6.4	50.2	63.8	96.3
40	7.0	53.3	67.1	97.3
41	7.6	56.4	70.3	98.0
42	8.3	59.7	73.4	98.6
43	9.0	62.9	76.5	99.1
44	9.8	66.1	79.4	99.4
45	10.6	69.4	82.2	99.6
46	11.6	72.5	84.8	99.8
47	12.6	75.6	87.3	99.9
48	13.6	78.6	89.5	99.9
49	14.8	81.4	91.4	100
50	16.0	84.1	93.2	100
51	17.4	86.6	94.7	100
52	18.8	88.9	95.9	100

Abbreviations: DBP, diastolic blood pressure; SBP, systolic blood pressure; WBC, white blood cell.

The high incidence rate in our study participants may be attributed to the following reasons: First, we repeated blood pressure three times during the 6-year period, which would increase the incidence rates. Second, the 6-year incidence of hypertension in the US communities was 28 and 30% in African-American men and women aged 50–64 years, respectively,[37] indicating whites had a lower hypertension rate. In fact, the original Framingham Heart Study showed that the lifetime risk for developing hypertension were 90% in both 55- and 65-year-old participants who were free of hypertension at baseline during the 1976–1998,[38] indicating a high incidence rate in middle and elderly individuals. In addition, inclusion of diabetes cases may be one explanation for a high hypertension incidence. A 55-year-old men with 22 kg m^{-2} BMI and 120/80 mm Hg would developed an 81% probability of the 10-year hypertension incidence.

Our data showed that the association between DBP and the risk of hypertension did not change significantly by age group. The likelihood ratio test comparing the model with the interaction terms of age and DBP and without those terms did not reach significant level ($P = 0.18$). Our findings did not support previous evidence on the bimodal effect of age as an effect modifier for DBP on the risk of hypertension, as reported in previous studies.[4,39] The possible explanation was due to truncated age distribution and high incident hypertension rates in our study participants.

Biochemical risk factors

For biochemical measure, Wang *et al*.[9] demonstrated that some biomarkers were related to new-onset hypertension in the 1456 adults from a Framingham offspring cohort for 3 years of follow-up. Our biochemical models included white blood cell count, fasting glucose and uric acid, which may be obtained in a laboratory-based mass screening. Our findings did not support a family history of hypertension as a significant role for further developing hypertension in our population. In another study based on young John Hopkins medical students for >50 years of follow-up, family history of hypertension was associated with a multivariate 1.8- to 2.4-fold risk for hypertension.[2] The non-significant association in our population was partly attributed to the relatively older age of the participants. Although our prediction model did not include family history of hypertension, family history may be more valuable in risk prediction among younger adults than in older adults.[2,8] In addition, our data did not support plasma lipid levels, including total cholesterol, triglyceride, high-density lipoprotein and low-density lipoprotein cholesterol, as significant predictors for hypertension risk, in contrast to the findings from 16130 women for 10.8 years of follow-up, which showed that hyperlipidemia had a 1.34-fold risk of hypertension.[5] In this Women' Health study, lack of controlling baseline blood pressure and other biochemical variables, such as inflammatory markers and uric acid, may overestimate the role of lipids in the risk of hypertension, although apolipoprotein B, lipoprotein(a) and C-reactive protein were included.[10] Furthermore, the ethnic difference in the metabolic syndrome components should be taken into consideration. We also incorporated uric

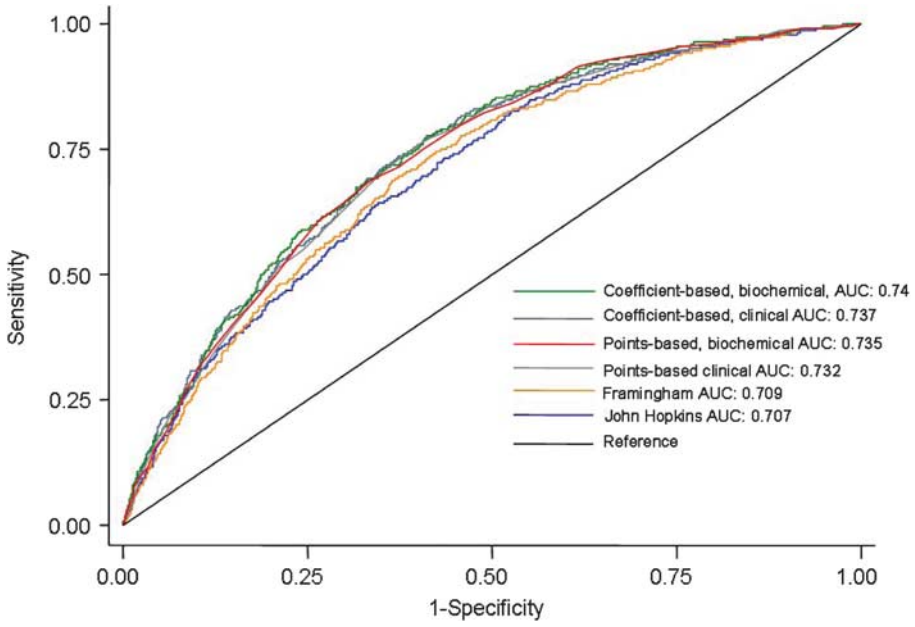

Figure 1 Receiver-operating characteristic curves for various models applied to the study population. Green, coefficient-based biochemical (AUC, 0.741); dark blue, coefficient-based, clinical (AUC, 0.737), red, point-based, biochemical (AUC, 0.735); grey, point-based, clinical (AUC, 0.732); yellow, Framingham (AUC, 0.709); blue, John Hopkins, (AUC, 0.707); black, reference.

Table 5 Summary of statistics comparing risk prediction algorithms to prediction based on the models

Model comparison	*NRI(%)*	*95% CI*		P-*value*	*IDI (%)*	*95% CI*		P-*value*
Point-based vs coefficient-based, clinical	2.0	−1.5	5.4	0.30	0.6	0.3	0.8	<0.0001
Clinical point-based vs biochemical point-based	7.0	3.7	10.3	0.0002	1.0	0.7	1.3	<0.0001
Coefficient-based vs point-based, biochemical	3.7	0.7	6.7	0.015	0.9	0.6	1.2	<0.0001
Biochemical coefficient-based vs clinical coefficient-based	−1.3	−4.2	1.7	0.40	0.5	0.2	0.8	0.001

Abbreviations: CI, confidence interval; IDI, integrated discrimination improvement; NRI, net reclassification improvement.
NRI with *a priori* 5-year cumulative risk categories according to <20, 20–40, 40–60 and ⩾60%.

acid concentrations in our biochemical model, because uric acid has been strongly associated with new-onset hypertension.[40] Among the 2062 men free of hypertension in the Normative Aging Study for a 21-year follow-up period, uric acid increased new-onset hypertension risk by 1.05-fold,[40] similar to our estimate in the multivariate model. We believe that most of the variables included in our models were feasible in the usual care practice in primary prevention settings. In addition, our data showed that as age at study initiation increased, the incidence rates of hypertension risk and the net 20/10 mm Hg blood pressure change increased significantly, especially for women, indicating that age at onset of observation initiation is a critical factor in evaluating change of blood pressure over time. Moreover, we stratified the incident hypertension into two subtypes: isolated systolic hypertension (SBP ⩾140 mm Hg and DBP <90 mm Hg) and diastolic hypertension (defined as isolated diastolic and more commonly mixed diastolic/systolic hypertension), and we found that the cumulative rates of isolated systolic hypertension increased significantly as age increased; however, diastolic hypertension rates decreased as age progressed. The largest percentage of new-onset hypertension among elderly participants was attributed to isolated systolic hypertension.

The lack of lifestyle risk factors in the prediction model in our study, compared with previous data,[41] may be explained by the following reasons. First, potential measurement misclassification in assessing lifestyle risk factors could induce measurement errors and non-differential misclassifications reducing the power for predicting hypertension risk. Second, attenuation of lifestyle factor effects was induced by mediating factors, such as blood pressure, obesity and biochemical markers. When included in the prediction models, these clinical and biochemical variables made lifestyle factors less significant for hypertension outcome. Although our prediction models did not include lifestyle factors, we still emphasize the importance of lifestyle factors for the risk of hypertension.[41] As the baseline distribution of lifestyle factors in the study sample was identified as having the highest hypertension

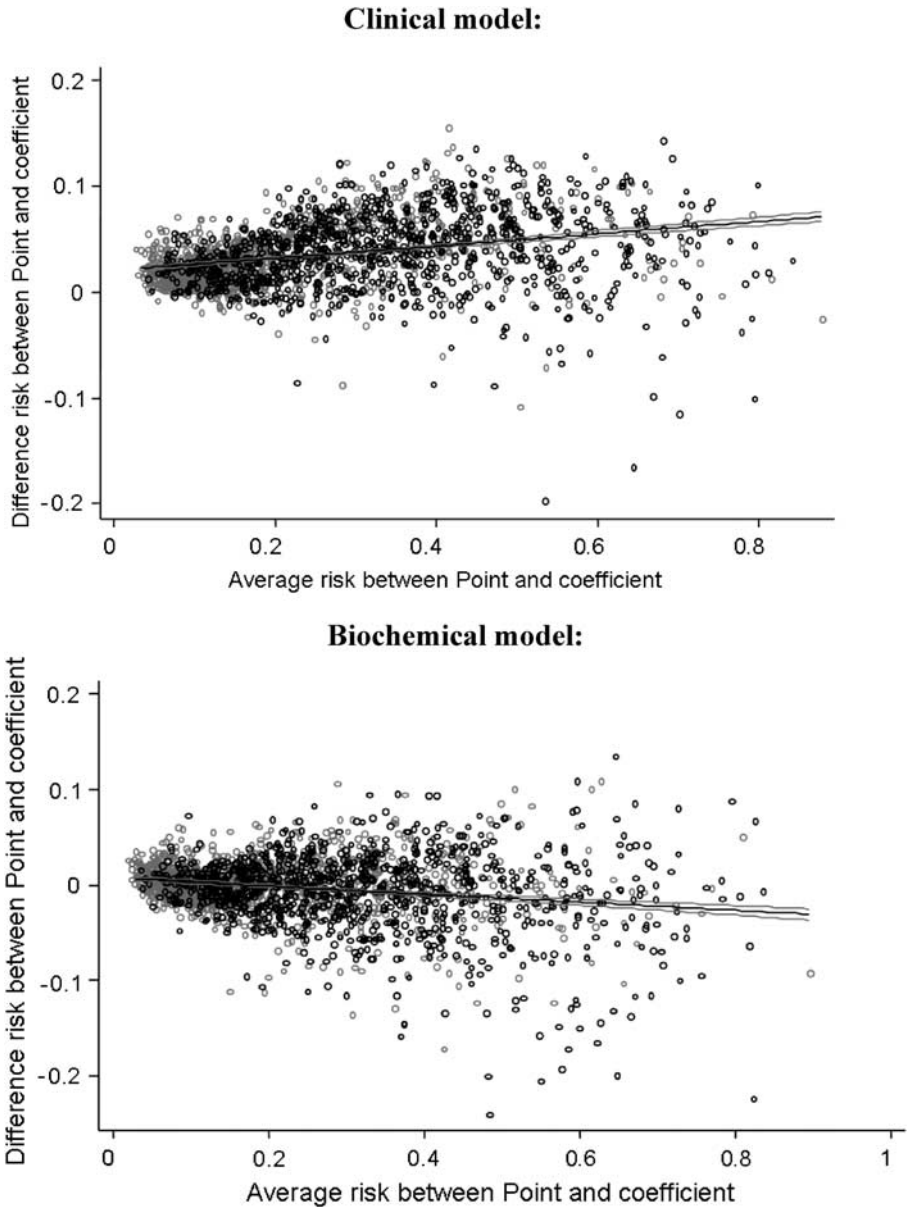

Figure 2 Bland–Altman plot of the difference between point-based and coefficient-based risks vs the average of these two risks in the clinical (upper) and biochemical (lower) models. (Black circle: subject developed hypertension; grey circle: subject did not develop hypertension.)

risk by the prediction models, the potential opportunities for lifestyle intervention during primary prevention should be identified early through prediction model screening.

Strengths and limitations

To our knowledge, this is the first hypertension prediction model specifically developed for an ethnic Chinese population. Owing to the large sample size, the estimates from our prediction models were found to be stable as demonstrated by the internal validation study. In addition, the use of a community-based population could reduce the possibility of selection bias. In addition, the constructed simple points system and predicted risk was still available among the participants without diabetes (Supplementary Tables S5 and S6) and cardiovascular disease at baseline. However, several potential limitations of this study should be mentioned. First, the point-based models were inferior to the coefficient-based model based on NRI and integrated discrimination improvement values, although the AUCs were similar. Second, we did not include extensive biomarker data in the model and the blood pressure ascertainment was performed once every 2 years. Third, we did not separate genders and did not include lifestyle factors in the prediction model. Lack of lifestyle factors in the model may decrease the application of

the prediction model in primary prevention. Finally, our study participants were middle and elderly Chinese population and the community is geographically unique, and so the external generalization to general population of our results was unknown. Furthermore, prevention strategies would be more effective on relatively young population. Therefore, further validation studies for these prediction models are warranted.

In conclusion, we have constructed the clinical and biochemical prediction models for predicting the 10-year incidence of hypertension among ethnic Chinese people. We recommend the point-based clinical model as the first step to identify high-risk populations for hypertension because of its simplicity and easily obtained measures in clinical practice and may be helpful to identify high-risk populations and improve prevention and treatment strategies for Chinese populations.

What is known about topic
- Hypertension incidence was associated with clinical and biochemical factors.
- Available prediction models were constructed in Caucasians.

What this study adds
- A simple point-based prediction model incorporating clinical measures was useful for Chinese.
- Gender, age, body mass index, SBP and DBP were assigned in this prediction model.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We thank the staff of the Department of Internal Medicine, National Taiwan University Hospital, and the participants of the CCCC study for their contributions. This study was partly supported by the National Science Council (NSC 97-2314-B-002-130-MY3, 97-3112-B-002-034-). This study was also supported in part by Taiwan Department of Health Clinical Trial Research Center of Excellence (DOH 99-TD-B-111-004).

References

1 Mancia G, De Backer G, Dominiczak A, Cifkova R, Fagard R, Germano G *et al.* 2007 Guidelines for the management of arterial hypertension: the Task Force for the Management of Arterial Hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). *Eur Heart J* 2007; **28**: 1462–1536.
2 Wang NY, Young JH, Meoni LA, Ford DE, Erlinger TP, Klag MJ. Blood pressure change and risk of hypertension associated with parental hypertension: the Johns Hopkins Precursors Study. *Arch Intern Med* 2008; **168**: 643–648.
3 Janghorbani M, Amini M. Hypertension in type 2 diabetes mellitus in Isfahan, Iran: incidence and risk factors. *Diabetes Res Clin Pract* 2005; **70**: 71–80.
4 Parikh NI, Pencina MJ, Wang TJ, Benjamin EJ, Lanier KJ, Levy D *et al.* A risk score for predicting near-term incidence of hypertension: the Framingham Heart Study. *Ann Intern Med* 2008; **148**: 102–110.
5 Sesso HD, Buring JE, Chown MJ, Ridker PM, Gaziano JM. A prospective study of plasma lipid levels and hypertension in women. *Arch Intern Med* 2005; **165**: 2420–2427.
6 Mellen PB, Bleyer AJ, Erlinger TP, Evans GW, Nieto FJ, Wagenknecht LE *et al.* Serum uric acid predicts incident hypertension in a biethnic cohort: the atherosclerosis risk in communities study. *Hypertension* 2006; **48**: 1037–1042.
7 Sesso HD, Buring JE, Rifai N, Blake GJ, Gaziano JM, Ridker PM. C-reactive protein and the risk of developing hypertension. *JAMA* 2003; **290**: 2945–2951.
8 Pearson TA, LaCroix AZ, , Mead LA, , Liang KY. The prediction of midlife coronary heart disease and hypertension in young adults: the Johns Hopkins multiple risk equations. *Am J Prevent Med* 1990; **6**: 23–28.
9 Wang TJ, Gona P, Larson MG, Levy D, Benjamin EJ, Tofler GH *et al.* Multiple biomarkers and the risk of incident hypertension. *Hypertension* 2007; **49**: 432–438.
10 Paynter NP, Cook NR, Everett BM, Sesso HD, Buring JE, Ridker PM. Prediction of incident hypertension risk in women with currently normal blood pressure. *Am J Med* 2009; **122**: 464–471.
11 Kivimaki M, Batty GD, Singh-Manoux A, Ferrie JE, Tabak AG, Jokela M *et al.* Validating the Framingham Hypertension Risk Score: results from the Whitehall II study. *Hypertension* 2009; **54**: 496–501.
12 Krishnan E, Kwoh CK, Schumacher HR, Kuller L. Hyperuricemia and incidence of hypertension among men without metabolic syndrome. *Hypertension* 2007; **49**: 298–303.
13 Sundstrom J, Sullivan L, D'Agostino RB, Levy D, Kannel WB, Vasan RS. Relations of serum uric acid to longitudinal blood pressure tracking and hypertension incidence. *Hypertension* 2005; **45**: 28–33.
14 Schulze MB, Shai I, Manson JE, Li T, Rifai N, Jiang R *et al.* Joint role of non-HDL cholesterol and glycated haemoglobin in predicting future coronary heart disease events among women with type 2 diabetes. *Diabetologia* 2004; **47**: 2129–2136.
15 Chien K, Cai T, Hsu H, Su T, Chang W, Chen M *et al.* A prediction model for type 2 diabetes risk among Chinese people. *Diabetologia* 2009; **52**: 443–450.
16 Chien KL, Hsu HC, Sung FC, Su TC, Chen MF, Lee YT. Incidence of hypertension and risk of cardiovascular events among ethnic Chinese: report from a community-based cohort study in Taiwan. *J Hypertens* 2007; **25**: 1355–1361.
17 Lee YT, Lin RS, Sung FC, Yang CY, Chien KL, Chen WJ *et al.* Chin-Shan Community Cardiovascular Cohort in Taiwan: baseline data and five-year follow-up morbidity and mortality. *Clin Epidemiol* 2000; **53**: 836–846.
18 Chien KL, Hsu HC, Su TC, Chen MF, Lee YT, Hu FB. Apolipoprotein B and non-high-density lipoprotein cholesterol and risk of coronary heart disease in Chinese. *J Lipid Res* 2007; **48**: 2499–2505.

19 Chien KL, Hsu HC, Chen PC, Su TC, Chang WT, Chen MF *et al.* Urinary sodium and potassium excretion and risk of hypertension in Chinese: report from a community-based cohort study in Taiwan. *J Hypertens* 2008; **26**: 1750–1756.
20 Chien KL, Lee YT, Sung FC, Su TC, Hsu HC, Lin RS. Lipoprotein (a) level in the population in Taiwan: relationship to sociodemographic and atherosclerotic risk factors. *Atherosclerosis* 1999; **143**: 267–273.
21 Chien KL, Hsu HC, Sung FC, Su TC, Chen MF, Lee YT. Hyperuricemia as a risk factor on cardiovascular events in Taiwan: the Chin-Shan Community Cardiovascular Cohort Study. *Atherosclerosis* 2005; **183**: 147–155.
22 Fossati P, Prencipe L, Berti G. Use of 3,5-dichloro-2-hydroxybenzenesulfonic acid/4-aminophenazone chromogenic system in direct enzymic assay of uric acid in serum and urine. *Clin Chem* 1980; **26**: 227–231.
23 Chien KL, Chen MF, Hsu HC, Chang WT, Su TC, Lee YT *et al.* plasma uric acid and the risk of type 2 diabetes in a Chinese community. *Clin Chem* 2008; **54**: 310–316.
24 Schulze MB, Hoffmann K, Boeing H, Linseisen J, Rohrmann S, Mohlig M *et al.* An accurate risk score based on anthropometric, dietary, and lifestyle factors to predict the development of type 2 diabetes. *Diabetes Care* 2007; **30**: 510–515.
25 Sullivan LM, Massaro JM, D'Agostino Sr RB. Presentation of multivariate data for clinical use: the Framingham Study risk score functions. *Stat Med* 2004; **23**: 1631–1660.
26 Hosmer Jr DW, Lemeshow S. The multiple logistic regression model. In: David W Hosmer Jr, Stanley Lemeshow (eds). *Applied Logistic Regression*. 1st edn. John Wiley & Sons: New York, 1989, pp 25–37.
27 McGeechan K, Macaskill P, Irwig L, Liew G, Wong TY. Assessing new biomarkers and predictive models for use in clinical practice: a clinician's guide. *Arch Intern Med* 2008; **168**: 2304–2310.
28 Hanley JA, McNeil BJ. A method of comparing the areas under receive operating characteristic curves derived from the same cases. *Radiology* 1983; **148**: 839–843.
29 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988; **44**: 837–845.
30 Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* 2007; **115**: 928–935.
31 Efron B, Tibshirani RJ. *An introduction to the Bootstrap*. Chapman & Hall/CRC: New York, 1994. p 456.
32 Pencina MJ, D' Agostino Sr RB, D' Agostino Jr RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008; **27**: 157–172.
33 Bland JM, Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. *Lancet* 1995; **346**: 1085–1087.
34 D'Agostino Sr RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM *et al.* General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation* 2008; **117**: 743–753.
35 Chen PC, Sung FC, Su TC, Chien KL, Hsu HC, Lee YT. Two-year change in body mass index and subsequent risk of hypertension among men and women in a Taiwan community. *J Hypertens* 2009; **27**: 1370–1376.
36 Czernichow S, Mennen L, Bertrais S, Preziosi P, Hercberg S, Oppert JM. Relationships between changes in weight and changes in cardiovascular risk factors in middle-aged French subjects: effect of dieting. *Int J Obes Relat Metab Disord* 2002; **26**: 1138–1143.
37 Apostolides AY, Cutter G, Daugherty SA, Detels R, Kraus J, Wassertheil-Smoller S *et al.* Three-year incidence of hypertension in thirteen US communities. On behalf of the Hypertension Detection and Follow-up Program cooperative group. *Prev Med* 1982; **11**: 487–499.
38 Vasan RS, Beiser A, Seshadri S, Larson MG, Kannel WB, D'Agostino RB *et al.* Residual lifetime risk for developing hypertension in middle-aged women and men: the Framingham Heart Study. *Jama* 2002; **287**: 1003–1010.
39 Franklin SS, Gustin IV W, Wong ND, Larson MG, Weber MA, Kannel WB *et al.* Hemodynamic patterns of age-related changes in blood pressure. The Framingham Heart Study. *Circulation* 1997; **96**: 308–315.
40 Perlstein TS, Gumieniak O, Williams GH, Sparrow D, Vokonas PS, Gaziano M *et al.* Uric acid and the development of hypertension: the normative aging study. *Hypertension* 2006; **48**: 1031–1036.
41 Forman JP, Stampfer MJ, Curhan GC. Diet and lifestyle risk factors associated with incident hypertension in women. *Jama* 2009; **302**: 401–411.

Supplementary Information accompanies the paper on the Journal of Human Hypertension website (http://www.nature.com/jhh)